Evolution of massive binary black holes

Qingjuan Yu[★]

Princeton University Observatory, Peyton Hall, Princeton, NJ 08544-1001, USA

ABSTRACT

Since many or most galaxies have central massive black holes (BHs), mergers of galaxies can form massive binary black holes (BBHs). In this paper we study the evolution of massive BBHs in realistic galaxy models, using a generalization of techniques used to study tidal disruption rates around massive BHs. The evolution of BBHs depends on BH mass ratio and host galaxy type. BBHs with very low mass ratios (say, $\lesssim 0.001$) are hardly ever formed by mergers of galaxies, because the dynamical friction time-scale is too long for the smaller BH to sink into the galactic centre within a Hubble time. BBHs with moderate mass ratios are most likely to form and survive in spherical or nearly spherical galaxies and in high-luminosity or high-dispersion galaxies; they are most likely to have merged in low-dispersion galaxies (line-of-sight velocity dispersion $\lesssim 90\,\mathrm{km\,s^{-1}}$) or in highly flattened or triaxial galaxies.

The semimajor axes and orbital periods of surviving BBHs are generally in the range 10^{-3}–$10\,\mathrm{pc}$ and 10–$10^5\,\mathrm{yr}$; they are also larger in high-dispersion galaxies than in low-dispersion galaxies, larger in nearly spherical galaxies than in highly flattened or triaxial galaxies, and larger for BBHs with equal masses than for BBHs with unequal masses. The orbital velocities of surviving BBHs are generally in the range 10^2–$10^4\,\mathrm{km\,s^{-1}}$. The methods of detecting surviving BBHs are also discussed.

If no evidence of BBHs is found in AGNs, this may be either because gas plays a major role in BBH orbital decay or because nuclear activity switches on soon after a galaxy merger, and ends before the smaller BH has had time to spiral to the centre of the galaxy.

Key words: black hole physics – galaxies: evolution – galaxies: interactions – galaxies: kinematics and dynamics – galaxies: nuclei.

1 INTRODUCTION

It is believed that many or most galaxies house massive black holes (BHs) at their centres (e.g. Magorrian et al. 1998). Mergers of galaxies, which are a central part of the galaxy formation process in the hierarchical structure formation picture, will inevitably form binary black holes (BBHs) if every galaxy has a central BH and the BH in-spiral time is less than a Hubble time. The questions of whether the binaries merge and how long they survive before the merger are relevant to resolving a number of problems in extragalactic astronomy, such as the detection of gravitational waves (e.g. Folkner 1998), the bending and precession of radio jets from active galactic nuclei (AGNs) (Begelman, Blandford & Rees 1980), the shapes of Fe Kα emission-line profiles from AGNs (Yu & Lu 2001), etc. On the other hand, so far there is not much observational evidence that BBHs exist, which might be expected if (i) the lifetime of the BBH is much shorter than the Hubble time, (ii) the hierarchical model of galaxy formation is incorrect, or (iii)

the BHs are efficiently ejected from the galaxy by three-body interactions, or other processes. The purpose of this paper is to study the lifetime of BBHs in realistic galaxy models.

Consider two massive BHs with masses $m_1 \geq m_2$ in the core of a galaxy. The merger of the BBH mainly involves four processes (Begelman et al. 1980). First, the dynamical friction stage: each BH sinks independently towards the centre of the common gravitational potential on the Chandrasekhar dynamical friction time-scale:

$$t_{\mathrm{df}} \sim \frac{4 \times 10^6}{\log N} \left(\frac{\sigma_{\mathrm{c}}}{200\,\mathrm{km\,s^{-1}}}\right) \left(\frac{r_{\mathrm{c}}}{100\,\mathrm{pc}}\right)^2 \left(\frac{10^8\,\mathrm{M_\odot}}{m_i}\right)\,\mathrm{yr} \tag{1}$$

$$(i = 1, 2),$$

where the core is assumed to have a one-dimensional velocity dispersion σ_{c} and a radius r_{c}, and to contain N stars. Second, the non-hard binary stage:[1] the two BHs eventually form a bound

[★]E-mail: yqj@astro.princeton.edu

---

[1] This stage is not called the 'soft' binary stage because, as Quinlan (1996) pointed out, the term 'soft' is best restricted to the familiar sense given by Heggie's law, 'soft binaries grow softer' (i.e., gain energy and semimajor axis) and a massive BBH in a galaxy core never grows softer.

binary system that continues to lose energy by dynamical friction. As the binary shrinks, the effectiveness of dynamical friction slowly declines with the increase of the BH velocity and the shortening of its orbital period, and the three-body interactions between the BBH and the stars passing in their vicinity gradually become the dominant factor that makes the BBH lose energy. Third is the hard binary stage: when

$$a \simeq a_h \equiv \frac{Gm_2}{4\sigma_c^2} = 2.8 \left(\frac{m_2}{10^8 \, \mathrm{M_\odot}}\right) \left(\frac{200 \, \mathrm{km \, s^{-1}}}{\sigma_c}\right)^2 \mathrm{pc} \qquad (2)$$

(G: gravitational constant), the BBH becomes hard (Quinlan 1996). Hard BBHs lose energy mainly by interacting with stars passing in their vicinity, most of which will be expelled from the BBH with an energy gain after one or more encounters with it. In such expulsions the average relative change in the BBH energy, $\Delta E/E$, is independent of E (Heggie 1975). The final stage is the gravitational radiation stage: as the BBH continues to harden, its semimajor axis decreases to the point at which gravitational radiation becomes the dominant dissipative force. A BBH on a circular orbit will then merge within the time (Peters 1964):

$$t_{\mathrm{merge}}(a) \sim \frac{5}{256} \frac{c^5 a^4}{G^3 \, \mu_{12} M_\bullet^2}$$

$$= 5.8 \times 10^6 \left(\frac{a}{0.01 \, \mathrm{pc}}\right)^4 \left(\frac{10^8 \, \mathrm{M_\odot}}{m_1}\right)^3 \frac{m_1^2}{m_2(m_1 + m_2)} \, \mathrm{yr}, \qquad (3)$$

where $\mu_{12} \equiv m_1 m_2/(m_1 + m_2)$ and $M_\bullet \equiv m_1 + m_2$ are the reduced mass and total mass of the binary, respectively, and $m_2 \leq m_1$. If $m_2/m_1 \ll 1$ (hence a_h in equation 2 is small), gravitational radiation can become the dominant dissipative force before the hard binary stage begins (see Fig. 2 or Section 5.3.4 below).

The largest uncertainty in the BBH lifetime comes during the second (non-hard binary) stage or the third (hard binary) stage, either of which can be the slowest of the above four processes. The time spent in the second or third stage may vary by several orders of magnitude, depending on whether or not the low-angular-momentum core stars that interact most strongly with the BBH are depleted before the BBH decays to the gravitational radiation stage.

In this paper we will estimate BBH evolution time-scales in realistic galaxy models obtained from observations of the central regions of nearby galaxies. To determine the BBH evolution time-scale at the slowest stage, the crucial thing is to get the BBH hardening time-scale associated with interactions with stars in the hard binary stage (see equation 9, which is longer than the hardening time-scale in the non-hard binary stage. The BBH hardening time-scale in the hard binary stage gives the slowest evolution time-scale or its upper limit, no matter whether the slowest stage is the hard binary stage or the non-hard binary stage. So our analysis starts from the beginning of the third of the above stages, when the BBH becomes hard. The theory we shall develop in Sections 2 and 3 has been applied to study tidal disruption rates around massive BHs, the evolution of galactic nuclei, etc., and was first developed to describe the evolution of globular clusters containing massive black holes (Lightman & Shapiro 1977; Cohn & Kulsrud 1978; Shapiro 1985). A simple analysis of the time-scales in the first two stages is given in Section 4. We will consider only purely stellar dynamical processes, ignoring the uncertain, but usually small, contribution from gas. We will also ignore the response of the core structure to the BBH evolution, and assume that the stellar system is in a fixed steady state. The justification for this assumption is that the total stellar mass removed from the galactic core during the BBH hardening is generally less than the core mass (see equations 12 and 45 or Figs 1 and 6 below). We present the results of the evolution of BBHs in realistic galaxy models in Section 5, and the estimated observational properties of surviving BBHs in Section 6. Finally, discussion and conclusions are given in Section 7.

2 LOSS REGION

In this section we introduce the concept of 'loss region' for stellar systems with BBHs; more specifically, we shall use the terms 'loss cone' for spherical systems and 'loss wedge' for axisymmetric systems.

2.1 Spherical systems

In a stellar system with a single stellar mass m_* and radius r_*, the distribution function (DF) $f(x, v)$ is defined so that $f(x, v) \, \mathrm{d}^3 x \, \mathrm{d}^3 v$ is the number of stars within a phase-space volume $\mathrm{d}^3 x \, \mathrm{d}^3 v$ of (x, v). We will consider the generalization to a distribution of masses and radii in Section 5.6. By Jeans's theorem, in a spherical system the DF depends on (x, v) only through the orbital binding energy per unit mass $\mathcal{E} = \psi(x) - \frac{1}{2}v^2$ and angular momentum per unit mass $J = |x \times v|$, where $\psi(r)$ is related to the gravitational potential $\Phi(r)$ through $\psi(r) \equiv -\Phi(r)$. If a BH has mass $M_\bullet$, stars that pass within a distance $r_t \sim (M_\bullet/m_*)^{1/3} r_*$ of the BH will be tidally disrupted if $r_t \gtrsim r_s \equiv 2GM_\bullet/c^2$ and swallowed whole if $r_t \lesssim r_s$, where c is the speed of light, and r_s is the Schwarzschild radius. The region in the (specific energy, specific angular momentum) phase-space where stars can be tidally disrupted at pericentre is called the 'loss cone' (cf. Frank & Rees 1976), and is given by

$$J^2 \leq J_{\mathrm{lc}}^2(\mathcal{E}) \equiv 2r_t^2 [\psi(r_t) - \mathcal{E}] \simeq 2GM_\bullet r_t \qquad (\mathcal{E} \ll GM_\bullet/r_t). \quad (4)$$

Now, consider a hard BBH with the masses $m_1 \geq m_2 \gg m_*$ and semimajor axis $a \lesssim a_h$ (see equation 2). The radius of the sphere of influence of the BH, a_H, is defined implicitly in terms of the intrinsic one-dimensional velocity dispersion of the galaxy $\sigma(r)$ through

$$\sigma^2(a_H) = GM_\bullet/a_H \equiv \sigma_H^2. \qquad (5)$$

Note that a_H is different from $a_h = (m_2/4M_\bullet)a_H$, which represents the semimajor axis at which the BBH becomes hard. Given a total BBH mass $M_\bullet$, a_h depends on the mass of the smaller BH m_2, while a_H does not. We set the core velocity dispersion $\sigma_c = \sigma_H$, since in the core the one-dimensional velocity dispersion $\sigma(r)$ generally varies only slowly outside a_H (e.g., the 'η' models of the density distribution for $1 \leq \eta \leq 3$ in Dehnen 1993 and Tremaine et al. 1994). At large radii ($r \gg a$), the BBH acts as a central point with mass $M_\bullet = m_1 + m_2$, and the stars move in the potential of the stars and the central mass point $M_\bullet$, with no systematic loss or gain of energy. The stars interacting most strongly with the BBH have low angular momenta with pericentres $\lesssim f_a a$ (f_a is a dimensionless factor ~ 1). They may either lose or gain energy after their first encounters with the hard BBH, but eventually most of the stars will be expelled from the potential of BBH with an energy gain (Quinlan 1996). As will be seen from equation (17) below, the energy gain of a star is usually large enough for it to

escape from the core when the BBH is hard, especially when $a \ll a_h$. By replacing r_t with $f_a a$ in equation (4), the loss cone in the BBH system is given as

$$J^2 \leq J_{lc}^2(\mathcal{E}, f_a a) \simeq 2GM_\bullet f_a a = 2G(m_1 + m_2)f_a a$$
$$(\mathcal{E} \ll GM_\bullet/f_a a);$$
(6)

thus J_{lc} will decrease as the BBH hardens. The BBH interacts most strongly with the core stars in the loss cone, and the stars escaping from the core with energy gain are considered as being removed from the loss cone.[2] The BBH energy loss rate is determined by the rate of removal of stars from the loss cone (the 'clearing rate'). At first, the depletion of the initial population of stars in the loss cone determines the clearing rate. As stars are removed from the loss cone, new stars are scattered into the loss cone by two-body relaxation (Binney & Tremaine 1987) and resonant relaxation (Rauch & Tremaine 1996; Rauch & Ingalls 1998), and eventually the clearing rate reaches a steady state controlled by the balance between the loss rate and the rate at which stars refill the loss cone. If the rms angular momentum transferred to or from the stars per orbital period is larger than J_{lc}, then the stars will refill the loss cone as fast as it is depleted and the loss cone remains full; otherwise, the stars will slowly diffuse into the loss cone and the loss cone remains nearly empty. Thus, at large radii the loss cone is full ('pinhole' regime), and at small radii the loss cone is nearly empty ('diffusion' regime) (the terms come from Lightman & Shapiro 1977). The transition radius or binding energy between the two regimes is denoted by r_{lc} or $\mathcal{E}_{lc}$ [$\psi(r_{lc}) \equiv \mathcal{E}_{lc}$]. If most of the stars contributing to the BBH orbital decay have energy $\mathcal{E} \ll GM_\bullet/f_a a$, which is generally a good approximation as will be seen from Section 5.3.4, the rate of refilling the loss cone in a BBH system can be determined from existing tidal disruption calculations, simply by replacing the tidal disruption radius r_t with $f_a a$. The refilling rate caused by two-body relaxation can be obtained by solving the steady-state Fokker–Planck equation (e.g., equation 17 in Magorrian & Tremaine 1999, hereafter MT). The loss cone in the BBH system can be approximated as static, even though the loss cone shrinks slowly as the BBH orbit decays, because the time-scale of the stellar diffusion into the loss cone is generally shorter than the BBH hardening time-scale (see Section 5.3.4). Resonant relaxation is effective only within a radius enclosing a mass $\sim 0.1M_\bullet$ of stars, and does not contribute significantly to the total refilling rate; therefore it will be neglected (see MT or Section 5.3.4 below).

2.2 Axisymmetric systems and triaxial systems

In an axisymmetric galaxy, the DF involves three integrals: $f(\mathcal{E}, J_z, J_3)$, where J_z is the component of angular momentum along the symmetry (z-)axis, and J_3 is a third integral. If the axisymmetric galaxy is nearly spherical, the third integral J_3 may be approximated by the angular momentum J (Binney & Tremaine 1987). In axisymmetric galaxies, there exist centrophilic orbits such as box orbits, which pass arbitrarily close to the centre and

have low angular momentum, as well as centrophobic orbits such as loop orbits, which avoid the centre and have high angular momentum. Here, we introduce J_s to mark the transition from centrophilic ($J \lesssim J_s$) to centrophobic ($J \gtrsim J_s$) orbits. Stars on centrophilic orbits with $|J_z| < J_{lc}$ can process into the loss cone, while those on centrophobic orbits with $J > J_{lc}$ cannot. Hence, in an axisymmetric galaxy, the loss cone is replaced by a loss wedge with $|J_z| < J_{lc}$ (MT). The BBH energy loss rate is determined initially by the clearing rate from the loss wedge, and in the steady state by the rates at which two-body relaxation and tidal forces refill the loss wedge.

Likewise, in a triaxial galaxy, we may expect that there also exists some characteristic angular momentum J_s which marks the transition from centrophilic to centrophobic orbits, and most of the stars with $J < J_s$ can process into the loss cone. We shall call the region $J < J_s$ the 'loss region' in triaxial galaxies. Thus the BBH energy loss rate is determined initially by the clearing rate from the loss region, and in the steady state by the rates at which two-body relaxation and tidal forces refill the loss region.

3 EVOLUTION TIME-SCALES AFTER BBHS BECOME HARD

The stars contributing to the orbital decay of a hard BBH come from the loss region. Assuming that $F(\mathcal{E}, a) \, d\mathcal{E}$ is the clearing rate from the loss region for stars with energy $\mathcal{E} \rightarrow \mathcal{E} + d\mathcal{E}$, the energy loss rate of the BBH is given by

$$\left|\frac{dE}{dt}(a)\right| = -m_* \int \Delta\mathcal{E} F(\mathcal{E}, a) \, d\mathcal{E},$$
(7)

where $\Delta\mathcal{E}$ is the average specific energy change of the stars escaping from the BBH during an interaction with the BBH, and the semimajor axis of the BBH is a. The quantity $\Delta\mathcal{E}$ is only weakly dependent on $\mathcal{E}$ for hard binaries. Hence we assume

$$\Delta\mathcal{E} = -KG\mu_{12}/a,$$
(8)

where μ_{12} is the reduced mass of the BBH, and K is a constant which will be determined from Quinlan (1996) in Section 3.1. Thus, after the BBH becomes hard, its hardening time due to stellar interactions is given as

$$t_h(a) = \left|\frac{a}{\dot{a}}\right| = \left|\frac{E}{\dot{E}}\right| = \frac{M_\bullet}{2Km_*} \frac{1}{\int F(\mathcal{E}, a) \, d\mathcal{E}}.$$
(9)

The time-scale associated with gravitational radiation is given by (Peters 1964)

$$t_{gr}(a) = \left|\frac{a}{\dot{a}}\right| = \frac{5}{64} \frac{c^5 a^4 (1-e^2)^{7/2}}{G^3 \mu_{12} M_\bullet^2 (1 + 73e^2/24 + 37e^4/96)},$$
(10)

where e is the BBH orbital eccentricity. The merger time-scale in equation (3) is related to the time-scale in equation (10) by $t_{merge}(a) = \int_0^a t_{gr}(a, e=0) \, da/a$.

As will be seen in Section 3.1, the hardening time-scale $t_h(a)$ of a hard BBH is independent of its orbital eccentricity e. The gravitational radiation time-scale in equation (10) depends weakly on e when e is small [e.g., $t_{gr}(a, e=0.3)/t_{gr}(a, e=0) \simeq 0.6$]; however, for large e, it decreases significantly [e.g., $t_{gr}(a, e=0.8)/t_{gr}(a, e=0) \simeq 0.01$]. In this paper we assume that the initial eccentricity when the BBH becomes hard is small (say, $e \lesssim 0.3$). In this case, in the hard binary stage, the BBH eccentricity hardly grows as the BBH hardens (Quinlan 1996); in the gravitational radiation stage,

---

[2] The assumption that stars escaping from the core are removed from the loss cone was also made by Begelman et al. (1980), Quinlan (1996) and others. This assumption is plausible since such stars have much longer periods and are much more susceptible to external torques than core stars. However, in some circumstances even a small fraction of stars returning to the loss cone could significantly enhance the BBH decay rate. Therefore this assumption deserves further investigation.

the eccentricity decays exponentially. So, in our calculations we will always set $e = 0$.

Combining equations (9) and (10), we find that the evolution time-scale of the hard BBH $t_{evol}(a)$ is given by

$$\frac{1}{t_{evol}(a)} = \frac{1}{t_h(a)} + \frac{1}{t_{gr}(a)}, \tag{11}$$

and gravitational radiation becomes dominant when $a \lesssim a_{gr}$, where a_{gr} is defined by $t_h(a_{gr}) = t_{gr}(a_{gr})$. This marks the transition between the hard binary and gravitational radiation stage. The total mass of the stars interacting with the BBH during the hard binary stage ($a_{gr} \leq a \leq a_h$), which can be derived from equation (9), is given by

$$M_*^{interact} = m_* \int \int F(\mathcal{E}, a) \, d\mathcal{E} dt = \frac{M_\bullet}{2K} \ln \left(\frac{a_h}{a_{gr}} \right). \tag{12}$$

3.1 K

Quinlan (1996) studied the dynamical evolution of massive BBHs by scattering experiments using the restricted three-body approximation. Consider a hard BBH in a galaxy core with uniform density ρ and Maxwellian velocity dispersion σ_c. The stars in the core start at a radius $r \gg a$ from the BBH, with velocity $\boldsymbol{v}$. Assume that v_r and v_t are the radial and tangential components of the vector $\boldsymbol{v}$. At large radii, the low-angular-momentum stars that contribute to the orbital decay of the hard BBH are in nearly radial motion, and the number flux into the loss cone is

$$\int F(\mathcal{E}, a) \, d\mathcal{E} = \int_0^\infty f(v_r, v_t = 0, r) 4\pi^2 J_{lc}^2(f_a a) |v_r| dv_r, \tag{13}$$

where

$$f(v_r, v_t = 0, r) = \frac{\rho}{m_*} \frac{1}{(2\pi\sigma_c^2)^{3/2}} \exp(-v_r^2/2\sigma_c^2). \tag{14}$$

With $J_{lc}^2(f_a a) \simeq 2GM_\bullet f_a a$ at large radii, we have from equation (9) the hardening time-scale of the BBH:

$$t_h(a) = \left| \frac{a}{\dot{a}} \right| = \frac{1}{4\sqrt{2\pi}Kf_a} \frac{\sigma_c}{Ga\rho}. \tag{15}$$

In Quinlan's simulation, the hardening time is given as

$$t_h = \left| \frac{a}{\dot{a}} \right| = \frac{\sigma_c}{G\rho aH} = 2.8 \times 10^7 \, \text{yr} \left(\frac{\sigma_c}{200 \, \text{km s}^{-1}} \right)$$
$$\times \left(\frac{10^3 \, M_\odot \, \text{pc}^{-3}}{\rho} \right) \left(\frac{0.1 \, \text{pc}}{a} \right) \left(\frac{16}{H} \right), \tag{16}$$

where for a hard BBH, H is a constant (~ 16) independent of the BH mass ratio and orbital eccentricity. Combining equations (15) and (16), we have the constant $Kf_a = 0.10H = 1.60$ (where $H = 16$). In our calculations, we will always set $K = 1.60$ and $f_a = 1$. Thus, combining equations (2) and (8), the average specific energy change is given by

$$\Delta\mathcal{E} = K \frac{G\mu_{12}}{a} \simeq 2 \left(\frac{K}{1.60} \right) \left(\frac{2m_1}{m_1 + m_2} \right) \left(\frac{a_h}{a} \right) \left(\frac{3\sigma_c^2}{2} \right), \tag{17}$$

which is generally large enough for the star to escape from the core when the BBH becomes hard, especially when $a \ll a_h$.

3.2 $F(\mathcal{E}, a)$

Most of this subsection is a summary and generalization of the tidal disruption rates calculated in MT.

3.2.1 *Spherical galaxies*

In a spherical galaxy with a DF $f(\mathcal{E}, J^2)$, when the BBH first becomes hard (the time t is set to 0), the total stellar mass in the loss cone is given by

$$M_{lc}(a_h) \simeq m_* \int 4\pi^2 \eta(\mathcal{E})f(\mathcal{E}, J^2 = 0)J_{lc}^2(\mathcal{E}, f_a a_h)P(\mathcal{E}) \, d\mathcal{E} \tag{18}$$

$$0 \leq \eta(\mathcal{E}) \leq 1,$$

where $P(\mathcal{E})$ is the radial period of an orbit with energy $\mathcal{E}$ and zero angular momentum, and $\eta(\mathcal{E})$ is a dimensionless factor. If $\eta(\mathcal{E}) = 1$ (or 0) for all $\mathcal{E}$, the loss cone is full (or empty). The clearing rate per unit energy at time t is at first controlled mainly by the draining rate from the loss cone:

$$F^{drain}(\mathcal{E}, a; t) \, d\mathcal{E}$$
$$\simeq \begin{cases} 4\pi^2 \eta(\mathcal{E})f(\mathcal{E}, J^2 = 0)J_{lc}^2(\mathcal{E}, f_a a) \, d\mathcal{E} & \text{for } t < P(\mathcal{E}) \\ 0 & \text{for } t > P(\mathcal{E}) \end{cases}. \tag{19}$$

After time $P(\mathcal{E})$, when the loss cone of stars with energy $\mathcal{E}$ is depleted, the loss cone is refilled by two-body relaxation. The steady-state diffusion rate of stars into the loss cone is given by the steady-state solution of the Fokker–Planck equation (13) in MT, which is a generalization of the equation in Cohn & Kulsrud (1978) to a non-Keplerian potential:

$$F^{lc}(\mathcal{E}, a) \, d\mathcal{E} = \frac{F^{max}(\mathcal{E}) \, d\mathcal{E}}{\ln R_0^{-1}(\mathcal{E}, f_a a)}, \tag{20}$$

where

$$F^{max}(\mathcal{E}) \equiv 4\pi^2 \bar{f}(\mathcal{E})J_c^2(\mathcal{E})\bar{\mu}(\mathcal{E})P(\mathcal{E}), \tag{21}$$

$$R_0(\mathcal{E}, f_a a) \equiv \frac{J_{lc}^2(\mathcal{E}, f_a a)}{J_c^2(\mathcal{E})}$$

$$\times \begin{cases} \exp[-q(\mathcal{E}, f_a a)] & \text{for } q(\mathcal{E}, f_a a) > 1 \\ \exp[-0.186q(\mathcal{E}, f_a a) - 0.824\sqrt{q(\mathcal{E}, f_a a)}] & \text{for } q(\mathcal{E}, f_a a) < 1 \end{cases}, \tag{22}$$

$$q(\mathcal{E}, f_a a) \equiv \bar{\mu}(\mathcal{E})P(\mathcal{E})J_c^2(\mathcal{E})/J_{lc}^2(\mathcal{E}, f_a a), \tag{23}$$

$$\bar{\mu}(\mathcal{E}) \equiv 2 \int_{r_-}^{r_+} \frac{\mu \, dr}{v_r}/P(\mathcal{E}), \qquad \mu \equiv \frac{2r^2\langle\Delta v_t^2\rangle}{J_c^2(\mathcal{E})}, \tag{24}$$

$J_c(\mathcal{E})$ is the specific angular momentum of a circular orbit at energy $\mathcal{E}$, $\bar{f}(\mathcal{E})$ is the 'isotropized' DF defined by $\bar{f}(\mathcal{E}) \equiv \int_0^{J_c^2} f(\mathcal{E}, J^2) \, dJ^2/J_c^2(\mathcal{E})$, r_+ and r_- are the apocentre and pericentre of loss-cone orbits, $\langle\Delta v_t^2\rangle$ is the diffusion coefficient for tangential velocity (see equation 8–64 of Binney & Tremaine 1987), and $\bar{\mu}(\mathcal{E})$ is the orbit-averaged diffusion coefficient. Note that the stellar encounters responsible for the velocity diffusion can be treated as elastic encounters only outside the 'collision radius'

$$r_{coll} \simeq 7 \times 10^{10}(M_\bullet/m_*)(r_*/R_\odot) \, \text{cm} \tag{25}$$

at which the velocity dispersion $\sim\sqrt{GM_\bullet/r}$ is comparable with the escape velocity from typical stars (Frank & Rees 1976). When $r \lesssim r_{coll}$, two stars cannot deflect each other's velocities through a large angle without coming so close that they actually collide. Generally, most stars have $R_0(\mathcal{E}, f_a a) \ll 1$, and the isotropized DF is close to the real stellar DF and is related to the stellar density $\rho(r)$

by

$$\rho(r) \approx 4\pi m_* \int \bar{f}(\mathcal{E})\sqrt{2[\Psi(r) - \mathcal{E}]}\,\mathrm{d}\mathcal{E}. \qquad (26)$$

The transition between pinhole ($\mathcal{E} < \mathcal{E}_{\mathrm{lc}}$) and diffusion ($\mathcal{E} > \mathcal{E}_{\mathrm{lc}}$) regimes is at $q(\mathcal{E}_{\mathrm{lc}}, f_a a) \simeq -\ln[J_{\mathrm{lc}}^2(\mathcal{E}_{\mathrm{lc}}, f_a a)/J_c^2(\mathcal{E})]$. From equations (20)–(24) it can be seen that in the diffusion regime the flux into the loss cone $F^{\mathrm{lc}}(\mathcal{E}, a)$ is insensitive to a, since it depends only logarithmically on a.

Considering both the draining and the refilling of the loss cone, at a certain time t, the clearing rate is given as

$$F(\mathcal{E}, a; t) = \max[F^{\mathrm{drain}}(\mathcal{E}, a; t), F^{\mathrm{lc}}(\mathcal{E}, a)]. \qquad (27)$$

3.2.2 Axisymmetric galaxies

In spherical systems, stars are all on centrophobic loop orbits. In axisymmetric galaxies, there also exist centrophilic orbits which can enhance the clearing rates. The dynamics of eccentric orbits in non-spherical potentials can be studied by a simple symplectic map constructed by Touma & Tremaine (1997). The mapping models the evolution of eccentric orbits as a two-step process: (i) precession of the orientation of the orbit in a non-spherical potential, and (ii) a kick to the angular momentum of the orbits at apocentre, where the star spends most of its time, and the torques are likely to be strongest. Following MT, we use this mapping to study the dynamics of centrophilic orbits with $J_z = 0$ in axisymmetric galaxies:

$$Y_n' = Y_n - \tfrac{1}{2}\epsilon \sin 2\theta_n$$
$$\theta_{n+1} = \theta_n + g(Y_n')$$
$$Y_{n+1} = Y_n' - \tfrac{1}{2}\epsilon \sin 2\theta_{n+1}, \qquad (28)$$

where $Y = J_{\pm}/J_c$, $J_{\pm}$ is the scalar angular momentum ($|J_{\pm}| = J$), which can be positive or negative, θ is the colatitude at each apocentre passage, ϵJ_c is the time-integral of the torque over one radial period, and $g(Y)$ is the advance in θ per radial period. The parameter ϵ is related to the axis ratio of the potential b and the mass ellipticity ϵ'; for example,

$$\epsilon = \sqrt{2\pi \mathrm{e}}(1 - b) = \sqrt{2\pi \mathrm{e}}\epsilon'/3 \qquad (b < 1) \qquad (29)$$

in the logarithmic potential (Touma & Tremaine 1997)

$$\Phi = \tfrac{1}{2}\log(x_1^2 + x_2^2/b^2). \qquad (30)$$

Of course, the potential of realistic galaxies with central BHs is not scale-free, and so the precession rate and the mapping in equation (28) are dependent on the energy $\mathcal{E}$. As shown in fig. 4 of MT, in realistic galaxies, the centrophilic orbits are regular at large $\mathcal{E}$ and stochastic below some critical energy $\mathcal{E}_{\mathrm{crit}}$. Above $\mathcal{E}_{\mathrm{crit}}$, the peak angular momenta of the regular orbits passing the points $(Y = 0, \theta = 0, \pi)$ in the Y–θ surfaces of section are denoted by $J_{\mathrm{m}}(\mathcal{E})$, and the area of (Y, θ) phase-space covered by the regular orbits with peak angular momenta less than J_{m} is $4J_{\mathrm{m}}(\mathcal{E})/J_c(\mathcal{E})$. Below $\mathcal{E}_{\mathrm{crit}}$, the areas of (Y, θ) phase-space occupied by the stochastic orbits are denoted by $2\pi J_l(\mathcal{E})/J_c(\mathcal{E})$. Those stars on centrophilic orbits with $|J_{\pm}| < J_{\mathrm{m}}$ (for regular orbits) or $|J_{\pm}| < J_l$ (for stochastic orbits) can process into the loss cone, so long as $|J_z| < J_{\mathrm{lc}}$. Here, we use one symbol $J_{\mathrm{s}}(\mathcal{E})$ to represent $2J_{\mathrm{m}}(\mathcal{E})/\pi$ for regular orbits or $J_l(\mathcal{E})$ for stochastic orbits.

If the DF of the axisymmetric galaxy is a function of two integrals $f(\mathcal{E}, J_z)$, when the BBH first becomes hard, the total mass

in the loss wedge is given by

$$M_{\mathrm{lw}}(a_{\mathrm{h}}) \simeq m_* \int 4\pi^2 \eta(\mathcal{E}) f(\mathcal{E}, J_z = 0)$$
$$\times \max[J_{\mathrm{lc}}^2(\mathcal{E}, f_a a_{\mathrm{h}}), 2J_{\mathrm{lc}}(\mathcal{E}, f_a a_{\mathrm{h}})J_{\mathrm{s}}(\mathcal{E})]P(\mathcal{E})\,\mathrm{d}\mathcal{E}$$
$$0 \le \eta(\mathcal{E}) \le 1, \qquad (31)$$

where, as in equation (18), the dimensionless factor $\eta(\mathcal{E}) = 1$ (or 0) for a full (or empty) loss wedge. If $J_{\mathrm{lc}}(\mathcal{E}, f_a a)$ is larger than $2J_{\mathrm{s}}(\mathcal{E})$, we use the clearing rates obtained in the spherical case. After $J_{\mathrm{lc}}(\mathcal{E}, f_a a)$ becomes less than $2J_{\mathrm{s}}(\mathcal{E})$ at a time given by $t \equiv T_{\mathrm{s}}(\mathcal{E})$ ($t = 0$ at $a = a_{\mathrm{h}}$), we will consider the effects of flattening as follows. First, the loss wedge in a two-integral model drains at a rate

$$F^{\mathrm{drain}}(\mathcal{E}, a; t)\,\mathrm{d}\mathcal{E}$$
$$\simeq \begin{cases} 4\pi^2 \eta(\mathcal{E}) f(\mathcal{E}, J_z = 0)J_{\mathrm{lc}}^2(f_a a)\,\mathrm{d}\mathcal{E} & \text{for } t < T^{\mathrm{drain}}(\mathcal{E}) \\ 0 & \text{for } t > T^{\mathrm{drain}}(\mathcal{E}), \end{cases} \qquad (32)$$

where $T^{\mathrm{drain}}(\mathcal{E})$ satisfies

$$\int_0^{\min[P(\mathcal{E}), T_{\mathrm{s}}(\mathcal{E})]} \mathrm{d}t'(a') \frac{2J_{\mathrm{lc}}(f_a a)J_{\mathrm{s}}(\mathcal{E})}{J_{\mathrm{lc}}^2(f_a a')} F^{\mathrm{drain}}(\mathcal{E}, a'; t')\,\mathrm{d}\mathcal{E}$$
$$+ \int_{\min[P(\mathcal{E}), T_{\mathrm{s}}(\mathcal{E})]}^{T^{\mathrm{drain}}(\mathcal{E})} \mathrm{d}t'(a') \frac{J_{\mathrm{lc}}(f_a a)J_{\mathrm{s}}(\mathcal{E})}{J_{\mathrm{lc}}(f_a a')J_{\mathrm{s}}(\mathcal{E})} F^{\mathrm{drain}}(\mathcal{E}, a'; t')\,\mathrm{d}\mathcal{E}$$
$$\simeq 8\pi^2 \eta(\mathcal{E}) f(\mathcal{E}, J_z = 0)J_{\mathrm{lc}}(f_a a)J_{\mathrm{s}}(\mathcal{E})P(\mathcal{E})\,\mathrm{d}\mathcal{E}. \qquad (33)$$

In the above equation, the right-hand-side term represents the total number of stars (with energy $\mathcal{E} \to \mathcal{E} + \mathrm{d}\mathcal{E}$) which can process into the loss cone $J < J_{\mathrm{lc}}(f_a a)$ by tidal forces. If $T_{\mathrm{s}}(\mathcal{E}) < P(\mathcal{E})$, the left-hand-side terms in equation (33) approximately represent the numbers of stars (with energy $\mathcal{E} \to \mathcal{E} + \mathrm{d}\mathcal{E}$) removed from the loss wedge before and after $J_{\mathrm{lc}}(\mathcal{E}, f_a a) = 2J_{\mathrm{s}}(\mathcal{E})$; also if $T_{\mathrm{s}}(\mathcal{E}) > P(\mathcal{E})$, we have $T^{\mathrm{drain}}(\mathcal{E}) \simeq P(\mathcal{E})$.

After the loss wedge is depleted, the rate at which the loss wedge is refilled is given by a Fokker–Planck analysis in MT:

$$F^{\mathrm{lw}}(\mathcal{E}, a)\,\mathrm{d}\mathcal{E} = 4\pi^2 J_{\mathrm{lc}}^2(\mathcal{E}, f_a a) \frac{\bar{f}(\mathcal{E})\,\mathrm{d}\mathcal{E}}{1 + J_c/(4q_z J_{\mathrm{s}})}, \qquad (34)$$

where $q_z = q/8$ (see q in equation 23), and $\bar{f}(\mathcal{E})$ is the 'isotropized' DF defined by $\bar{f}(\mathcal{E}) \equiv \int_0^{J_c} f(\mathcal{E}, J_z)\,\mathrm{d}J_z/J_c(\mathcal{E})$. If the galaxy is close to spherical, the isotropized DF $\bar{f}(\mathcal{E})$ is close to the real stellar DF and $\bar{f}(\mathcal{E})$ is also related to the stellar density by equation (26). The transition between the 'diffusion' and 'pinhole' regions occurs at $q \simeq 2J_c/J_{\mathrm{s}}$. From equations (23) and (34) it can be seen that in the diffusion regime the flux into the loss wedge $F^{\mathrm{lw}}(\mathcal{E}, a)$ is independent of a.

Considering both the draining and refilling of the loss wedge or the loss cone, at a certain time t, the clearing rate is given as

$$F(\mathcal{E}, a; t) = \max[F^{\mathrm{drain}}(\mathcal{E}, a; t), F^{\mathrm{lw}}(\mathcal{E}, a), F^{\mathrm{lc}}(\mathcal{E}, a)]. \qquad (35)$$

3.2.3 Triaxial galaxies

Similarly, in triaxial galaxies, there will be some characteristic angular momentum $J_{\mathrm{s}}(\mathcal{E})$ inside which most orbits are centrophilic and the BBH decay rate is determined by the clearing rates from the loss regions $J(\mathcal{E}) < J_{\mathrm{s}}(\mathcal{E})$. Here, we approximate the $J_{\mathrm{s}}(\mathcal{E})$ obtained in the axisymmetric cases as the characteristic angular momentum $J_{\mathrm{s}}(\mathcal{E})$ in triaxial galaxies. Merritt & Quinlan (1998)

argue that the triaxiality of galaxies with central BHs decays due to the evolution of stochastic orbits, so that most galaxies become axisymmetric at any given radius within a time $t \equiv T^{\text{trans}} \sim 10^2$ local orbital periods. When the BBH first becomes hard, we set the time $t = 0$, and the age of the galactic triaxiality is defined as $T^{\text{tri}}_{\text{age}}$ at this time. We give the mass in the loss region at $a = a_{\text{h}}$ as

$$M_{\text{lr}}(a_{\text{h}}) \simeq m_* \int 4\pi^2 \eta(\mathcal{E}) f(\mathcal{E}) \max[J^2_{\text{lc}}(\mathcal{E}, f_a a_{\text{h}}), J^2_s(\mathcal{E})] P(\mathcal{E}) \, d\mathcal{E}$$

$$0 \leq \eta(\mathcal{E}) \leq 1. \tag{36}$$

When $0 \leq t \leq T_{\text{tri}}(\mathcal{E}) \equiv \max[T^{\text{trans}} - T^{\text{tri}}_{\text{age}}, 0]$, we consider the clearing rates in triaxial galaxies as follows. If $J_s(\mathcal{E}) < J_{\text{lc}}(\mathcal{E})$, the clearing rates follow the same formula as in the spherical case (see Section 3.2.1). When $J_s(\mathcal{E})$ becomes larger than $J_{\text{lc}}(\mathcal{E})$ at a time given by $t \equiv T_s(\mathcal{E})$, the clearing rates from the loss region are first controlled by the draining rates from the loss region:

$$F^{\text{drain}}(\mathcal{E}, a; t) \, d\mathcal{E}$$

$$\simeq \begin{cases} 4\pi^2 \eta(\mathcal{E}) f(\mathcal{E}) J^2_{\text{lc}}(f_a a) \, d\mathcal{E} & \text{for } t < \min[T^{\text{drain}}(\mathcal{E}), T_{\text{tri}}(\mathcal{E})] \\ 0 & \text{for } t > \min[T^{\text{drain}}(\mathcal{E}), T_{\text{tri}}(\mathcal{E})] \end{cases},$$

$$\tag{37}$$

where T^{drain} satisfies

$$\int_0^{\min[P(\mathcal{E}), T_s(\mathcal{E})]} dt'(a') \frac{J^2_s(f_a a)}{J^2_{\text{lc}}(f_a a')} F^{\text{drain}}(\mathcal{E}, a'; t') \, d\mathcal{E}$$

$$+ \int_{\min[P(\mathcal{E}), T_s(\mathcal{E})]}^{T^{\text{drain}}} dt'(a') F^{\text{drain}}(\mathcal{E}, a'; t') \, d\mathcal{E}$$

$$\simeq 4\pi^2 \eta(\mathcal{E}) f(\mathcal{E}) J^2_s(\mathcal{E}) P(\mathcal{E}) \, d\mathcal{E}. \tag{38}$$

Equation (38) has similar meanings to equation (33).

After the loss region is depleted, the refilling of the loss region can be described by equation (20) or (34). The clearing rate is given as equation (35).

When $t \geq T_{\text{tri}}(\mathcal{E})$, we will assume axisymmetry and follow the calculation in Section 3.2.2.

4 EVOLUTION TIME-SCALES BEFORE BBHS BECOME HARD

In this section we carry out a simple analysis of the evolution time-scales before the BBH becomes hard. Assume that a BH of mass m_1 is located at the centre of a spherical galaxy with stellar mass density $\rho(r)$ and one-dimensional velocity dispersion $\sigma(r)$. Suppose that a BH m_2 ($m_2 \leq m_1$) that used to be at the centre of a galaxy is orbiting with a velocity $\boldsymbol{v}_2$ and spiralling into the centre of the parent galaxy by dynamical friction. Because the in-spiralling BH m_2 is accompanied by the stars bound to it with total mass $M^{m_2}_*$, dynamical friction brings the two BHs together much more rapidly than if m_2 is 'naked' (e.g. Milosavljević & Merritt 2001). The dynamical friction force on m_2 and its accompanying stars $M^{m_2}_*$ is given by (Binney & Tremaine 1987)

$$(m_2 + M^{m_2}_*) \left(\frac{d\boldsymbol{v}_2}{dt}\right)_{\text{df}} = -4\pi \ln\Lambda G^2 (m_2 + M^{m_2}_*)^2 \rho(r)$$

$$\times \frac{\boldsymbol{v}_2}{v^3_2} \left[\text{erf}(X) - \frac{2X}{\sqrt{\pi}} e^{-X^2}\right], \tag{39}$$

where $X = v_2/\sqrt{2}\sigma(r)$, erf is the error function, and m_2 and its bound stars $M^{m_2}_*$ are assumed to act as a single body. The logarithm of the ratio of the maximum and the minimum impact parameter $\ln\Lambda$ is set to unity (which overestimates the dynamical friction time-scale, but not by more than a small logarithmic factor). Assuming that the orbit of m_2 is circular, the force (equation 39) is tangential and causes m_2 to lose angular momentum per unit mass L at a rate

$$\frac{dL(r)}{dt} = r\left(\frac{dv_2}{dt}\right)_{\text{df}}, \tag{40}$$

where

$$L(r) = rv_2 = \sqrt{G[m_1 + M_*(r)]r}, \tag{41}$$

and $M_*(r)$ is the stellar mass in the parent galaxy within a radius r. The dynamical friction time-scale is given by

$$t_{\text{df}} = \left|\frac{r}{\dot{r}}\right| = r\left|\frac{dL/dr}{dL/dt}\right|, \tag{42}$$

which can be calculated from equations (39)–(41).

In the potential of the parent galaxy with BH m_1, the stars around the BH m_2 are tidally truncated at the radius

$$r_{\text{t}} \simeq \left[\frac{G(m_2 + M^{m_2}_*)}{4\Omega^2 - \kappa^2}\right]^{1/3} \simeq \left(\frac{GM^{m_2}_*}{4\Omega^2 - \kappa^2}\right)^{1/3} \quad M^{m_2}_* \gg m_2, \tag{43}$$

where Ω and κ are the circular frequency and epicycle frequency in the parent galaxy. Note that the tidal forces approach zero in a nearly homogeneous core, in which $\kappa = 2\Omega$. We use this formula to obtain a crude estimate of the ratio of the bound stellar mass $M^{m_2}_*$ to the BH mass m_2. If the distribution of the stars around each of the BHs is a singular isothermal sphere and their one-dimensional velocity dispersions are σ_1 and σ_2, we have $\Omega^2 = 2\sigma^2_1/r^2$, $\kappa^2 = 4\sigma^2_1/r^2$, $M^{m_2}_* = 2\sigma^2_2 r_{\text{t}}/G$ and $r_{\text{t}} = (\sigma_2/\sqrt{2}\sigma_1)r$. If the velocity dispersions σ_1 and σ_2 follow the same relation with central BH masses m_1 and m_2 as equation (50) (see Section 5.1), the ratio of the bound stellar mass $M^{m_2}_*$ to the BH mass m_2 is given by

$$\frac{M^{m_2}_*}{m_2} = \frac{\sqrt{2}r\sigma^3_2}{Gm_2\sigma_1} = 300 \left(\frac{r}{2\,\text{kpc}}\right) \left(\frac{m_2}{m_1}\right)^{0.25} \left(\frac{10^8\,\text{M}_\odot}{m_1}\right)^{2.14}$$

$$M^{m_2}_* \gg m_2. \tag{44}$$

The stellar mass bound to m_2 is effective in speeding the in-spiralling of m_2 only when m_2 is at large radii (e.g., $\gtrsim 10-100\,\text{pc}$ in Fig. 3 below).

When m_2 sinks to the radius at which $M_*(r) \approx M_\bullet$ (e.g., $r \sim 1-10a_{\text{H}}$ in Fig. 5 below), the two BHs form a 'bound' binary system (this marks the transition from the dynamical friction stage to the non-hard binary stage). (Here, the definition of 'bound' is somewhat arbitrary; we only intend to mean that the gravitational force on m_2 is dominated by m_1 rather than by the stars.) Although the BBH orbital decay in the non-hard binary stage comes from both dynamical friction from distant stars and three-body interactions between the BBH and the stars passing the BBH vicinity, we still use equation (42) to estimate the BBH evolution time-scales in both the dynamical friction stage and the non-hard binary stage because scattering experiments with the restricted three-body approximation basically give a similar hardening time-scale to equation (42) (see equations 2, 16 and 18 in Quinlan 1996).

Before the BBH becomes bound [$M_*(r) \gtrsim M_\bullet$], the removed

mass from the galactic core can be ignored, since both the mass and binding energy of the smaller BH are less than the total mass and kinetic energy of the core stars interior to the BBH orbit; thus the in-spiral has negligible effect on the dynamics of the core stars. However, during the non-hard binary stage, depletion of the stars that interact with the BBH can be significant in some galaxies. The mass removing rate from the galactic core caused by interactions between stars and one or both of the BHs, $J_{rm}(a)$, is defined by

$$J_{rm}(a) \equiv \frac{1}{M_\bullet} \frac{dM_{rm}}{d\ln(1/a)}, \qquad (45)$$

where M_{rm} is the stellar mass removed from the galactic core. In Section 2 we introduced the concept of loss region (loss cone for spherical galaxies and loss wedge for axisymmetric galaxies) for a hard BBH, but equations (18), (31) and (36) (i.e., the total stellar mass in the loss region) can be generalized to the non-hard binary stage to reflect the mass of the stars which may have close encounters with the BBH. Here, the stars which may have close encounters with the BBH can be either the original low-angular-momentum stellar population or those precessed from high-angular-momentum orbits by tidal forces. Before the BBH becomes hard, if the mass of the initial stellar population in the loss region is not enough to provide the removed stellar mass from the galactic core M_{rm} (see equation 45), the BBH would lose energy mainly by interactions with distant stars, which becomes more and more inefficient as the BBH hardens, and by three-body interactions with the low-angular-momentum stars which are diffused from high-angular-momentum orbits by two-body relaxation, which will dominate the BBH hardening time-scales in the hard binary stage (see equations 9, 20 and 34). Thus the BBH hardening time-scale should be higher than those estimated from equations (39)–(42) in the non-hard binary stage, and smoothly increase to connect the hardening time-scale in the hard binary stage. Here, we use the linear approximation [for the $\ln(t_h)-\ln(a)$ relation] to obtain the BBH hardening time-scale at $a_h \leq a \leq a_{dp}$:

$$\ln[t_h(a)] = \ln[t_h(a_h)]$$
$$+ \ln[t_{df}(a_{dp})/t_h(a_h)] \ln(a/a_h)/\ln(a_{dp}/a_h), \qquad (46)$$

where a_{dp} is the semimajor axis where the initial population of stars in the loss region are all removed from the loss region (which will be determined in Section 5.2), $t_h(a_h)$ is obtained from Section 3, and $t_{df}(a)$ is obtained from equations (39)–(42). Considering the possible BBH energy loss caused by gravitational radiation, the BBH evolution time-scale $t_{evol}(a)$ is given as equation (11) and $t_h(a) = t_{gr}(a)$ may happen before the BBH becomes hard.

5 RESULTS

5.1 Galaxy properties

In the past several years, images from the *Hubble Space Telescope* (*HST*) have revealed many details about the central regions of nearby galaxies, with a resolution of 0.1 arcsec, corresponding to distances of several pc or $10^5-10^6(M_\bullet/10^8 \, M_\odot)$ Schwarzschild radii for typical target galaxies (Byun et al. 1996). The inner surface brightness profiles of the galaxies are well fitted with a five-parameter fitting function – the Nuker law:

$$I(r) = 2^{\frac{\beta-\gamma}{\alpha}} I_b \left(\frac{r}{r_b}\right)^{-\gamma} \left[1 + \left(\frac{r}{r_b}\right)^\alpha\right]^{-\frac{\beta-\gamma}{\alpha}}. \qquad (47)$$

The asymptotic logarithmic slope inside r_b is $-\gamma$, the asymptotic

outer slope is $-\beta$, and the parameter α parametrizes the sharpness of the break. The break radius r_b is the point of maximum curvature in log–log coordinates. The 'break surface brightness' I_b is the surface brightness at r_b. Elliptical galaxies and spiral bulges (hot galaxies) can be classified into two main types according to their inner surface brightness profiles ($I \propto r^{-\gamma}$ when $r \to 0$). Hot galaxies with a quite shallow inner power-law slope ($\gamma \lesssim 0.3$) are classified as 'core' galaxies, and those with a steep slope ($\gamma \gtrsim 0.5$) are labelled as 'power-law' galaxies. The difference between core and power-law profiles has no direct connection with the presence of a disc, whether seen edge-on or face-on, although power-law galaxies have discy isophotes (Faber et al. 1997). Core galaxies are luminous ($M_V < -20.5$) with large central BHs, and power-law galaxies are faint ($M_V > -22$) with small central BHs. Early discussions of the demography of the central BHs focused on the correlation of BH mass with galaxy luminosity. It is now believed that BH mass is correlated more tightly with galactic velocity dispersion, which suggests a causal connection between the formation and evolution of the BH and the bulge (Ferrarese & Merritt 2000; Gebhardt et al. 2000).

We shall apply the theoretical results of the preceding sections to estimate what the evolution time-scales of the BHs in the centres of nearby galaxies would be if they were binary.

We use the *HST* sample of hot galaxies compiled in the paper by Faber et al. (1997). Among the 61 galaxies in that paper, 41 galaxies (Table 1) have values for the Nuker-law parameters, the stellar mass-to-light ratio Υ (constant for each galaxy) and the half-light radius, where Υ is determined by normalizing to the central velocity dispersion based on spherical and isotropic models fitted to the Nuker-law profile. The isotropized DF $\bar{f}(\mathcal{E})$ of the galaxies (see equation 26) can be obtained by the Eddington formula (Binney & Tremaine 1987):

$$\bar{f}(\mathcal{E}) \approx \frac{1}{\sqrt{8}\pi^2} \frac{d}{d\mathcal{E}} \int_{-\infty}^{\mathcal{E}} \frac{1}{m_*} \frac{d\rho}{d\psi} \frac{d\psi}{\sqrt{\psi - \mathcal{E}}}, \qquad (48)$$

where

$$\rho(r) = \Upsilon j(r) = -\frac{\Upsilon}{\pi} \int_r^\infty \frac{dI}{dR} \frac{dR}{\sqrt{R^2 - r^2}}. \qquad (49)$$

The total mass of the central BHs is estimated by

$$M_\bullet = 1.2(\pm 0.2) \times 10^8 \, M_\odot \left(\frac{\sigma_e}{200 \, \text{km s}^{-1}}\right)^{3.75(\pm 0.3)}, \qquad (50)$$

where σ_e is the luminosity-weighted line-of-sight velocity dispersion inside the half-light radius (Gebhardt et al. 2000). The BH mass–velocity dispersion relation due to Merritt & Ferrarese (2001) uses central velocity dispersion and the velocity dispersion has a different exponent, $4.72(\pm 0.36)$. In this paper we use equation (50) to estimate the central BH mass. The results will not be affected much if we equate the central velocity dispersion to σ_e and use the different exponent 4.72. Of these 41 galaxies in Table 1, nine DFs obtained from equation (48) are negative, perhaps because the spherical and isotropic dynamical model is not appropriate or because the Nuker law is a poor fit, and these are deleted from our sample. All of these nine galaxies are core galaxies with shallow central cusps ($0 \lesssim \gamma \lesssim 0.04$). Generally, σ_e is determined by the gravitational potential of stars and affected little by the gravitational potential of the central BH, but that is not true for two galaxies with $\gamma > 1$, which are also deleted from our sample. Thus a total of 30 galaxies is left in our study.

Table 1. Galaxy Sample. The second column gives the Hubble types of the galaxies: 1 = E, 2 = E/S0 or S0, 3 = Sa or Sb, 4 = dE or dSph. The r_b and μ_b are the break radius and break surface brightness (corrected for Galactic extinction, in V mag arcsec^{-2}) in the Nuker law (equation 47); α, β and γ are the exponents describing the sharpness of the break, the outer and inner slopes of the Nuker law. The r_e is the half-light radius and Υ_V is the mass-to-light ratio in the V band. $M_\bullet$ is the central black hole mass, and σ_e is the luminosity-weighted line-of-sight velocity dispersion inside r_e. All the parameters except $M_\bullet$ and σ_e are adapted from Faber et al. (1997), where $H_0 = 80\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$. The σ_e comes from the calculation based on a spherical and isotropic model in this paper, and $M_\bullet$ is obtained from the correlation of BH mass with galactic velocity dispersion (equation 50). The nine galaxies with '$DF < 0$' are deleted from our study because their distribution functions obtained from the Eddington formula are negative (equation 48), and so are the last two galaxies with $\gamma > 1$ because their σ_e are significantly affected by the gravitational potential of central BHs (see Section 5.1).

Name	Type	$\log_{10}(r_b)$ (pc)	μ_b	α	β	γ	$\log_{10}(r_e)$ (pc)	Υ_V ($\mathrm{M_\odot/L_\odot}$)	Distance (Mpc)	$M_\bullet$ ($\mathrm{M_\odot}$)	σ_e $\mathrm{km\,s^{-1}}$
NGC596	1	2.56	18.03	0.76	1.97	0.55	3.49	4.2	21.2	3.5e7	140
NGC720	1	2.55	17.50	2.32	1.66	0.06	3.64	8.2	22.6	2.0e8	230
NGC1172	1	2.55	18.61	1.52	1.64	1.01	3.75	2.6	29.8	4.5e6	84
NGC1399	1	2.43	17.06	1.50	1.68	0.07	3.56	12.7	17.9	6.8e8	320
NGC1426	1	2.23	17.53	3.62	1.35	0.85	3.44	4.9	21.5	2.4e7	130
NGC1600	1	2.88	18.38	1.98	1.50	0.08	4.06	14.3	50.2	6.5e8	310
NGC3115	2	2.07	16.17	1.47	1.43	0.78	3.17	7.1	8.4	3.0e8	260
NGC3377	1	0.64	12.85	1.92	1.33	0.29	3.21	2.9	9.9	1.1e7	110
NGC3379	1	1.92	16.10	1.59	1.43	0.18	3.23	6.9	9.9	1.5e8	210
NGC3599	2	2.12	17.58	13.01	1.66	0.79	3.47	2.1	20.3	1.4e6	61
NGC3605	1	1.94	17.25	9.14	1.26	0.67	3.23	4.1	20.3	7.5e6	96
NGC4168	1	2.65	18.33	0.95	1.50	0.14	3.90	7.5	36.4	1.1e8	200
NGC4239	4	1.98	18.37	14.53	0.96	0.65	3.08	3.4	15.3	2.4e6	70
NGC4365	1	2.25	16.77	2.06	1.27	0.15	3.79	8.4	22.0	3.2e8	260
NGC4387	1	2.52	18.89	3.36	1.59	0.72	3.06	5.3	15.3	1.2e7	110
NGC4434	1	2.25	18.21	0.98	1.78	0.70	3.14	4.7	15.3	9.1e6	100
NGC4458	1	0.95	14.49	5.26	1.43	0.49	3.30	4.0	15.3	2.7e6	72
NGC4464	1	1.95	17.35	1.64	1.68	0.88	2.60	4.8	15.3	1.5e7	120
NGC4478	1	1.10	15.40	3.32	0.84	0.43	3.02	5.0	15.3	6.8e7	170
NGC4486	1	2.75	17.86	2.82	1.39	0.25	3.89	17.7	15.3	1.4e9	390
NGC4551	1	2.46	18.83	2.94	1.23	0.80	3.12	7.3	15.3	2.3e7	130
NGC4564	1	1.59	15.70	0.25	1.90	0.05	3.21	4.8	15.3	4.0e7	150
NGC4621	1	2.34	17.20	0.19	1.71	0.50	3.54	6.7	15.3	1.8e8	220
NGC4636	1	2.38	17.72	1.64	1.33	0.13	3.88	10.4	15.3	1.6e8	220
NGC4649	1	2.42	17.17	2.00	1.30	0.15	3.74	16.2	15.3	1.2e9	370
NGC4697	1	2.12	16.93	24.86	1.04	0.74	3.58	6.8	10.5	1.4e8	210
NGC4874	1	3.08	19.18	2.33	1.37	0.13	4.44	15.0	93.3	4.3e8	280
NGC4889	1	2.88	18.01	2.61	1.35	0.05	4.15	11.2	93.3	8.7e8	340
NGC5813	1	2.04	16.42	2.15	1.33	0.08	3.82	7.1	28.3	1.2e8	200
NGC6166	1	3.08	19.35	3.32	0.99	0.08	4.49	15.6	112.5	1.1e9	360
NGC221	1	−0.26	11.77	0.98	1.36	0.01	2.18	2.3	0.8	*DF < 0*	
NGC1316	3	1.55	14.43	1.16	1.00	0.00	3.84	2.6	17.9	*DF < 0*	
NGC1400	1	1.54	15.41	1.39	1.32	0.00	3.60	10.7	21.5	*DF < 0*	
NGC1700	1	1.19	13.95	0.90	1.30	0.00	3.61	4.0	35.5	*DF < 0*	
NGC2636	1	1.17	15.68	1.84	1.14	0.04	2.86	3.0	33.5	*DF < 0*	
NGC3608	1	1.44	15.45	1.05	1.33	0.00	3.54	7.0	20.3	*DF < 0*	
NGC4472	1	2.25	16.66	2.08	1.17	0.04	3.89	9.2	15.3	*DF < 0*	
NGC4552	1	1.68	15.41	1.48	1.30	0.00	3.35	7.7	15.3	*DF < 0*	
NGC7768	1	2.30	16.99	1.92	1.21	0.00	4.18	9.5	103.1	*DF < 0*	
NGC221V	1	2.43	20.42	1.72	3.55	1.21	2.18	1.3	19.2	σ_e is affected by $M_\bullet$	
NGC4742	2	1.93	16.69	48.60	1.99	1.09	2.85	1.8	12.5	σ_e is affected by $M_\bullet$	

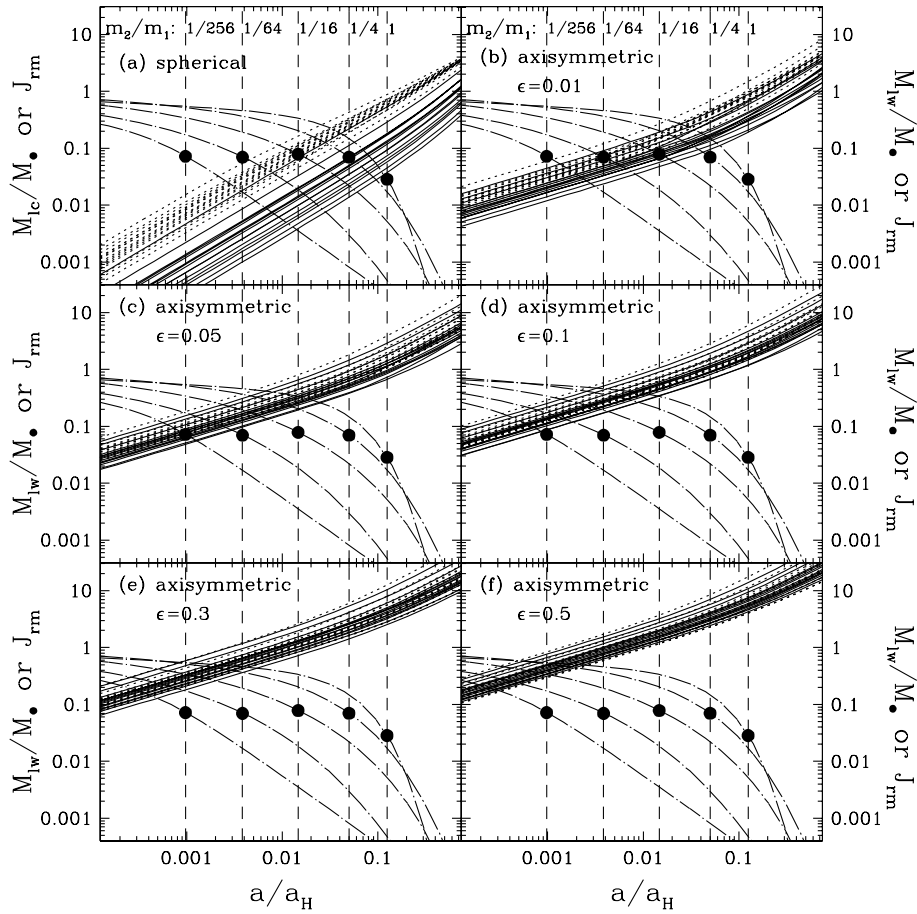## 5.2 Total stellar mass in the full loss region

Before studying the details of the BBH evolution time-scales, we will first generalize equations (18) and (31) to the non-hard binary stage to obtain the total stellar mass in the full loss cone/wedge (for spherical/axisymmetric galaxies) to see if the total stellar mass in the full loss cone/wedge is large enough to provide the removed stellar mass from galactic core before the BBH becomes hard (Fig. 1).

Quinlan's simulation (1996) gives the mass removing rate from the galactic core caused by interactions between a BBH and stars (J_{rm}, see equation 45) as a function of σ_c/v_c (or a/a_h) for five different BH mass ratios ($m_2/m_1 = 1/256, 1/64, 1/16, 1/4$ and 1), where $v_c \equiv \sqrt{GM_\bullet/a}$. We can roughly obtain the removing rates for other BH mass ratios by interpolating the rates of those five mass ratios. As shown in Fig. 1 (or see fig. 5 in Quinlan 1996), the

removing rate J_{rm} is small at large semimajor axes (e.g., $J_{rm} \lesssim 10^{-3}$ at $a/a_H \simeq 0.3$) and increases steeply as the BBH hardens, and we may approximate equation (45) as $M_{rm}(a) \simeq M_\bullet J_{rm}(a)$. The removing rate is about $0.03 - 0.1$ when the BBH becomes hard, and then increases towards a constant in the range $0.1 - 1$ as the BBH hardens. This is roughly consistent with the factor $1/(2K) \sim 0.3$ ($Kf_a = 1.60, f_a \sim 1$) in equation (9) or the energy gain in equation (17) (i.e., soon after the BBH becomes hard, especially when $a \ll a_h$, most stars having close encounters with the BBH will be removed from the galactic core).

Fig. 1(a) shows the total stellar mass in the full loss cone $M_{lc}(a)$ and the removing rate as a function of a if the galaxies in Table 1 are spherical. As seen from Fig. 1(a), with the decrease of a, $M_{lc}(a)$ decreases and the mass removing rate from the galactic core $J_{rm}(a)$ increases. For the BBH with equal masses ($m_2/m_1 = 1$), we have $M_{lc}/M_{rm} \simeq M_{lc}/M_\bullet J_{rm} \gtrsim 1$ at $a \gtrsim a_h$ [i.e., in Fig. 1a the solid

Figure 1. The total stellar mass in the full loss cone/wedge (for spherical/axisymmetric galaxies) and the removing rate from the galactic core as a function of a/a_H (a_H represents the radius of the sphere of influence of the BH $M_\bullet$, cf. equation 2). The solid lines (for core galaxies in Table 1) and the dotted lines (for power-law galaxies in Table 1) represent the total stellar mass in the full cones $M_{lc}(a)$ in panel (a) (cf. equation 18) and the total stellar mass in the full loss wedges $M_{lw}(a)$ in panels (b)–(f) (cf. equation 31). The dot-dashed lines represent the removing rate from the galactic core J_{rm} (see equation 45) for the five different m_2/m_1 ratios ($m_2/m_1 = 1/256, 1/64, 1/16, 1/4$ and 1) which can be obtained from Quinlan's simulation (1996). The vertical dashed lines represent the place where the BBH first becomes hard, and the corresponding m_2/m_1 ratio is labelled at the right-hand side of each dashed line. For each given m_2/m_1 ratio, the crossing of the vertical dashed line ($a = a_h$) and the removing rate curve, J_{rm}, is labelled as a solid circle. The total mass in the full loss region increases with increasing ϵ. The $M_{lc}(a_h)$ and $M_{lw}(a_h)$ also increase with increasing m_2/m_1 ratios. For the cases in which the solid or dotted lines are higher than the solid circles at $a = a_h$, the removed stellar mass from galactic cores during the non-hard binary stage can be ignored, and the loss region is approximately full at $a = a_h$ (see details in Section 5.2).

circle representing the stellar removing rate $J_{rm}(a = a_h)$ for $m_2/m_1 = 1$ is located below all the $M_{lc}(a \gtrsim a_h)/M_\bullet$ curves]; so the mass of the removed stars from the galactic core during the non-hard binary stage can be ignored and the loss cone is approximately full at $a = a_h$. Compared with the BBH with equal masses, the BBH with low m_2/m_1 ratio becomes hard at smaller semimajor axis and the removed mass may be larger than the mass in the full loss cone before the hard binary stage [i.e., in Fig. 1a, solid circles for low m_2/m_1 ratios are located above some of the $M_{lc}(a \gtrsim a_h)/M_\bullet$ curves]. For the BBHs whose loss cone is depleted before the hard binary stage begins (e.g., 20 out of our sample of 30 galaxies for $m_2/m_1 = 1/16$), we obtain the depletion semimajor axes a_{dp} approximately by setting $M_{lc}(a_{dp})/M_\bullet = J_{rm}(a_{dp})$, and we will use equation (46) to obtain the BBH hardening time-scales at $a_h < a < a_{dp}$ in Section 5.3. Note that in Fig. 1(a), at a given $a(\lesssim 0.1a_H)$, $J_{rm}(a)$ decreases with BH mass ratios, and hence a_{dp} also decreases with BH mass ratios.

Figs 1(b)–(f) show the total mass in the full loss wedge $M_{lw}(a)$ and the removing rate as a function of a if the galaxies in Table 1 are axisymmetric. The total mass in the full loss wedge depends on

the degree of flattening ϵ (defined in Section 3.2.2), which is assumed to be a constant here. The average mass ellipticity of the sample in our study, which can be obtained from table 2 in the paper by Faber et al. (1997), is about 0.26. Their average ϵ is ~0.36 if equation (29) is applied. As seen from Figs 1(b)–(f) ($\epsilon = 0.01, 0.05, 0.1, 0.3$ and 0.5), $M_{lw}(a)$ increases with the increase of ϵ. If ϵ is very small (e.g., $\epsilon < 0.01$), the loss wedge is approximately full at $a = a_h$ only for $m_2/m_1 \simeq 1$, as in the spherical case. When $\epsilon \gtrsim 0.1$, the loss wedge is approximately full at $a = a_h$ for $m_2/m_1 \gtrsim 1/256$ (i.e., the corresponding solid circle is located below all the $M_{lw}/M_\bullet$ curves at $a \gtrsim a_h$ in Figs 1d–f).

For triaxial galaxies, the maximum mass of the stars which may have close encounters with the BBH is not less than the mass in the full loss wedge of axisymmetric galaxies with the same parameter ϵ. In the study below of the evolution of BBHs in triaxial galaxies (see Section 5.5 below), we will consider mainly BBHs with equal masses. In this case, the stellar mass removed during the non-hard binary stage can be ignored, and the loss region is approximately full at $a = a_h$ as in axisymmetric galaxies (cf. Figs 1b–f). For other cases in triaxial galaxies (varying BH mass ratios at a given ϵ), the

variation of the BBH evolution should follow similar trends to those in axisymmetric galaxies, and we need not consider the maximum mass of the stars which may have close encounters with the BBH in triaxial galaxies.

5.3 Spherical galaxies

Assuming that the galaxies in Table 1 are spherical, we obtain the BBH evolution time-scales as a function of BBH semimajor axis
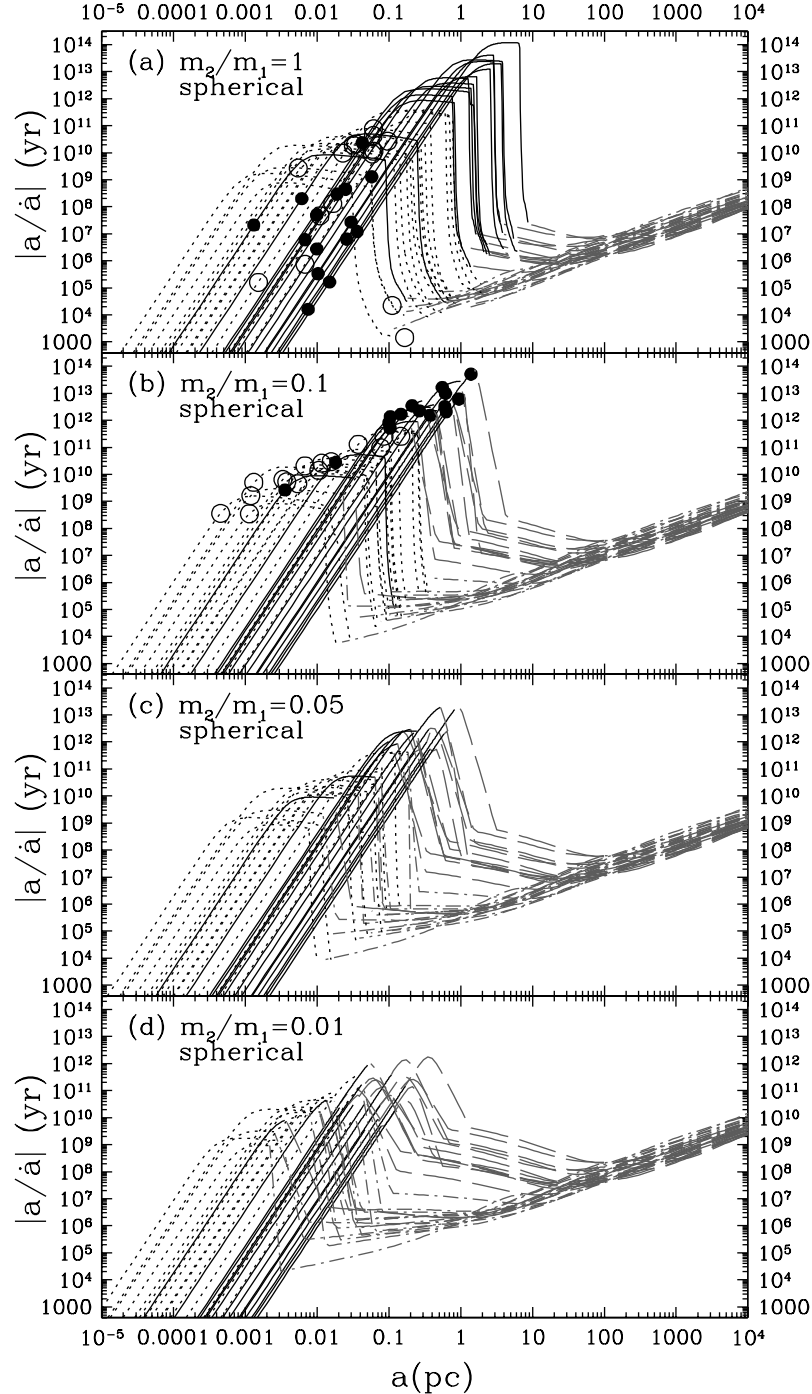
Figure 2. BBH evolution time-scales for spherical galaxies. The panels show different mass ratios m_2/m_1, and the curves represent time-scales $|a/\dot{a}|$ for the galaxies in Table 1. The long-dashed lines (for core galaxies: $\gamma \lesssim 0.3$) and the dot-short-dashed lines (for power-law galaxies: $\gamma \gtrsim 0.5$) represent the time-scales before the hard binary stage; the solid lines (for core galaxies) and the dotted lines (for power-law galaxies) represent the time-scales after the BBHs become hard. In panel (a), the solid circles (for core galaxies) and the open circles (for power-law galaxies) represent the maximum Brownian motion magnitudes of the BBH centres of mass ($r_{\rm ran}$ in equation 53). In panel (b), the solid circles (for core galaxies) and the open circles (for power-law galaxies) mark the boundary between the hard binary and gravitational radiation stage or the place where gravitational radiation becomes dominant in the non-hard binary stage (i.e., $a = a_{\rm gr}$ when $t_{\rm h} = t_{\rm gr}$). BBHs in power-law galaxies have shorter lifetimes than those in core galaxies. With decreasing mass ratios, the evolution time-scales at large radii (e.g., $a \gtrsim 10^2$ pc) increase and hence BBHs with very low mass ratios (say, $\lesssim 0.001$) cannot be formed by mergers of galaxies. The bottleneck stages shift to smaller radii as the mass ratio decreases. In most core galaxies the bottleneck time-scales become shorter as the mass ratio decreases, while in most power-law galaxies they are affected little by mass ratios. The lifetime of BBHs is not affected by their Brownian motion. See details in Section 5.3.

(or the separation of the two BHs before they become bound) and the BH mass ratio $m_2/m_1 \leq 1$. The results are shown in Fig. 2, where the BBH evolution curves are composed of two parts: one comes from equations (9)–(11) and (19)–(27) in Section 3 describing the evolution in the hard binary stage and the gravitational radiation stage (dotted or solid lines), and the other from equations (39)–(44) or equation (46) in Section 4 describing the evolution in the dynamical friction stage and the non-hard binary stage (long-dashed or dot-short-dashed lines). A characteristic feature of all these plots is a pronounced peak in the evolution time-scale between 0.001 and 10 pc, where dynamical friction and gravitational radiation are both relatively ineffective. This is always the slowest evolution stage if we do not consider the possibly longer time-scales at very large radii, $a \gtrsim 10^4$ pc. We shall call this stage the 'bottleneck'.

To study the effect of different BH mass ratios on the BBH evolution, we will first consider BBHs with equal masses (in Section 5.3.1), and then consider those with unequal masses (in Section 5.3.2), although actually there is no strict dividing line between these cases.

5.3.1 Evolution of BBHs with equal BH masses

In this subsection we will consider the evolution of BBHs with equal BH masses ($m_2/m_1 = 1$, Fig. 2a) for which the removed stellar mass from the core during the non-hard binary stage can be ignored and the loss cone is approximately full at $a = a_h$ [$\eta(\mathcal{E}) \simeq 1$ in equation 18] (see Section 5.2 or Fig. 1a).

As seen from Fig. 2(a), the overall evolution of the BBHs has the following trends: starting at large radii (~ 10 kpc), the evolution
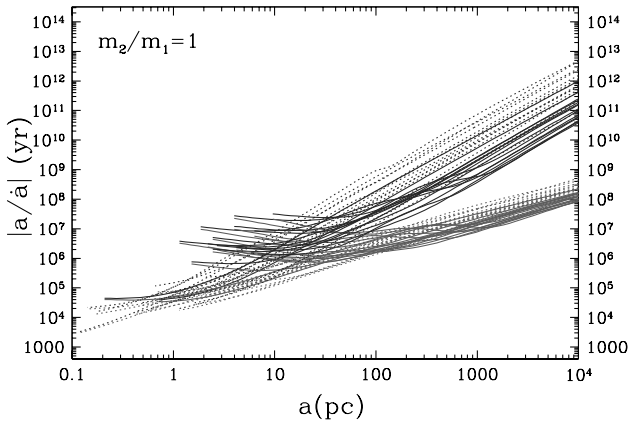
Figure 3. Evolution time-scales of BBHs (with equal BH masses) as a function of the BBH separation or the orbit semimajor axes in the dynamical friction stage and non-hard binary stage. The solid lines represent core galaxies, and the dotted lines represent power-law galaxies. At small semimajor axes, the curves are cut off at $a = a_h$. The curves are composed of two groups, which are easy to distinguish at large radii ($a \sim 10^3$–10^4 pc). The lower group ($|a/\dot{a}| \sim 10^8$–10^9 yr at $a \simeq 10^4$ pc) are the evolution time-scales including the bound stars around m_2, and the higher group ($|a/\dot{a}| \sim 5 \times 10^{10}$–$5 \times 10^{12}$ yr at $a \simeq 10^4$ pc) are the evolution time-scales without including the bound stars around m_2 (see Section 4). At small radii ($a \rightarrow a_h$), the two groups have almost the same time-scales. At large radii ($a \gtrsim 100$ pc), the evolution time-scales decrease with decreasing a; at small radii ($a \lesssim 100$ pc), the evolution time-scales increase for most core galaxies and decrease for most power-law galaxies (see Section 5.3.1). The evolution time-scales at large radii are shorter than the Hubble time 10^{10} yr if and only if we assume that the BHs are accompanied by the tidally truncated remnant of stars from their original host galaxy.

time-scales first decrease with decreasing radii, then increase at some intermediate radii to reach the bottleneck, and finally, when the BBH semimajor axis a is small enough ($\lesssim 10^{-3}$–1 pc), the evolution time-scales decrease with decreasing a. The bottleneck occurs during the hard binary stage.

Before the BBH becomes hard, the evolution time-scales are always shorter than the Hubble time $t_{\text{Hubble}} = 10^{10}$ yr; these relatively short time-scales reflect the enhancement of dynamical friction due to the bound stellar mass around m_2. Fig. 3 shows that without including the bound stellar mass around m_2, the dynamical friction time-scales would be much larger than t_{Hubble} at large radii (e.g. $\sim 5 \times 10^{10}$–5×10^{12} yr at 10 kpc). The dynamical friction time-scale decreases with the decrease of r, and when r decreases to the place where m_1 dominates the potential, we have $|r/\dot{r}| \propto v_{m_2}^3(r)/\rho(r) \propto r^{-1.5}/r^{-\gamma-1} \propto r^{\gamma-0.5}$ (see equations 39–42). Thus, at small radii ($a \lesssim 100$ pc), with the decrease of r, the evolution time-scale continues decreasing for power-law galaxies ($\gamma \gtrsim 0.5$), but increases for core galaxies ($\gamma \lesssim 0.3$).

As seen from Fig. 2(a), after the BBH becomes hard, starting at $a = a_h$, the BBH evolution time-scale increases steeply with decreasing separation a, then becomes nearly flat, and then decreases as a^4 once the BBH reaches the gravitational radiation stage. Most of the BBH evolution time-scales at the bottleneck are longer than the Hubble time. In these galaxies we might expect to find BBHs with semimajor axes in the range 0.01–10 pc. BBHs in the core galaxies generally reach the bottleneck at larger semimajor axes ($a_h \sim 1$–10 pc) and need more time to reach the gravitational radiation stage than those in the power-law galaxies. At the beginning of the hard binary stage, the BBH orbital decay rate is controlled by the draining rate of the stars initially in the loss cone, and the BBH evolution time-scale increases steeply as the stars in the loss cone are depleted. The loss cones are depleted (generally when $a \lesssim 0.1a_H$) before the BBHs are close enough for gravitational radiation to be effective. Thereafter, the clearing rates are controlled by diffusion of stars into the loss cone. Fig. 4 shows that the inner parts of most of the galaxies are in the diffusion regime ($r < r_{lc}$) before the gravitational radiation stage begins (generally $a_{gr}/a_h \gtrsim 0.001$ in Fig. 5), although the transition radii r_{lc} decrease with the shrinking of the loss cone. As mentioned in Section 3.2.1, in the diffusion regime the flux of stars into the loss cone is insensitive to a, and hence the BBH evolution time-scales controlled by the stellar diffusion are nearly constant. As seen from Fig. 2(a), the bottleneck time-scales in core galaxies are larger than those in power-law galaxies, which reflects the fact that core galaxies have lower central densities (hence the stellar diffusion rates into the loss cone are slower) and more massive BHs [hence the average relative change in the BBH energy caused by interactions with a star passing the BBH vicinity is smaller ($|\Delta E/E| = 2Km_*/M_\bullet$)] (see equation 9). Gravitational radiation becomes dominant at larger semimajor axes in core galaxies, because their BBH bottleneck time-scales are longer and the gravitational radiation time-scale (equation 10) is smaller for large BHs.

5.3.2 Evolution of BBHs with unequal BH masses

For the BBH with unequal masses, as discussed in Section 5.2, low-angular-momentum stars may have been depleted before the BBH first becomes hard, because the BBH becomes hard at smaller semimajor axes and the total stellar mass in the full loss cone decreases with the decrease of the semimajor axes. Figs 2(b)–(d) show the evolution of the BBH with unequal masses ($m_2/m_1 < 1$).
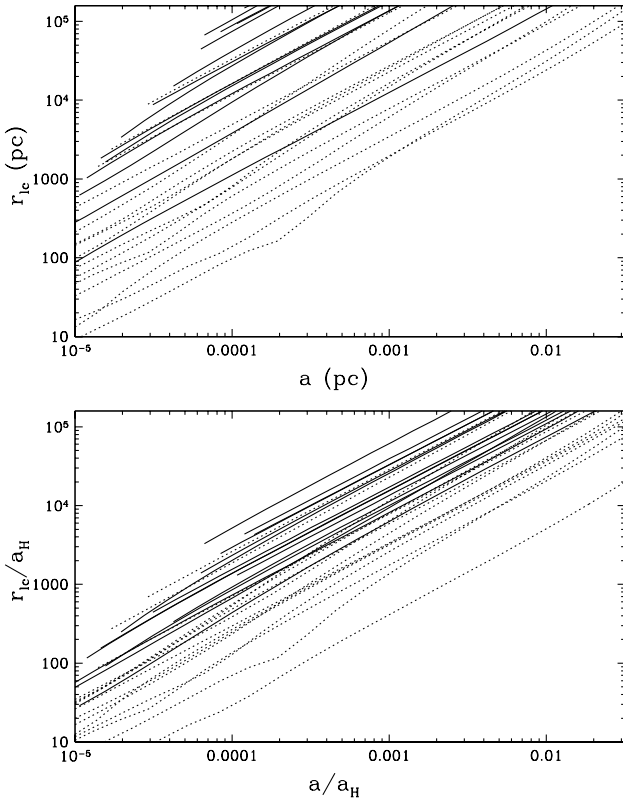
Figure 4. The semimajor axes a of the BBHs versus the transition radii r_{lc} between the pinhole ($r > r_{lc}$) and the diffusion ($r < r_{lc}$) regimes in spherical galaxies. The lower ends of the lines are cut off at the Schwarzschild radius $r_s = 2GM_\bullet/c^2$. The solid and dotted lines have the same meaning as those in Fig. 2. At a given a, most core galaxies have larger transition radii r_{lc} than power-law galaxies, because core galaxies generally have more massive BHs and lower central densities. In the hard binary stage, if the loss cone in spherical galaxies is depleted, the inner parts of the galaxies are generally in the diffusion regime.

As seen from Figs 2(b)–(d), roughly, the overall evolution trend of the BBHs with unequal masses looks like those with equal masses; but the bottleneck stages of the BBHs with unequal masses can be either in the non-hard binary stage or in the hard binary stage, and tend to occur at smaller semimajor axes. The bottleneck time-scales of most of the BBHs are still longer than the Hubble time. Generally, the bottleneck time-scales decrease with decreasing mass ratios in most of the core galaxies, but remain unchanged in most of the power-law galaxies.

Before the BBHs become hard, at large radii ($r \sim 10^2$–10^4 pc), the BBH evolution time-scales increase with decreasing m_2/m_1, but most of them are still smaller than the Hubble time $t_{\text{Hubble}} = 10^{10}$ yr (Figs 2b–d). However, when $m_2/m_1 \lesssim 0.01$, the BBH evolution time-scales in some of the galaxies exceed the Hubble time $t_{\text{Hubble}} = 10^{10}$ yr at $r \simeq 10$ kpc. So, in BBHs with $m_2/m_1 \lesssim 0.01$, the orbit of the smaller BH would not decay to the galactic centre from the outer edge of the galaxy within a Hubble time.

For the BBH with unequal masses, if the removed mass from the galactic core during the non-hard binary stage can be ignored and the loss cone is approximately full when the BBH becomes hard, the evolution of the BBH after $a = a_h$ would be like that of the BBH with equal masses, which is first controlled by the depletion of the loss cone (when the evolution time-scale increases sharply) and then controlled by the stellar diffusion (see flat parts of the BBH evolution time-scale curves in Figs 2b–d) and gravitational
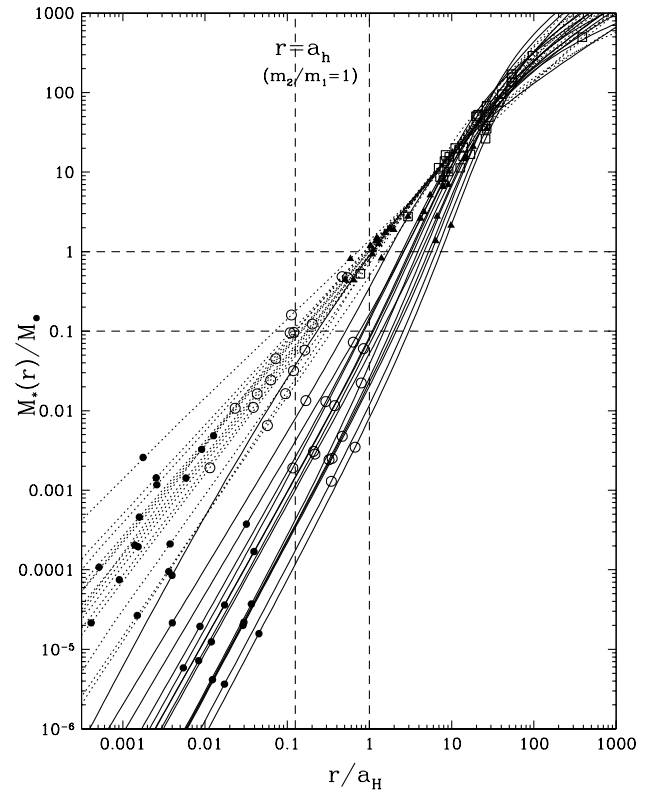
Figure 5. The ratios of enclosed stellar mass within radius r to BBH total mass. The solid and dotted lines have the same meaning as those in Fig. 2. The left vertical dashed line represents the semimajor axis a_h where the BBH with equal masses first becomes hard. The open squares represent the break radii of the surface brightness profile r_b (cf. equation 47); the open circles represent the collision radii r_{coll} (equation 25); the solid circles represent the boundary between the hard binary and gravitational radiation stage a_{gr} when $m_2/m_1 = 1$ in Fig. 2(a). In the hard binary stage, if the loss cone in spherical galaxies is depleted, most of the stars that diffuse into the loss cone have the energy $\mathcal{E}_{\text{peak}} \equiv \psi(r_{\text{peak}})$, and the apocentres of the radial orbits, r_{peak}, are labelled as solid triangles. For hard BBHs, we have $r_{\text{peak}} \gg \max(r_{\text{coll}}, a_h) \gtrsim f_a a (f_a \simeq 1)$, which helps to justify the generalization of solution of the Fokker–Planck equation in the tidal disruption context to the BBH systems.

radiation. The bottleneck stage of the BBH with unequal masses shifts to smaller semimajor axes, because it becomes hard at smaller semimajor axis. Note that the flat part (controlled by stellar diffusion) of the BBH evolution curve of each galaxy, if any, has almost the same height in all the panels of Fig. 2, which reflects the fact that the BH mass ratio does not affect the hardening time t_h (see equation 9).

For the BBH with unequal masses, if the loss cone is depleted when the BBH first becomes hard [$\eta(\mathcal{E}) \simeq 0$ in equation 18], the following BBH evolution will be directly controlled by the stellar diffusion to the loss cone (see the flat part of the BBH evolution curves, mostly for power-law galaxies) and then gravitational radiation, or directly controlled by gravitational radiation (no flat part in the BBH evolution curves, mostly for core galaxies). Before the BBH becomes hard, we obtain the place a_{dp} where the low-angular-momentum stars are depleted by setting $M_{lc}(a_{\text{dp}}) = M_{\text{rm}} \simeq M_\bullet J_{\text{rm}}(a_{\text{dp}})$ (see equation 45). Also, we use equation (46) (log-linear approximation) to obtain the BBH hardening time-scales from a_{dp} to a_h, because the evolution at $a_h < a < a_{\text{dp}}$ is an intermediate process between the evolution controlled by dynamical friction and the evolution controlled by three-body

interactions with the stars that diffuse from high-angular-momentum orbits; then we use equation (11) to obtain the BBH evolution time-scales. As seen from Figs 2(b)–(d), the evolution time-scale curves of these BBHs increase steeply from $a = a_{dp}(> a_h)$ and smoothly connect to the time-scales in the hard binary stage. If the BBH evolution is controlled by stellar diffusion to the loss cone when it just becomes hard, its lifetime is not affected by mass ratios. For some of the BBHs, gravitational radiation may become the dominant dissipative force when $a > a_h$, and in the non-hard binary stage their evolution time-scales show a peak or maximum (i.e., the bottlenecks) at $a = a_{gr}$ where $t_{gr}(a_{gr}) = t_h(a_{gr})$. With decreasing BH mass ratio, the merger time-scale becomes shorter because the gravitational time-scale t_{gr} (equation 10) decreases as a^4, although at a given a, t_{gr} is inversely proportional to mass ratio m_2/m_1 (when $m_2/m_1 \ll 1$). Here, as the BH mass ratio decreases, the bottleneck stage shifts to smaller semimajor axes, because either a_h or a_{dp} decreases.

5.3.3 Sharp change of the BBH evolution time-scale at $a = a_h$

As seen from Fig. 2(a), the BBH evolution time-scale increases sharply at $a = a_h$. This sharp increase arises because the loss cone is depleted almost as soon as the BBH becomes hard. A similar sharp increase starting at $a = a_h$ also exists in Figs 2(b)–(d) for those galaxies whose loss cones are approximately full when the BBH first becomes hard. In this subsection we will see that the coincidence between the increase in the evolution time-scale and a_h is caused by some simplification in our analysis; however, even if more realistic treatment is considered, the sharp increase would also start at some place close to $a = a_h$, though not exactly at $a = a_h$, and our conclusions would not be significantly affected.

Note that in the analysis in Sections 2 and 3, before the BBH becomes hard, the velocity dispersion of the low-angular-momentum stars not removed from the galactic core is assumed to be unaffected by interactions with the BBH; and after the BBH becomes hard, each star having had close encounters with the BBH is removed from the galactic core with an energy gain. In fact, the heating of the stars is a continuous physical process, i.e., in the non-hard binary stage, the stars should also be heated as the BBH hardens, even though they are not removed from the galactic core.

From scattering experiments with the restricted three-body approximation, the BBH energy change due to stellar encounters in the non-hard stage is generally very small [e.g., $|\Delta E(a = 10a_h)| \simeq 0.03|\Delta E(a = a_h)| \simeq 0.1m_*\sigma_c^2$, cf. fig. 1 in Quinlan 1996] and decreases steeply with increasing a (which is consistent with the steep decrease of the stellar removing rate from the core with increasing a at $a > a_h$ in Fig. 1); only when a is close to a_h does the relative change in the BBH energy, $\Delta E/E$, increase steeply towards a constant (independent of E) as the BBH hardens (cf. fig. 1 in Quinlan 1996). Hence, in the non-hard binary stage, the energy increase of a star occurs through a slow diffusion process. If all the energy loss of the BBH is used to heat the low-angular-momentum stars, the heating would become significant after $Gm_1m_2/2a \simeq 3M_{lc}(a)\sigma_c^2/2$, i.e., after

$$a \simeq 5a_h\left(\frac{2m_1}{m_1+m_2}\right)^{1/2}\left[\frac{0.03M_\bullet}{M_{lc}(a_h)}\right]^{1/2} \lesssim 5a_h, \quad (51)$$

where $M_{lc}(a) \simeq M_{lc}(a_h)a/a_h$ for $a \ll a_H$ (cf. equation 18) and $M_{lc}(a_h)/M_\bullet \sim 0.03-1$ for those galaxies whose loss cones are approximately full at $a = a_h$ in Fig. 1(a). Thus the heating of the stars becomes significant only at some place close to a_h and the

sharp increase of the BBH evolution time-scales would also start not far away from $a = a_h$. The semimajor axis a_h is also the point where the typical energy gain of a scattered star is just large enough for it to escape from the galactic core. Thus we expect a sharp increase in the evolution time-scale near $a \simeq a_h$.

5.3.4 Testing assumptions

We can use the results obtained above to test some of the assumptions that we have made in our analysis.

(i) In the hard binary stage, if the loss cone is depleted, the BBH decay is controlled by the stellar diffusion into the loss cone. We define $\mathcal{E}_{peak}$ as the energy at the peak of stellar diffusion rates $F^{lc}(\mathcal{E}, a)$ (see equation 20); $\mathcal{E}_{peak}$ is generally near or smaller than $\psi(a_H)$, not at $\mathcal{E}_{lc}(a)$ because it is the factor $F_{max}(\mathcal{E})$ that dominates the shape of $F^{lc}(\mathcal{E}, a)$ in equation (20) (see also MT). As seen from Fig. 5, most of the stars contributing to the BBH orbital decay have energy $\mathcal{E}_{peak} \ll GM_\bullet/f_a a$ or $r_{peak} \gg \max(r_{coll}, a_h) \gtrsim f_a a$, where r_{peak} is defined by $\mathcal{E}_{peak} = \psi(r_{peak})$ and r_{coll} (see equation 25) is the collision radius, which helps to justify the generalization of solution of the Fokker–Planck equation in the tidal disruption context to the BBH systems. Fig. 5 also shows that the stellar mass within r_{peak} is higher than $0.1M_\bullet$ for all the galaxies, which confirms that resonant relaxation is not important here.

(ii) In the hard binary stage, when stellar diffusion dominates the BBH orbital decay, the diffusion time-scale into the loss cone at $\mathcal{E} \sim \mathcal{E}_{peak}$ is given by

$$t_{diff} \sim \frac{M_*(r_{peak})}{\int F(\mathcal{E}, a)\,d\mathcal{E}}\frac{J_{lc}^2(\mathcal{E}_{peak}, f_a a)}{J_c^2(\mathcal{E}_{peak})}$$

$$\lesssim (0.06 \sim 0.6)\left(\frac{Kf_a}{1.60}\right)\left(\frac{100 \sim 10}{r_{peak}/a_h}\right)\left(\frac{a}{a_h}\right)\left|\frac{a}{\dot{a}}\right|, \quad (52)$$

which is less than the evolution time-scale $|a/\dot{a}|$. Thus our use of time-independent solutions of the Fokker–Planck equation is justified.

(iii) With equation (12), we may also obtain the total mass of the stars strongly interacting with the BBHs during the hard binary stage, which is less than the core mass as seen from Fig. 6 ($m_2/m_1 = 1$) and is much less when $m_2/m_1 < 1$. The mass removed from the galactic core during the non-hard binary stage is also small ($\sim 0.1M_\bullet$), as seen from Fig. 1. So the approximation that we ignore the response of the galactic structure to BBH evolution is reasonable.

(iv) The centre of mass of BBHs is not expected to be located exactly at the galactic centre (Bahcall & Wolf 1976). The BBH centre of mass should be in equipartition with the stars, which implies that its rms radial excursion in core galaxies is

$$r_{ran} \sim \sqrt{\frac{\sigma_\bullet^2}{G\rho}} \sim \sqrt{\frac{m_*}{M_\bullet}}r_c = 0.01\,\text{pc}\sqrt{\frac{m_*}{M_\odot}\frac{10^8\,M_\odot}{M_\bullet}}\left(\frac{r_c}{100\,\text{pc}}\right), \quad (53)$$

where $\sigma_\bullet$ is the velocity dispersion of the BBH centre of mass. The 'break radius' r_b is roughly equivalent to the core radius r_c in core galaxies. Power-law galaxies have no well-defined core, but replacing r_c with the break radius r_b in equation (53) should give an upper limit to the likely motion of the BH. Hence, for both the core galaxies and the power-law galaxies, we set $r_c = r_b$ to obtain the wandering amplitude of the centre of mass of BBHs. Wandering of the BBH centre of mass might enlarge the loss cone, increase the clearing rates, and decrease the BBH hardening time (e.g. Quinlan
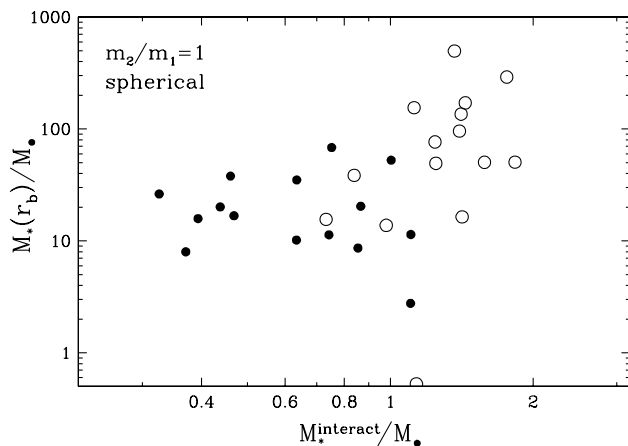
Figure 6. The core mass versus the total mass of stars interacting strongly with BBHs ($m_2/m_1 = 1$, in spherical galaxies) during the hard binary stage $a_{gr} \leq a \leq a_h$ (see equation 12 for $M_*^{interact}$, Fig. 5 for a_h/a_{gr} or Fig. 2a for the BBH evolution). The solid circles represent core galaxies, and the open circles represent power-law galaxies. For almost all of the galaxies, $M_*(r_b) > M_*^{interact}$, which helps to justify our neglect of the reaction of the core to the in-spiral of the BHs.

& Hernquist 1997). We believe that this effect is generally not important in these spherical galaxies for the following reasons. (1) For all the core galaxies, gravitational radiation becomes dominant before their BBH ($0.01 \lesssim m_2/m_1 \lesssim 1$) semimajor axes decrease to the value r_{ran} (e.g. Fig. 2a). (2) For most of the power-law galaxies, $a = r_{ran}$ are located at the flat parts of the evolution curves (see Fig. 2a) where stellar diffusion dominates the BBH orbital decay, or $a = r_{ran}$ are less than or around the semimajor axes where the loss cones are depleted (e.g., for the BBHs with lower mass ratios $0.01 \lesssim m_2/m_1 < 1$), and so the BBH lifetime is not sensitive to the wandering of the BBH. In Fig. 2(a), for only two power-law galaxies, $a = r_{ran}$ happens before their loss cones are depleted, but their BBH lifetimes are shorter than the Hubble time even without considering the wandering of the BBHs. (3) With decreasing m_2/m_1 ratios (say, $m_2/m_1 \lesssim 0.001$), although $a = r_{ran}$ may happen before the loss cone is depleted, the dynamical friction time-scales at large radii ($\gtrsim 10$ kpc) are longer than the Hubble time, and the smaller BH cannot sink into the centre of the galaxy with m_1.

5.4 Axisymmetric galaxies

Assuming that the galaxies in Table 1 are axisymmetric, we use equations (9)–(11) and (28)–(35) to obtain the BBH evolution time-scales as a function of BBH semimajor axes after the BBH becomes hard (equations 19–27 may also be used if the galaxies are nearly spherical). Before the hard binary stage, the BBH evolution time-scales are obtained from the equations in Section 4, as in spherical galaxies. We will first consider the evolution of BBHs with equal masses ($m_2/m_1 = 1$) in the galaxies with different degrees of flattening ϵ (defined in Section 3.2.2) (Fig. 7), and then consider the evolution of BBHs with different m_2/m_1 ratios for a given value of $\epsilon (= 0.3)$ (we estimated in Section 5.2 that the average ϵ in our sample was ~ 0.36) (Fig. 8).

As seen from Fig. 1 or Section 5.2, if the BBH has equal BH masses, the removed stellar mass from the galactic core during the non-hard binary stage can be ignored, and the loss wedge is approximately full when the BBH becomes hard [$\eta(\mathcal{E}) \simeq 1$ in equation 31]. As seen from Fig. 7 ($m_2/m_1 = 1$), the BBH evolution

time-scales follow similar trends (small evolution time-scales at large and small radii, and bottlenecks at intermediate radii) to those in spherical galaxies. The bottleneck occurs during the hard binary stage. After the BBHs become hard, the effects of flattening on the hardening time-scales depend on the value of ϵ. Fig. 7 shows that the bottleneck time-scales in the most nearly spherical galaxies (e.g., $\epsilon = 0.01$, 0.05 and 0.1 in panels a, b and c) are still longer than the Hubble time, but the bottleneck time-scales in most highly flattened galaxies (e.g., 21 out of our sample of 30 galaxies in panel d with $\epsilon = 0.5$) become shorter than the Hubble time. If ϵ is small (e.g., $\lesssim 0.01$), for most stars on centrophilic orbits, the characteristic angular momenta J_s (see Section 3.2.2) is not significantly larger than $J_{lc}(a)$, and the evolution reduces to the spherical case (see most of the power-law galaxies or some core galaxies in Fig. 7a and the galaxies in Fig. 2a). As ϵ increases, the effects of flattening become significant with the increase of the total stellar mass in the loss wedge (cf. Fig. 1). As seen from Figs 7(b)–(d), the semimajor axes of the BBHs decrease more in axisymmetric galaxies than in spherical galaxies before they reach the bottleneck stages, and the bottleneck stages shift to smaller semimajor axes as ϵ increases. For many BBHs in nearly spherical or low-ϵ axisymmetric galaxies (e.g., most of the BBHs in power-law galaxies in Figs 7a–c or some BBHs in core galaxies in Fig. 7a whose evolution curves have flat parts at intermediate radii $\sim 10^{-3} - 1$ pc), gravitational radiation becomes dominant still after the loss wedge/cone is depleted, and the lifetime is controlled by the stellar refilling rate to the loss wedge/cone. As in spherical galaxies, after the loss wedge/cone is depleted, the inner parts of most of the axisymmetric galaxies are also in the diffusion regime before the gravitational radiation stage begins. In the diffusion regime, the stellar diffusion rate to the loss wedge is also insensitive to a, and hence the BBH evolution time-scales controlled by the stellar diffusion are nearly constant. As seen from Figs 7(a)–(c), the flat parts of the evolution curves in Fig. 7 have almost the same heights as those in Fig. 2, and so the lifetimes of those BBHs are affected little by the flattening effect, which is consistent with the results in MT that the stellar tidal disruption rates by central BHs are not significantly changed by the flattening effect. With increasing ϵ, more and more BBH evolution curves have no flat parts, i.e., gravitational radiation becomes dominant before the loss wedge is depleted (e.g., BBHs in power-law galaxies in Fig. 7d and BBHs in core galaxies in Figs 7b–d), and their merger time-scales decrease, because the total stellar mass in the loss wedge increases and gravitational radiation becomes dominant at smaller semimajor axes.

Fig. 8 shows the evolution for the BBHs with unequal BH masses when the flattening parameter of the galaxies ϵ is 0.3. In this case the stellar mass removed from the galactic core during the non-hard binary stage can be ignored, and the loss wedge is approximately full when the BBHs (with $m_2/m_1 \gtrsim 0.01$) become hard (see Fig. 1e). As seen from Fig. 8, the bottleneck stages (also in the hard binary stage) shift to smaller semimajor axes as the mass ratio decreases, but not so obviously as those in spherical cases (Fig. 2) since the stellar mass in the loss wedges of the galaxies are generally so much not to be depleted even when the BBHs reach the gravitational radiation stage. As in spherical galaxies (Fig. 2), the merger time-scales decrease with decreasing mass ratios, because the gravitational radiation time-scale t_{gr} (equation 10) decreases as a^4. When m_2/m_1 decreases to 0.01, most of the BBH merger time-scales are shorter than the Hubble time (Fig. 8d).

The justifications of our assumptions in Section 5.3.4 can be generalized to axisymmetric galaxies (and triaxial cases below).

After the loss wedge is depleted, most of the stars contributing to the BBH decay have energy $\mathcal{E}_{peak} \ll GM_{\bullet}/f_a a$ or $r_{peak} \gg \max(r_{coll}, a_h) \gtrsim f_a a$; the stellar diffusion time-scales to the loss wedges are smaller than the BBH hardening time-scales, and the removed stellar mass from the core is also less than the total core mass. If ϵ is small (e.g., $\epsilon = 0.01$ in Fig. 7a), the wandering of BBHs is not important, just as in spherical galaxies. When ϵ is large (e.g., $\epsilon = 0.3$ or 0.5 in Fig. 8 or 7d), the wandering of BBHs is also not important in core galaxies. In a few of the power-law galaxies whose lifetimes are longer than the Hubble time, $a = r_{ran}$ happens before the loss wedges are depleted and their BBH evolution time-scales at $a = r_{ran}$ are shorter than the Hubble time, so the wandering of BBHs might decrease the BBH lifetime in those galaxies.
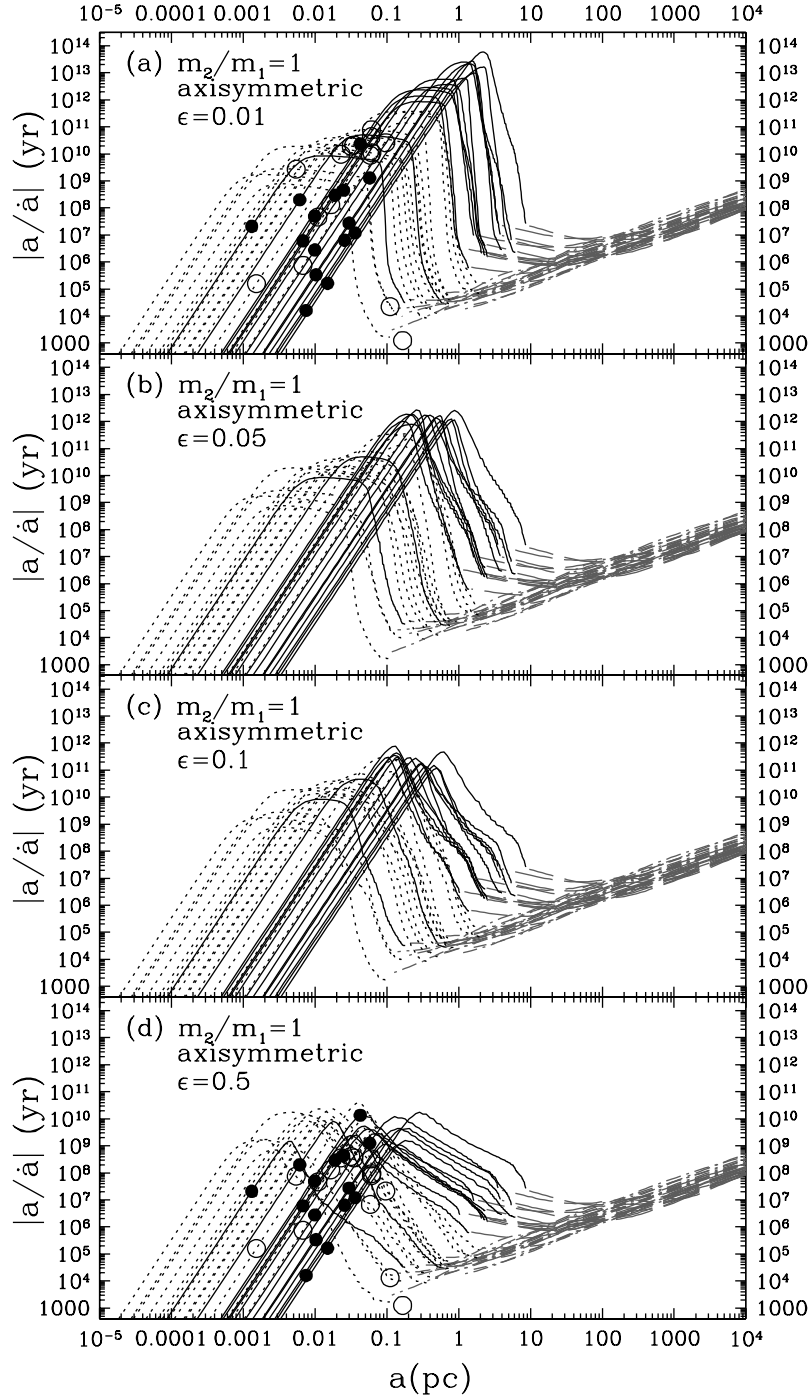
Figure 7. BBH evolution time-scales for axisymmetric galaxies with $m_2/m_1 = 1$ and varying degrees of flattening ϵ. The curves have the same meanings as those in Fig. 2. In panels (a) and (d), the solid circles (for core galaxies) and the open circles (for power-law galaxies) represent the maximum Brownian motion magnitudes of the BBH centres of mass. The evolution curves before the hard binary stage are the same as those in Fig. 2. The bottleneck is in the hard binary stage. When ϵ is small (panel a), the BBH evolution reduces to the spherical case (cf. Fig. 2a). As ϵ increases, the bottlenecks shift to smaller semimajor axes, and many BBH lifetimes become shorter. The Brownian motion of BBHs does not affect the BBH lifetime in nearly spherical galaxies (e.g. panel a) or in core galaxies (e.g. panels a and d), but may decrease the BBH lifetime in highly flattened power-law galaxies (e.g. panel d). See details in Section 5.4.
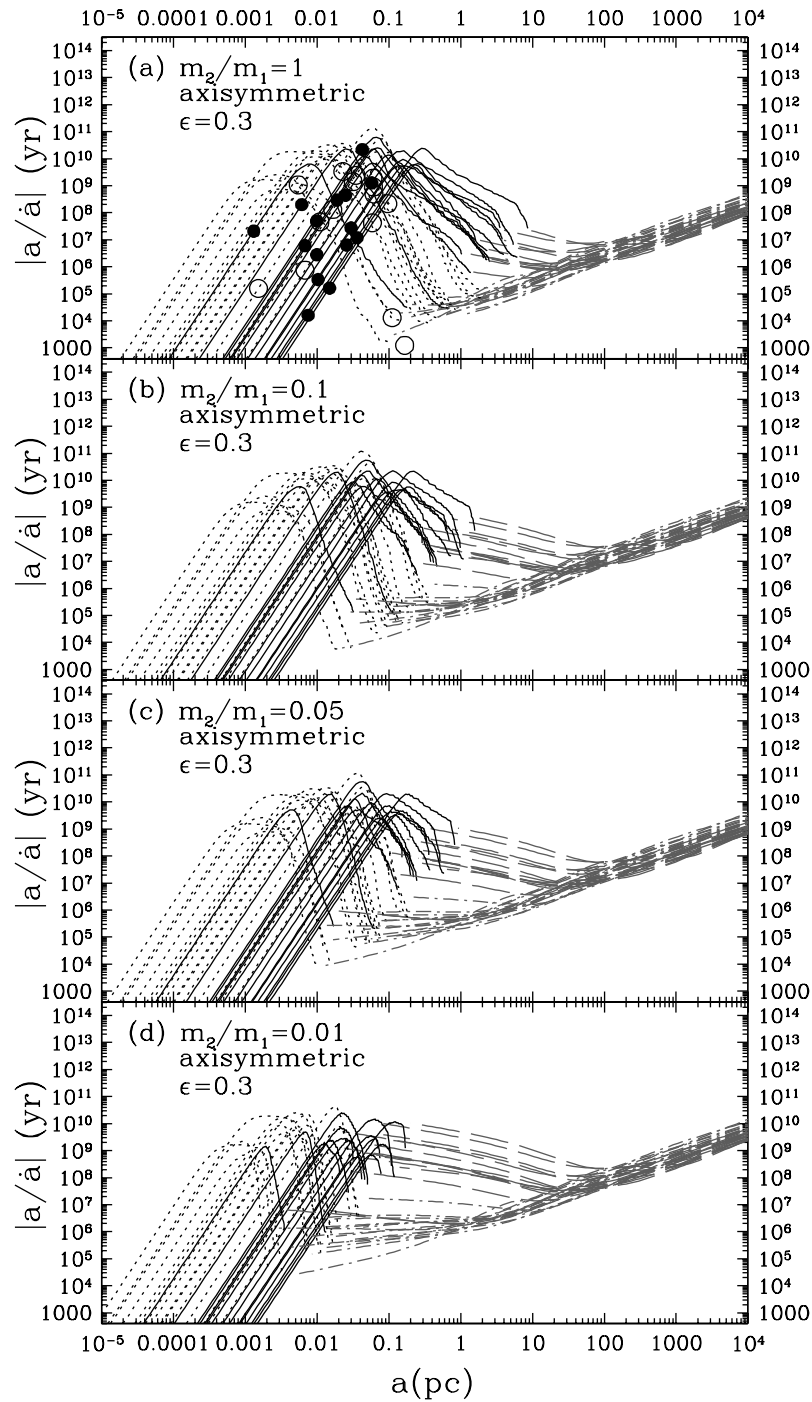
Figure 8. BBH evolution time-scales for axisymmetric galaxies with $\epsilon = 0.3$ and varying mass ratios. The curves have the same meanings as those in Fig. 2. In panel (a), the solid circles (for core galaxies) and the open circles (for power-law galaxies) represent the maximum Brownian motion magnitudes of the BBH centres of mass. As the mass ratio decreases, the bottlenecks shift to slightly smaller radii, and the merger time-scales become shorter. Brownian motion of BBHs does not affect the BBH lifetime in core galaxies, but may decrease the BBH lifetime in power-law galaxies.

5.5 Triaxial galaxies

Using the equations in Sections 3.2 and 4, we obtain the BBH evolution time-scales if the galaxies in Table 1 are triaxial. Here, we consider mainly the BBH with equal masses ($m_2/m_1 = 1$) to see how the evolution depends on the triaxiality ϵ. The trends of the BBH evolution with different m_2/m_1 are similar to those in axisymmetric galaxies (see Fig. 8 or Section 5.4). Noting that triaxial galaxies containing central BHs are likely to evolve

secularly toward axisymmetry within a time T^{trans} (see Section 3.2.3), we will first ignore the evolution toward axisymmetry by setting $T^{\mathrm{trans}} = \infty$, and see the maximum effect of the triaxiality on the BBH evolution. The results are shown in Fig. 9. Then, we will set finite T^{trans} to see the effect of triaxiality on the BBH evolution. The age of galactic triaxiality $T^{\mathrm{tri}}_{\mathrm{age}}$ is assumed to be zero at $a = a_{\mathrm{h}}$.

Fig. 9 shows the BBH evolution in triaxial galaxies without including secular evolution towards axisymmetry. As seen from Fig. 9, the BBH evolution curves in triaxial galaxies follow similar
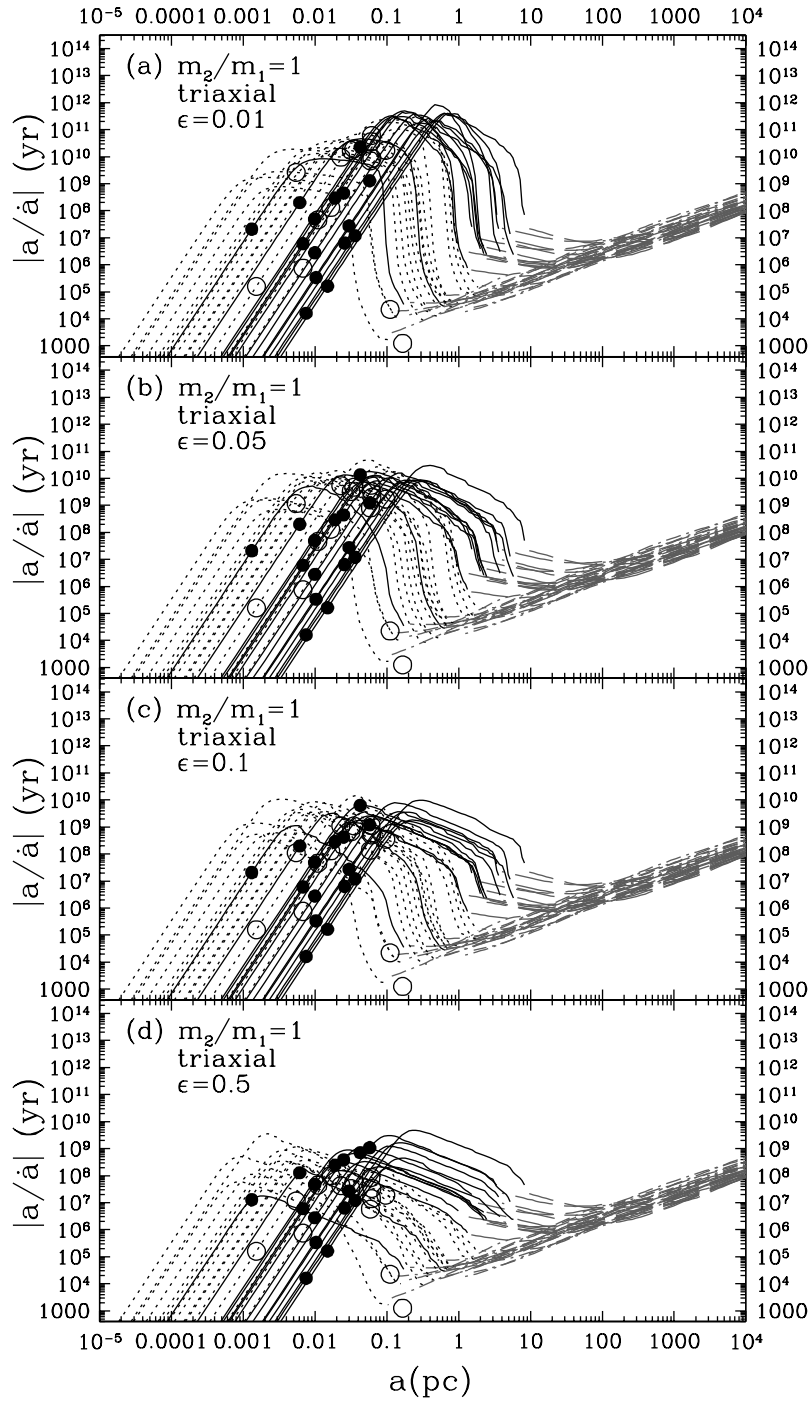
Figure 9. BBH evolution time-scales for triaxial galaxies with $m_2/m_1 = 1$ and varying degrees of triaxiality ϵ. Secular evolution towards axisymmetry is not included. The curves have the same meanings as those in Fig. 2. The solid circles (for core galaxies) and the open circles (for power-law galaxies) represent the maximum Brownian motion magnitudes of the BBH centres of mass. Comparing with spherical galaxies and axisymmetric galaxies (Figs 2 and 7), the draining of the loss regions in triaxial galaxies can significantly decrease the BBH lifetime. Almost all the BBHs can merge within a Hubble time when ϵ is large (panels c and d). The lifetime of BBHs in the power-law triaxial galaxies with large ϵ (e.g. $\gtrsim 0.05$ in panels b–d) may be decreased by their Brownian motion.

trends to those in spherical or axisymmetric galaxies, and the effects of triaxiality on the evolution time-scales also depend on the value of ϵ as in axisymmetric galaxies. If ϵ is small, the characteristic angular momentum J_s is not significantly larger than J_{lc}, and the BBH evolution would be similar to that in spherical or axisymmetric cases. For example, when $\epsilon = 0.01$ (Fig. 9a), most of the BBH evolution time-scales of the power-law galaxies are like those in the spherical or axisymmetric galaxies, and the bottleneck

stages are controlled by the stellar diffusion to the loss cones/wedges. When ϵ is large ($\gtrsim 0.05$, see Figs 9b–d), the difference between the loss region and the loss cone/wedge becomes significant. The loss regions in all the galaxies cannot be depleted before the gravitational radiation stage begins, because there are many more stars in the loss regions $J < J_s$ than in the loss cones/wedges. Most of the BBH lifetimes in Fig. 9(b) ($\epsilon = 0.05$) are comparable to or shorter than the Hubble time, and almost all

the BBHs in Figs 9(c)–(d) ($\epsilon = 0.1$ and 0.5) can merge within a Hubble time. As in axisymmetric galaxies, only the lifetime of BBHs in power-law triaxial galaxies with large ϵ ($\epsilon \gtrsim 0.05$ in panels b–d) might be decreased by their wandering motion.

After including secular evolution towards axisymmetry, the effect of triaxiality on the BBH evolution depends on the value of T^{trans}. For example, if $T^{\mathrm{trans}} = 10^2 P(\mathcal{E})$, after the BBHs become hard, they first follow similar evolution to those in Fig. 9; then with evolution towards axisymmetry, they evolve as those in axisymmetric galaxies (Fig. 7), and the BBHs have similar lifetimes to those in axisymmetric galaxies. If T^{trans} is smaller, the BBH evolution curves would be close to those in axisymmetric galaxies at earlier time; if T^{trans} is larger, they would be more like those in Fig. 9. Actually, if $T^{\mathrm{trans}} = 10^3 P(\mathcal{E})$, the BBH evolution curves in triaxial galaxies with large ϵ (e.g., $\gtrsim 0.05$) are quite similar to those in Figs 9(b)–(d).

In the discussion below on BBHs in triaxial galaxies, we will always use the results in Fig. 9, which gives the maximum effect of triaxiality on BBH evolution.

5.6 Generalization to a distribution of stellar masses and radii

In the calculations above, we assume a single stellar mass of $1\,M_\odot$ and a single stellar radius of $1\,R_\odot$. In this subsection we will see that generalizing to a distribution of stellar masses and radii does not significantly affect our results.

First, generalizing to a range of stellar mass does not affect the dynamical friction time-scales (see equations 39–42), which is affected by the mass density, not the mass of a single star. The value of Λ may be sensitive to the stellar mass, but the friction time-scale depends only logarithmically on Λ.

Second, in the hard binary stage, the BBH hardening time-scale is controlled by the mass clearing rate from the loss region (see equation 9). If the loss region is not depleted, generalizing to a range of stellar mass does not affect the BBH hardening time-scale, since the total mass cleared from the loss region per unit time is not changed. Only when the evolution is controlled by stellar diffusion to the loss region is the BBH evolution time-scale affected through the variation of two-body relaxation diffusion coefficient with stellar mass (see equation 24). According to equation (A6) in MT, the diffusion coefficient of two-body relaxation at constant mass density $\bar{\mu} \propto \int_{m_1}^{m_2} m_*^2 n(m_*)\,\mathrm{d}m_*$, where $n(m_*)$ is proportional to the probability of finding a star with mass $m_* \to m_* + \mathrm{d}m_*$ and $\int_{m_1}^{m_2} n(m_*)\,\mathrm{d}m_* = M_\odot$. Thus, if we use a Salpeter mass function $n(m_*) \propto m_*^{-2.35}$ with limits $m_1 = 0.08\,M_\odot$ and $m_2 = 1\,M_\odot$, the diffusion coefficient is reduced by a factor of 0.31 compared to a population of stars of mass $1\,M_\odot$. Therefore, when the stellar diffusion to the loss cone or loss wedge controls the BBH evolution time-scales, the evolution time-scales will be increased by only a factor of 3 or so, which does not affect our qualitative conclusions.

Third, if the loss region is depleted in the non-hard binary stage, the evolution of the BBH after the depletion and before the hard binary stage is controlled by both dynamical friction from distant stars and three-body interactions with low-angular-momentum stars diffused from high-angular-momentum orbits, which further reduces the effects of the dependence of the diffusion coefficient on the stellar mass distribution.

Finally, the range of stellar radii has relatively little effect on the collision radius r_{coll} (equation 25). Thus we still have $r_{\mathrm{coll}} \ll r_{\mathrm{peak}}$ in Fig. 5, and the BBH evolution time-scale is not affected much (see Section 5.2).

Since generalizing to a distribution of stellar masses and radii does not significantly affect the BBH evolution time-scales, we will continue to use the results with a single stellar mass and radius (obtained in Sections 5.3–5.5) to study the properties of surviving BBHs in Section 6.
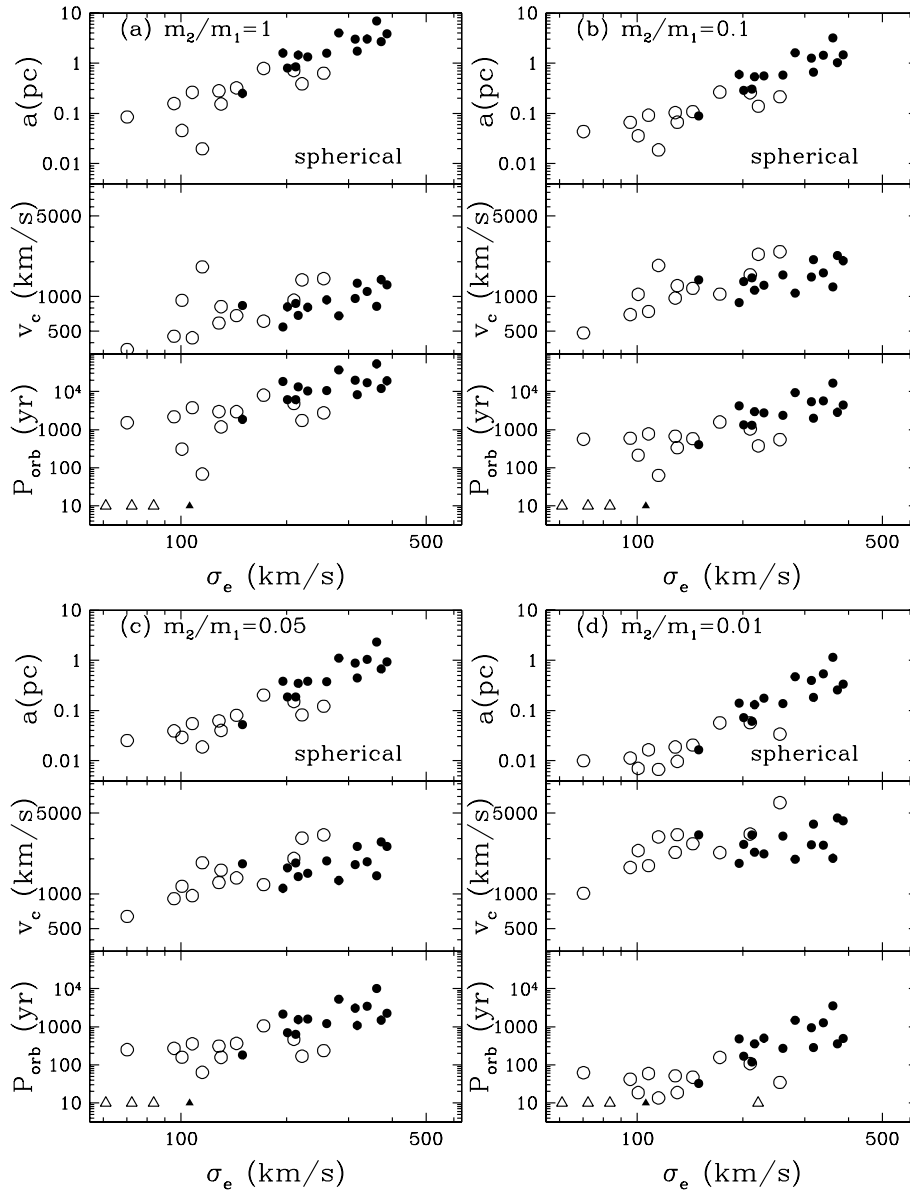
6 PROPERTIES OF SURVIVING BBHS

As seen from Section 5, BBHs can survive for a Hubble time in spherical galaxies, axisymmetric galaxies or some galaxies with small triaxiality ($\epsilon \lesssim 0.05$). Whether or not some of the various astronomical phenomena that have been associated with BBHs [e.g., double nuclei, double-peaked Balmer lines (Gaskell 1996); quasi-periodic radio, optical, X-ray or γ-ray variability (Sillanpää et al. 1988; Rieger & Mannheim 2000), etc.] can be observed depends on the current orbital properties of the BBHs, such as their semimajor axis a, circular speed $v_{\mathrm{c}} = \sqrt{GM_\bullet/a}$, orbital period $P_{\mathrm{orb}} = 2\pi a/v_{\mathrm{c}}$, etc.

Fig. 10 shows the relation between the galactic velocity dispersions σ_{e} and the BBH orbital parameters at the point where $|a/\dot{a}| = 10^{10}$ yr as a decreases for the first time in Fig. 2 (spherical galaxies). As seen from Fig. 10, the BBH orbital parameters depend on galactic velocity dispersions and BH mass ratios. The BBHs in most of the galaxies with $\sigma_{\mathrm{e}} \lesssim 90\,\mathrm{km\,s^{-1}}$ have merged. The semimajor axes and orbital periods of surviving BBHs are generally in the range 10^{-3}–10 pc and 10–10^5 yr; they are generally larger in high-dispersion galaxies (core galaxies with high central BH masses) than in low-dispersion galaxies (power-law galaxies with low central BH masses), and larger for BBHs with equal BH masses than for BBHs with unequal masses. The orbital velocities of surviving BBHs are generally in the range 10^2–$10^4\,\mathrm{km\,s^{-1}}$; they are generally larger in high-dispersion galaxies than in low-dispersion galaxies, and larger for BBHs with unequal masses than for BBHs with equal masses. The surviving BBH properties estimated above may help to identify appropriate methods to probe BBHs in different spherical galaxies.

(i) Double nuclei associated with BBHs should be easier to observe in high-dispersion or luminous galaxies than in low-dispersion or faint galaxies, because BBH semimajor axes are much larger in high-dispersion galaxies. The BBH sample identified from double nuclei would have a bias toward high m_2/m_1 (say, $\gtrsim 0.1$) BBHs for at least three reasons: (1) the two components have comparable brightness if the stellar or non-stellar luminosity is correlated with the BH mass, (2) the BBH semimajor axes are generally larger if the BH masses are equal, and (3) BBHs with very low m_2/m_1 ratios (say, $\lesssim 0.001$) may not exist in the galactic centre, since the smaller BH may not sink to the galactic centre in a Hubble time (see Fig. 2d) (here we consider only the BBHs formed by mergers of galaxies, and we do not consider the possible low-mass BH formation caused by some mechanism at galactic centres).

(ii) Broad-line regions (BLRs) associated with the two components of a BBH in AGNs may produce double-peaked emission lines (e.g., Balmer lines) varying with time-scales of the BBH orbital period, although the double peaks are not easily recognized for $v_{\mathrm{c}} \lesssim \sigma_{\mathrm{BLR}}$ where σ_{BLR} is the velocity dispersion of the clouds in BLRs (typically 3000–$5000\,\mathrm{km\,s^{-1}}$). Although the BBH orbital period can be as short as 10 yr for $m_2/m_1 = 0.01$ (Fig. 10d), for such an extreme mass ratio one component of double peaks would be too weak to be discernible compared to the other strong component. The double peaks from the BLRs of
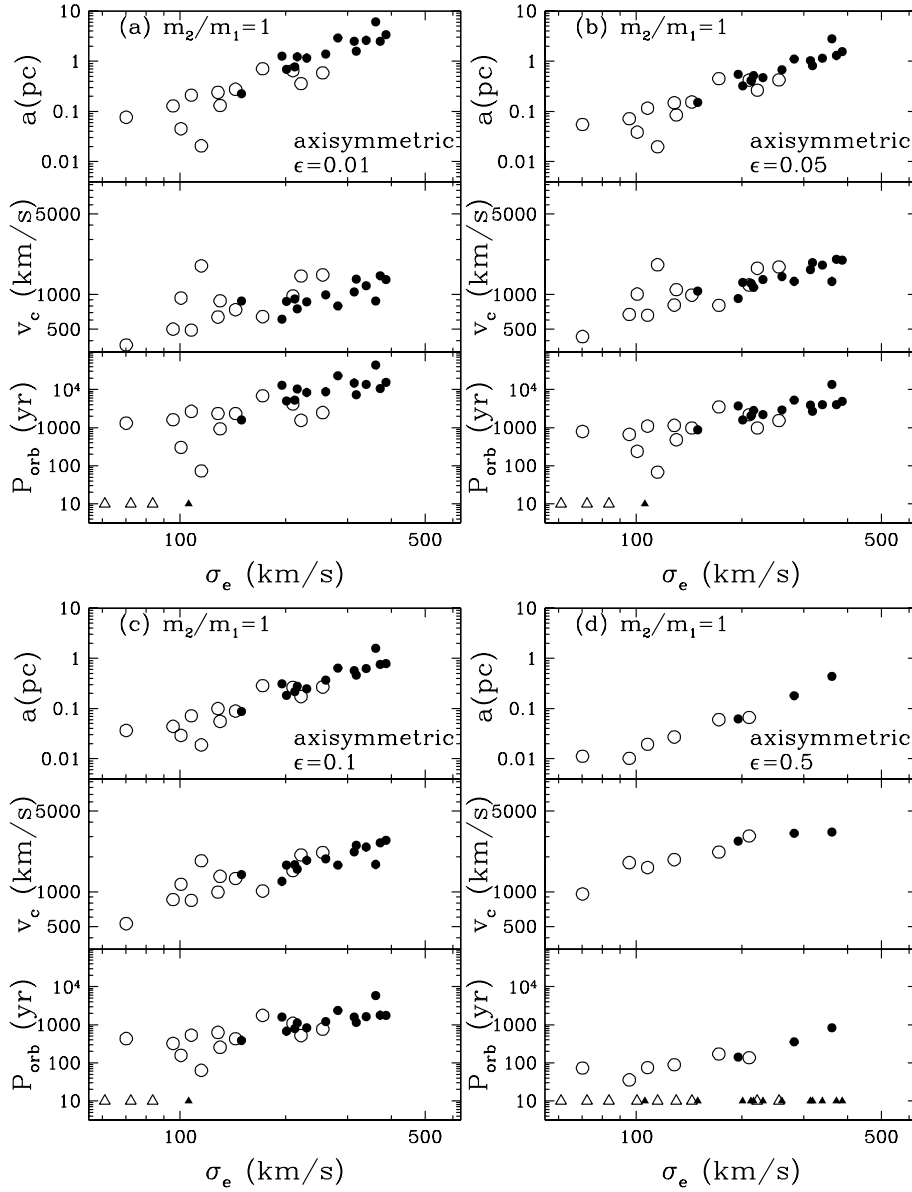
Figure 10. Estimated observational properties of surviving BBHs in spherical galaxies with varying mass ratios. The panels show the relations between the galactic velocity dispersion σ_e and the BBH semimajor axes a, circular speeds $v_c = \sqrt{GM_\bullet/a}$ and orbital periods $P_{orb} = 2\pi a/v_c$ at the point where $|a/\dot{a}| = 10^{10}$ yr for the first time in Fig. 2. The solid circles represent the core galaxies, and the open circles represent the power-law galaxies. The solid triangles (core galaxies) and open triangles (power-law galaxies) located at the bottom of each panel represent the dispersion of the galaxies in which the BBH evolution time-scales are less than the Hubble time of 10^{10} yr. BBHs in most of the low-dispersion ($\sigma_e \lesssim 90\,\mathrm{km\,s^{-1}}$) galaxies have merged within a Hubble time, and BBHs are more likely to survive in the galaxies with $\sigma_e \gtrsim 90\,\mathrm{km\,s^{-1}}$ (similarly in Figs 11–13 below). The semimajor axes and orbital periods of surviving BBHs are larger in high-dispersion galaxies than in low-dispersion galaxies (similarly in Figs 11–13 below) and larger for BBHs with equal masses than for BBHs with unequal masses (similarly in Fig. 12 below). The circular speeds are larger in high-dispersion galaxies than in low-dispersion galaxies, and larger for BBHs with unequal masses than for BBHs with equal masses (similarly in Fig. 12 below). Note that BBHs with very low mass ratios (say, $\lesssim 0.001$) arising from mergers of galaxies would not exist at galactic centres, because the smaller BH cannot sink into the galactic centre within a Hubble time (cf. Section 5.3.1).

high-m_2/m_1 BBRs would have comparable strengths, but it is hard to detect the line variability within a reasonable time for high-m_2/m_1 BBHs, since P_{orb} is at least 10^2–10^3 yr (see the case $m_2/m_1 = 0.1$ in Fig. 10b) even in low-dispersion galaxies ($\sigma_e \lesssim 150\,\mathrm{km\,s^{-1}}$). So, double-peaked emission lines from BLRs are unlikely to show periodic behaviour due to the long orbital period in BBH galaxies. For example, the quasar 3C 390.3 was introduced as a BBH candidate based on data from only ~20 years, which is much less than the estimated BBH orbital period $P_{orb} \sim 300$ yr. The radial velocity curve of ~20 yr for one peak of

emission line Hβ in 3C 390.3 is consistent with the expectation from the orbital motion of a BBH (Gaskell 1996), but the results from continuous observations of ~10 more years deviate from the expectation from a BBH model (Eracleous et al. 1997).

(iii) Some periodic behaviour observed in the radio, optical, X-ray or γ-ray light curves is possibly related to a BBH creating jet(s) aligned nearly along the line of sight or interacting with disc(s). The observed period is usually identified with the orbital period of the BBH or a fraction of the orbital period. For example, if a relativistic jet emerges from the less massive BH and is aligned
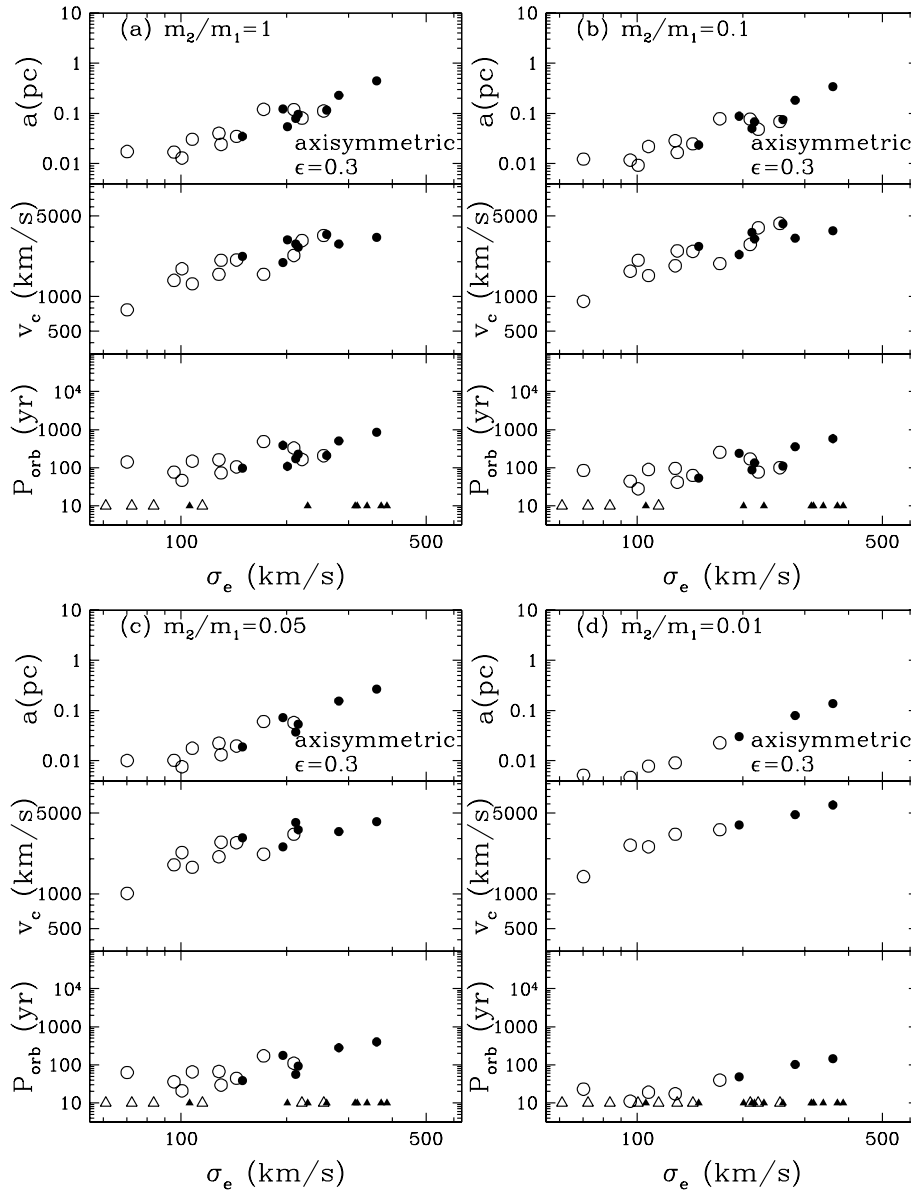
Figure 11. Estimated observational properties of surviving BBHs in axisymmetric galaxies with $m_2/m_1 = 1$ and varying degrees of flattening ϵ (see Fig. 7). The symbols have the same meanings as those in Fig. 10. The relations between the surviving BBH orbital properties and the galactic dispersion follow similar trends to the relations in spherical galaxies (Fig. 10). Most BBHs in highly flattened galaxies can merge within a Hubble time (panel d). The semimajor axes and the orbital periods of surviving BBHs are larger in nearly spherical galaxies than in highly flattened galaxies, while the circular speeds are larger in highly flattened galaxies than in nearly spherical galaxies (similarly in Fig. 13 below). In power-law or low-dispersion galaxies with large ϵ (e.g., panel c or d), Brownian motion of BBHs may further decrease the semimajor axis (at the same time, decrease the orbital period and increase the circular speed) shown in the figure or make the BBHs merge within a Hubble time (similarly in Figs 12–14 below).

nearly along the line of sight, the flux of the X/γ-ray radiation emitted from the jet may vary with the periodic change of the Doppler-boosting factor (due to the slight change of the jet inclination angle with the BH orbital motion) (Rieger & Mannheim 2000). Because of the same mechanism in the phenomenon of 'superluminal motion' (i.e., the observed time-scale is shorter than the intrinsic time-scale; see Rees 1966), the observed X/γ-ray variability time-scale is shorter than the orbital period (say, $\sim10^{-2}P_{\rm orb}$ in Rieger & Mannheim 2000), which can be identified within a much shorter time (~0.1–1 yr) in low-dispersion galaxies with $m_2/m_1 \sim 0.01$ ($P_{\rm orb} \sim 10$–100 yr when $\sigma_{\rm e} \lesssim 150\,{\rm km\,s}^{-1}$ in Fig. 10d). It is hard to use other variability phenomena with time-scales $\sim P_{\rm orb}$ to search for BBHs within a short time (say, <50 yr)

except for low-m_2/m_1 (~0.01) ones in low-dispersion galaxies ($\sigma_{\rm e} \lesssim 150\,{\rm km\,s}^{-1}$).

In axisymmetric or weakly triaxial ($\epsilon \lesssim 0.05$) galaxies (Figs 11–13), the relations between the orbital parameters of surviving BBHs and the properties of their host galaxy follow similar trends to the relations in spherical galaxies (Fig. 10), although the normalization may be different. The orbital properties of surviving BBHs are in the similar range as those in spherical galaxies, and they depend on the flattening or triaxiality parameter ϵ, as well as the galactic velocity dispersion and BH mass ratio. As seen from Fig. 11 (BBHs with mass ratio $m_2/m_1 = 1$ in axisymmetric galaxies), with increasing flattening parameter ϵ, BBH semimajor
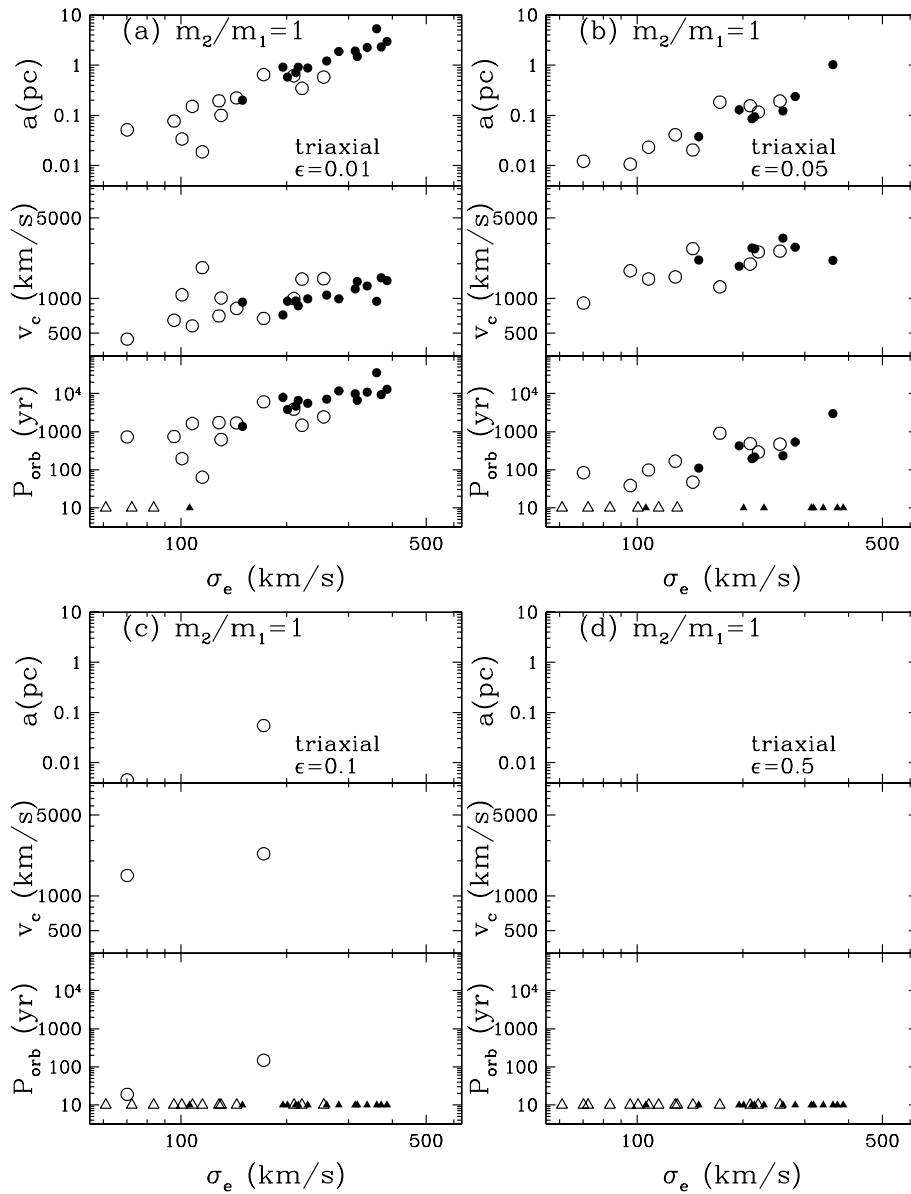
Figure 12. Estimated observational properties of surviving BBHs in axisymmetric galaxies with $\epsilon = 0.3$ and varying mass ratios (see Fig. 8). The symbols have the same meanings as those in Fig. 10. The relations between the surviving BBH orbital properties and mass ratios follow similar trends to those in spherical galaxies (Fig. 10). With decreasing BH mass ratios, more and more BBHs can merge within a Hubble time.

axes decrease (the upper limit is ~ 10 pc for $\epsilon = 0.01$ and ~ 1 pc for $\epsilon = 0.5$), which may require higher resolution to resolve double nuclei; BBH orbital periods decrease ($\sim 10^2 - 10^5$ yr for $\epsilon = 0.01$ and $\sim 50 - 10^3$ yr for $\epsilon = 0.5$), which may help in the identification of periodicities in X/γ-ray variability within a short time ($\sim 0.5 - 10$ yr if the time-scale of periodicities $\sim 10^{-2} P_{\text{orb}}$), but the orbital periods are still so long that the periodic variability of double-peaked emission lines from BLRs will be difficult to identify. As seen from Fig. 12 (flattened galaxies with $\epsilon = 0.3$), with decreasing m_2/m_1 ratios, the BBH semimajor axes and the orbital periods decrease as in spherical galaxies. When m_2/m_1 decreases to 0.01 (Fig. 12d), many BBHs (22 out of our sample of 30 galaxies) have merged within a Hubble time; most of the surviving BBHs in low-dispersion galaxies ($\sigma_e \lesssim 150$ km s^{-1}) have an orbital period $\sim 10 - 30$ yr, and could possibly be found by identifying some variability phenomenon associated with the orbital motion. It is difficult to observe BBHs even with equal

masses in more triaxial galaxies ($\epsilon \gtrsim 0.1$), because most of them have merged within a Hubble time (Figs 13c and d). In addition, the random motion of the centre of mass of BBHs might affect their observational properties in low-dispersion axisymmetric galaxies (with $\epsilon \gtrsim 0.1$) or low-dispersion triaxial galaxies (with $\epsilon \sim 0.05$). On the one hand, the random motion may decrease their semimajor axes and orbital periods, which would make variability phenomena a more efficient tool to search for BBH candidates; on the other hand, it may also make more BBHs merge within a Hubble time and decrease the number of surviving BBHs, which would decrease the probability of identifying BBHs by variability phenomena.

Note that the double-peaked emission lines from BLRs and radio, optical, X/γ-ray variability that we have discussed above are phenomena seen in the relatively rare and distant AGNs, not the common inactive nearby galaxies where BHs are known to exist.

If future observations do not reveal any phenomena associated

Figure 13. Estimated observational properties of surviving BBHs in triaxial galaxies with $m_2/m_1 = 1$ and varying degrees of triaxiality ϵ (see Fig. 9) and without including secular evolution towards axisymmetry. The symbols have the same meanings as those in Fig. 10. Except for some BBHs in weakly triaxial galaxies (e.g., $\epsilon = 0.01$ or 0.05 in panel a or b), most BBHs in triaxial galaxies can merge within a Hubble time.

with BBHs in AGNs, that might suggest that (i) the inner stellar distribution in AGNs is different from normal nearby early-type galaxies, (ii) gas in AGNs plays an important role in BBH orbital decay (e.g. Gould & Rix 2000), or (iii) nuclear activity switches on soon after a galaxy merger and the lifetime of nuclear activity is much less than the Hubble time, so the surviving BBHs in AGNs would not have had time to spiral together before the nuclear activity ceases.

To see if double nuclei associated with BBHs formed by mergers of galaxies can be observed with current telescope resolution, we show the predicted BBH semimajor axes *in arcsec* versus the galactic velocity dispersion σ_e for the nearby inactive galaxies (Table 1) in Fig. 14. The BBHs are assumed to have equal BH masses, which have larger semimajor axes than those with unequal masses. As seen from Fig. 14, spherical or nearly spherical and high-dispersion galaxies have surviving BBHs with the largest semimajor axes, which are just within the *HST* resolution, 0.1 arcsec.

7 DISCUSSION AND CONCLUSIONS

We have studied the evolution of BBHs in realistic galaxy models obtained from a sample of nearby early-type galaxies. We calculated the BBH evolution time-scales due to stellar interactions, taking into account the refilling of the loss cone by two-body relaxation and by tidal forces in non-spherical galaxies. The evolution of BBHs depends on BH mass ratio and host galaxy type. BBHs with low mass ratios (say, $m_2/m_1 \lesssim 0.001$) are only rarely formed by mergers of galaxies, because the dynamical friction time-scale is too long for the smaller BH to sink into the galactic centre within a Hubble time (cf. Fig. 3). In spherical galaxies, if the BBH has equal BH masses, low-angular-momentum stars are generally depleted quickly after the BBH becomes hard, and thereafter the BBH lifetime is controlled by the stellar diffusion rates into the loss cone (see Figs 1a, 2a and 4); if the BBH has unequal masses (so that the BBH becomes hard at
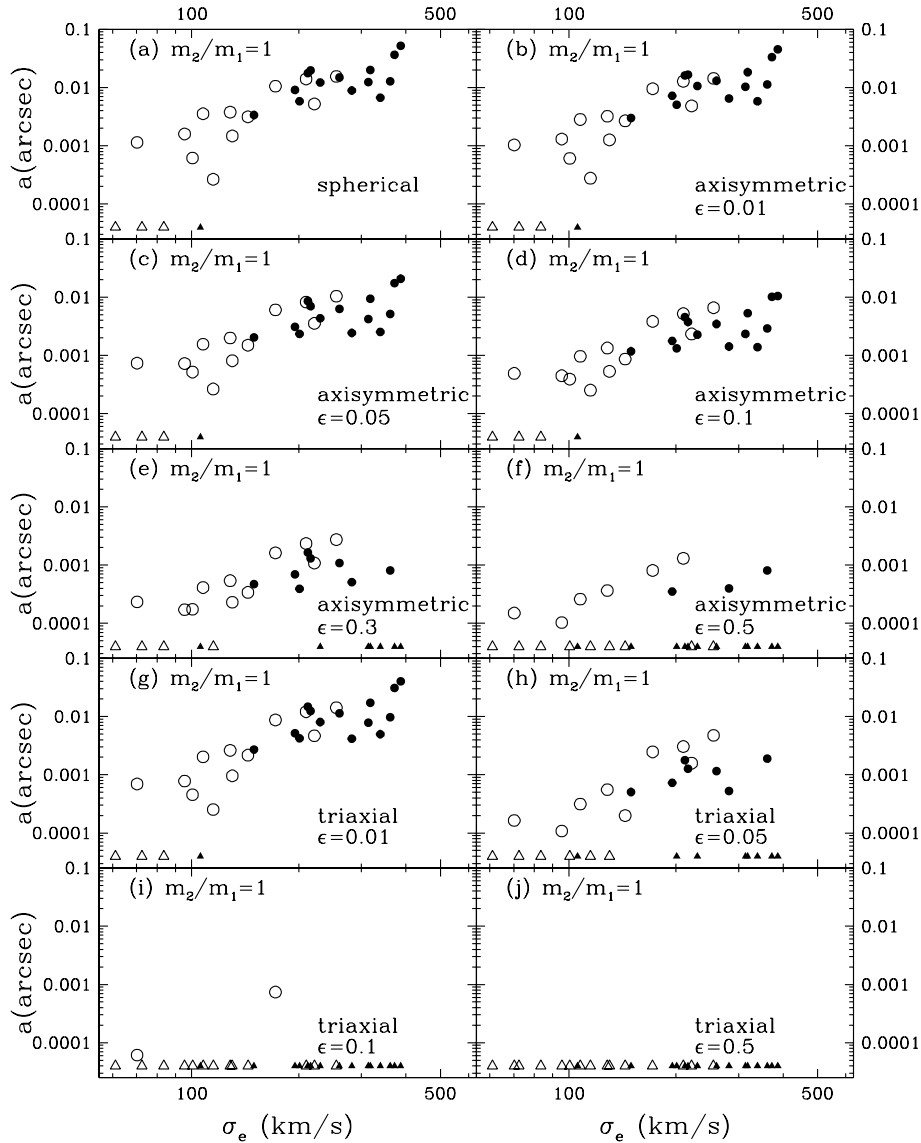
Figure 14. Estimated semimajor axes *in arcsec* of surviving BBHs (with $m_2/m_1 = 1$) versus galactic velocity dispersion σ_e. The symbols have the same meanings as in Fig. 10. The galaxies are assumed to be spherical in panel (a), axisymmetric in panels (b)–(f), and triaxial (without including secular evolution towards axisymmetry) in panels (g)–(j). The upper limit of the BBH semimajor axes is just within the *HST* resolution, 0.1 arcsec.

smaller semimajor axis), low-angular-momentum stars can be depleted before the BBH becomes hard, and possibly gravitational radiation can dominate the BBH orbital decay in the non-hard binary stage and the BBH lifetime decreases (see Figs 1a and 2b–d). Low-dispersion galaxies have smaller central BHs and higher central densities, and hence their BBHs generally have shorter hardening time-scales (associated with interactions with stars) in the hard binary stage (see equation 9) and shorter lifetimes (see Fig. 2). In axisymmetric and triaxial galaxies, many stars can process on to the low-angular-momentum orbits by tidal forces and decrease the lifetime of BBHs (see Figs 1b–f and 7–9). The wandering of the BBHs is generally not important, but might decrease the lifetime of BBHs in some axisymmetric or triaxial galaxies with low velocity dispersion (see Figs 7–9). Note that in the study, stars escaping from the core are considered as removed from the loss cone, which is usually a plausible assumption; however, in some circumstances, even a small fraction of stars returning to the loss cone could enhance the BBH decay rate, and this deserves further investigation.

Most of BBHs in the galaxies with velocity dispersion $\sigma_e \lesssim 90\,\mathrm{km\,s^{-1}}$ have merger time-scales shorter than a Hubble time (see Figs 11–13), and so do BBHs in highly flattened ($\epsilon \gtrsim 0.5$) or moderately triaxial ($\epsilon \gtrsim 0.05$) galaxies (see Figs 11d and 13b–d). BBHs with low m_2/m_1 (say, ~0.01) ratios, which have shorter lifetimes than the BBHs with equal masses, can also merge within a Hubble time in axisymmetric or triaxial galaxies (see Fig. 12d). Surviving BBHs are generally in the galaxies with velocity dispersion $\sigma_e \gtrsim 90\,\mathrm{km\,s^{-1}}$. Spherical, axisymmetric or weakly triaxial galaxies are all likely to have surviving BBHs, especially those with equal masses (see Figs 10–12 and 13a–b). We also estimated the observational properties of surviving BBHs. The semimajor axes and orbital periods of surviving BBHs are generally in the range 10^{-3}–10 pc and 10–10^5 yr; they are generally larger in high-dispersion galaxies than in low-dispersion galaxies, larger in spherical galaxies than in non-spherical galaxies, and larger for BBHs with equal BH masses than for BBHs with unequal masses. The orbital velocities of surviving BBHs are generally in the range 10^2–$10^4\,\mathrm{km\,s^{-1}}$; they are

generally larger in high-dispersion galaxies than in low-dispersion galaxies, larger in non-spherical galaxies than in spherical/nearly spherical galaxies, and larger for BBHs with unequal masses than for BBHs with equal masses.

In short, our study shows that BBHs arising from galaxy mergers are likely to have merged in low-dispersion or highly flattened/triaxial galaxies, but should still survive in spherical/nearly spherical and high-dispersion galaxies. The study in this paper may help to further explore the merger history of massive BHs.

The BBH mergers driven by stellar dynamics are also simply discussed by Gould & Rix (2000). They conclude that mergers driven by stellar dynamics are nearly impossible in a Hubble time and present a mechanism by which gas can drive BH mergers. However, their conclusion is based on much simpler galaxy models and dynamics than those used in this paper.

We have also discussed methods to detect surviving BBHs. For inactive galaxies, currently the only method to probe BBHs is to search for double nuclei at galactic centres. It would be easier to find double nuclei associated with BBHs in luminous (or high-dispersion), nearly spherical galaxies, and the BBHs found would have a bias toward equal BH masses. For active galaxies, we can also use periodic variability phenomena to search for BBHs. Within a short time (e.g., 1 month – 1 yr), it would be easier to find X/γ-ray periodic variability associated with the BBH orbital motion in faint or low-dispersion (e.g., $\sigma_e \lesssim 150\,\mathrm{km\,s^{-1}}$, but $\gtrsim 90\,\mathrm{km\,s^{-1}}$ since most BBHs have merged in the galaxies with $\sigma_e \lesssim 90\,\mathrm{km\,s^{-1}}$) galaxies than in luminous or high-dispersion (e.g., $\sigma_e \gtrsim 150\,\mathrm{km\,s^{-1}}$) galaxies. Here, the variability time-scale is usually a fraction of the BBH orbital period (e.g., 0.01 times of the BBH orbital period), and the BBHs found would have a bias toward low m_2/m_1 (e.g., ~ 0.01) ratios. By other variability phenomena with time-scales of the BBH orbital period, it is also easier to find the low-m_2/m_1 BBHs in faint or low-dispersion ($\sigma_e \lesssim 150\,\mathrm{km\,s^{-1}}$, but $\gtrsim 90\,\mathrm{km\,s^{-1}}$) galaxies, but it would take at least $\sim 10 – 100$ yr. It is hard to find BBHs by double-peaked emission lines from BLRs, either because it takes a long time ($\gtrsim 50$ yr) to identify the periodic variability of double peaks with time-scales of BBH orbital periods, or because the strength of one component is too weak to separate it from the other one.

The upper limit of the semimajor axes of surviving BBHs for all of the galaxies in our study is ~ 10 pc, which is just within the *HST* resolution, 0.1 arcsec, or several pc for typical galaxies in the sample used in this paper. None of these nearby common inactive galaxies has shown signs of surviving BBHs (close double nuclei with separation 0.1 arcsec) at their centres. It is not yet clear whether the absence of detected BBHs in nearby galaxies presents a problem, and if so whether the problem is with the calculations in this paper, the hierarchical structure formation model, or the observational detection techniques.

REFERENCES

Bahcall J. N., Wolf R. A., 1976, ApJ, 209, 214
Begelman M. C., Blandford R. D., Rees M. J., 1980, Nat, 287, 307
Binney J., Tremaine S., 1987, Galactic Dynamics. Princeton Univ. Press, Princeton
Byun Y. et al., 1996, AJ, 111, 5
Cohn H., Kulsrud R. M., 1978, ApJ, 226, 1087
Dehnen W., 1993, MNRAS, 265, 250
Eracleous M., Halpern J. P., Gilbert A. M., Newman J. A., Filippenko A. V., 1997, ApJ, 490, 216
Faber S. M. et al., 1997, AJ, 114, 1771
Ferrarese L., Merritt D., 2000, ApJ, 539, L9
Folkner W. M., ed., 1998, AIP Conf. Proc. 456, Laser-Interferometer Space Antenna. AIP, New York
Frank J., Rees M. J., 1976, MNRAS, 176, 633
Gaskell C. M., 1996, ApJ, 464, L107
Gebhardt K. et al., 2000, ApJ, 539, L13
Gould A., Rix H., 2000, ApJ, 532, L29
Heggie D. C., 1975, MNRAS, 173, 729
Lightman A. P., Shapiro S. L., 1977, ApJ, 211, 244
Magorrian J., Tremaine S., 1999, MNRAS, 309, 447 (MT)
Magorrian J. et al., 1998, AJ, 115, 2285
Merritt D., Ferrarese L., 2001, ApJ, 547, 140
Merritt D., Quinlan G. D., 1998, ApJ, 498, 625
Milosavljević M., Merritt D., 2001, ApJ, 563, 34
Peters P. C., 1964, Phys. Rev. B, 136, 1224
Quinlan G. D., 1996, New Astron., 1, 35
Quinlan G. D., Hernquist L., 1997, New Astron., 2, 533
Rauch K., Ingalls B., 1998, MNRAS, 299, 1231
Rauch K., Tremaine S., 1996, New Astron., 1, 149
Rees M. J., 1966, Nat, 211, 468
Rieger F. M., Mannheim K., 2000, A&A, 359, 948
Shapiro S. L., 1985, in Goodman J., Hut P., eds, Proc. IAU Symp. 113, Dynamics of Star Clusters. Reidel, Dordrecht, p. 373
Sillanpää A., Haarala S., Valtonen M. J., Sundelius B., Byrd G. G., 1988, ApJ, 325, 628
Touma J., Tremaine S., 1997, MNRAS, 292, 905
Tremaine S. et al., 1994, AJ, 107, 634
Yu Q., Lu Y., 2001, A&A, 377, 17

This paper has been typeset from a T$_{\rm E}$X/L$^{\rm A}$T$_{\rm E}$X file prepared by the author.