MERGERS OF UNEQUAL-MASS GALAXIES: SUPERMASSIVE BLACK HOLE BINARY EVOLUTION AND STRUCTURE OF MERGER REMNANTS

FAZEEL MAHMOOD KHAN1,2, MIQUEL PRETO1, PETER BERCZIK3,4,7, INGO BERENTZEN1,5, ANDREAS JUST1, and RAINER SPURZEM3,6,7

1 Astronomisches Rechen-Institut, Zentrum für Astronomie, University of Heidelberg, Mönchhof-Strasse 12-14, 69120 Heidelberg, Germany
2 Department of Physics, Government College University (GCU), 54000 Lahore, Pakistan
3 National Astronomical Observatories of China, Chinese Academy of Sciences, 20A Datun Rd., Chaoyang District, 100012 Beijing, China
4 Main Astronomical Observatory, National Academy of Sciences of Ukraine, 27 Akademika Zabolotnogo St., 03680 Kyiv, Ukraine
5 Heidelberg Institute for Theoretical Studies, Schlossolfsbrunnenweg 35, 69118 Heidelberg, Germany
6 The Kavli Institute for Astronomy and Astrophysics at Peking University (KIAA), Yi He Yuan Lu 5, Hai Dian Qu, Beijing 100871, China

Received 2011 September 27; accepted 2012 February 7; published 2012 April 3

ABSTRACT

Galaxy centers are residing places for supermassive black holes (SMBHs). Galaxy mergers bring SMBHs close together to form gravitationally bound binary systems, which, if able to coalesce in less than a Hubble time, would be one of the most promising sources of gravitational waves (GWs) for the Laser Interferometer Space Antenna. In spherical galaxy models, SMBH binaries stall at a separation of approximately 1 pc, leading to the “final parsec problem” (FPP). On the other hand, it has been shown that merger-induced triaxiality of the remnant in equal-mass mergers is capable of supporting a constant supply of stars on the so-called centrifugal orbits that interact with the binary and thus avoid the FPP. In this paper, using a set of direct N-body simulations of mergers of initially spherically symmetric galaxies with different mass ratios, we show that the merger-induced triaxiality is also able to drive unequal-mass SMBH binaries to coalescence. The binary hardening rates are high and depend only weakly on the mass ratios of SMBHs for a wide range of mass ratios \( q \). There is, however, an abrupt transition in the hardening rates for mergers with mass ratios somewhere between \( q \approx 0.05 \) and 0.1, resulting from the monotonic decrease of merger-induced triaxiality with mass ratio \( q \), as the secondary galaxy becomes too small and light to significantly perturb the primary, i.e., the more massive one. The hardening rates are significantly higher for galaxies having steep cusps in comparison with those having shallow cusps at centers. The evolution of the binary SMBH leads to relatively shallower inner slopes at the centers of the merger remnants. The stellar mass displaced by the SMBH binary on its way to coalescence is \( \sim 1–5 \times \) the combined mass of binary SMBHs. The coalescence timescales for SMBH binary with mass \( \sim 10^6 M_\odot \) from less than 1 Gyr and for those at the upper end of SMBH masses \( 10^9 M_\odot \) are \( 1–2 \) Gyr for less eccentric binaries whereas they are less than 1 Gyr for highly eccentric binaries. SMBH binaries are thus expected to be promising sources of GWs at low and high redshifts.

Key words: black hole physics – galaxies: evolution – galaxies: interactions – galaxies: nuclei – gravitational waves

Online-only material: color figures

1. INTRODUCTION

Supermassive black holes (SMBHs) are now a well-established component of galaxies having a sizable bulge (Ferrarese & Ford 2005). In the paradigm of the Λ cold dark matter cosmology, galaxies are formed via hierarchical merging. If more than one of these merging galaxies contain an SMBH, the formation of a binary SMBH is almost inevitable (Begelman et al. 1980). There are cases in which there is a clear observational evidence for two widely separated SMBHs as well as some circumstantial evidence for a true SMBH binary (Komossa 2006). Upon coalescence, SMBH binaries would be the highest signal-to-noise ratio sources of gravitational waves (GWs), detectable from essentially any cosmological redshift, for future space-borne GW detectors such as the Laser Interferometer Space Antenna (LISA; Hughes 2003; Barack & Cutler 2004; Danzmann et al. 2011).

The paradigm for SMBH binary evolution, after a merger of gas-poor galaxies, consists of three distinct phases (Begelman et al. 1980). First, the two SMBHs sink toward the center due to the dynamical friction exerted by the stars and the dark matter. Dynamical friction acts together with “gravitational slingshot effect” as SMBHs form a bound pair with semi-major axis \( a \approx r_b \), where \( r_b \) is the binary’s influence radius, which typically is defined to be the radius that encloses twice the mass of the binary in stars, \( M(<r_b) = 2 M_* \), where \( M_* \) is the mass of the binary. Dynamical friction stops being an effective driver of inspiral when the binary reaches the hard binary separation \( a \approx a_h \) (Quinlan 1996; Yu 2002)

\[
a_h := \frac{G \mu_r}{4 \sigma^2},
\]

(1)

where \( \mu_r \) is the binary’s reduced mass and \( \sigma \) is the one-dimensional velocity dispersion. In the second phase, slingshot ejection of stars is the dominant mechanism for taking energy and angular momentum away from SMBH binary. Third, the binary eventually reaches a separation \( a_{GW} \) at which the loss of orbital energy due to GW emission drives the final coalescence. The transition from the first to the second phase is prompt provided that the mass ratio of the remnants is not too small \( q \geq 0.1 \) (Callegari et al. 2011). In contrast, the subsequent transition from the second to the third phase could constitute a bottleneck for the binary evolution toward final coalescence caused by lack of stars that can interact with the binary. This is the so-called final parsec problem (FPP).
Yu (2002) early pointed out, based on an analysis based on the Hubble Space Telescope (HST) sample of nearby elliptical galaxies and spiral bulges from Faber et al. (1997), that flattening and non-axisymmetries would help stellar dynamics in bringing SMBH binaries to coalescence. Merritt & Poon (2004) built self-consistent cuspy, triaxial models with a single SMBH at the center and showed that such models could be concocted with a significant fraction of centrophilic orbits that would efficiently drive the hardening rate of a binary if it were present at the center. Berczik et al. (2006) showed, for the first time with N-body simulations of rotating King models, that the FPP could potentially be solved by purely stellar dynamical models that develop a significant amount of triaxiality. By including the post-Newtonian ($\mathcal{P}$N) terms into the equations of motion of the binary, Berentzen et al. (2009) explicitly showed that the coalescence times, if the latter models are scaled to real galaxies, are clearly shorter than a Hubble time. Preto et al. (2009) presented preliminary results suggestive that merger-induced triaxiality resulting from the merger of spherically symmetric models would lead to qualitatively similar results—namely, the prompt coalescence of resulting SMBH binaries. Khan et al. (2011, hereafter Paper I) and Preto et al. (2011, hereafter Paper II) confirmed and generalized the latter results by showing that galaxies that form via equal-mass mergers are mildly triaxial. They have shown that hardening rates of SMBH binaries resulting from equal-mass mergers are substantially higher than those found in spherical nuclei, and are independent of the total number $N$ of stars, thus allowing them to extrapolate results to real galaxies. Presumably, the nonspherical shapes of merger remnants reported in these studies result in a large population of stars on centrophilic orbits. Inside the influence radius of SMBH, the centrophilic orbit family includes saucer or cone orbits in the axisymmetric potentials (Sridhar & Touma 1999) and pyramids orbits (angular momentum close to zero) in triaxial potentials (Merritt & Vasiliev 2011). Outside the influence radius, most of the centrophilic orbits are chaotic in a triaxial potential (Poon & Merritt 2001).

The inspiral of a binary SMBH is expected to leave a characteristic imprint in the morphological and dynamical properties of the newly formed galactic nucleus following the merger: e.g., the bending and precession of radio jets (Roos et al. 1993), the variability in optical light curve of quasar (Valtonen et al. 2008), X-shaped radio lobes (Merritt & Ekers 2002) and, with some likelihood, a partially depleted core instead of a steep cusp in bright elliptical galaxies (Graham 2004; Kormendy et al. 2009), which could be explained by different models of SMBH binary evolution prior or after coalescence (Merritt et al. 2007; Gualandris & Merritt 2008).

Detailed surface photometry of ellipticals has revealed that their surface brightness is well described by a Sérsic profile, $\log I(r) \propto r^{1/\gamma}$, over most of the body of a bulge or early-type galaxy. However, systematic deviations from this profile are found in the innermost regions close to the central SMBH—either in the form of a light (mass) excess or deficit with respect to that which would result from the extrapolation of the Sérsic profile toward the center (Kormendy et al. 2009). In other words, observations seem to reveal a dichotomy in the central density profiles of bulges and early-type galaxies: giant, high-luminosity objects have relatively shallow central density profiles, while normal, low-luminosity ones show steeper density profiles in their center. The former are often called core galaxies and show density profiles with central logarithmic slopes $\gamma < 1$, and typically harbor SMBH of mass $\gtrsim 10^6 M_\odot$, while the latter are normally called power-law galaxies, have central $\gamma > 1$, and lighter SMBH $\lesssim 10^5 M_\odot$. The explanation of cores is non-trivial since dry mergers should preserve the highest density in phase space given that, by virtue of being devoid of appreciable amounts of gas, their dynamics are essentially collisionless (Binney & Tremaine 2008).

Galaxies having a single mass stellar population are expected to have Bahcall & Wolf (1976) cusp around the SMBH with $r(r) \propto r^{-\alpha}$, $\alpha \sim 7/4$ for $r \lesssim r_h$ after a relaxation time (Preto et al. 2004). In the more realistic case where there is a range of stellar masses, less massive objects follow a shallow profile, $\alpha \sim 3/2$, while heavier objects (e.g., stellar black holes) develop steeper cusps $\alpha \sim 2$ around the SMBH (Bahcall & Wolf 1977; Alexander & Hopman 2009). Whether a given, or the typical, nucleus has reached this relaxed state or not depends on its “initial” profile and on how long ago it was formed. Different studies have reached different conclusions regarding the most probable state of the stellar distribution around the SMBH after a relaxation time. Preto & Amaro-Seoane (2010), who performed Fokker–Planck and N-body simulations of a two-component galaxy nucleus model made of solar mass stars and stellar mass black holes, obtained mass-segregated stellar cusps in $\sim 1/4$ of a relaxation time. However, Gualandris & Merritt (2012) state roughly the opposite conclusion: Even after a few central relaxation times, the distribution of stars can be very different than in the steady-state, mass-segregated solutions.

The analysis of the number counts of spectroscopically identified, old stars in the sub-parsec region of our own Milky Way (Buchholz et al. 2009; Do et al. 2009) seems to favor the presence of a very shallow profile for the visible fraction of the old stellar population. This is consistent with the finding that the two-body relaxation time at the influence radius of the SMBH in the center of the Milky Way is longer than a Hubble time (Merritt 2010; Preto & Amaro-Seoane 2010). On the other hand, low-luminosity spheroids often contain nuclear star clusters that have central relaxation times shorter than a Hubble time (Merritt 2009).

In dry mergers, i.e., in the absence of a dynamically significant amount of gas, the slingshot ejection of stars by the central SMBH binary drives the binary inspiral. It has early been proposed that inspiraling binaries could carve a core inside $\sim r_h$, thus providing an explanation for the mass deficits present in the center of gas-poor giant ellipticals (Makino & Ebisuzaki 1996). On energetic grounds it can be shown that in order for the binary to reach coalescence, it needs to eject an amount of stellar mass of order of its own mass, $M_{\text{ej}} \sim \alpha \times M_\bullet$, where $\alpha = O(1)$ and $M_\bullet = M_{\bullet 1} + M_{\bullet 2}$ (Merritt & Milosavljević 2005; Perets & Alexander 2008). Merritt (2006a) studied the mass deficits created by SMBH inspirals evolving in spherical symmetric nuclei and found an average $\alpha \sim 0.5$. However, unless their mass $M_\bullet \lesssim \text{few} \times 10^6 M_\odot$, the evolution of SMBH binaries in gas-poor spherical galaxies tends to stall due to depletion, on the (local) dynamical timescale, of the pool of stars whose orbits intersect the binary (e.g., Milosavljević & Merritt 2003; Makino & Funato 2004; Berczik et al. 2005; Khan et al. 2011; Preto et al. 2011). The binaries studied by Merritt (2006a) did not reach coalescence due to the FPP. Therefore, they estimated mass deficits at the time when a hard binary forms in their simulations. Merritt et al. (2007) extended the latter study by following the inspiral of binaries up to coalescence in the lower mass range using an approximate Fokker–Planck description for stellar scattering. They find substantially higher mass deficits than before, albeit restricting their calculations to
spherical models and very shallow cusps (with initial $\gamma \sim 1/2$) in strong contrast with the larger $\gamma$ typical of compact low-mass nuclei. More precise estimates of $\alpha$ from more realistic galaxy merger studies are therefore certainly needed; in particular, ones that do not rely on any of the following approximations: spherical symmetry, low $\gamma$, crude model for inferring the mass ejected from the increase in the binary’s orbital energy, or those generally inherent to the Fokker–Planck formalism. In this respect, preliminary studies of mass deficits created by the merger of equal-mass galaxies, with moderate SMBH mass ratio, moderate central slope $\gamma \sim 1$, and that generate mildly triaxial remnants, suggest that the resulting mass deficits could be substantially higher $M_{\text{ej}} \sim 2 M_\odot$ within $\lesssim 3 r_0$ (Preto et al. 2009), a result that we confirm and generalize in this paper. Note, however, that alternative scenarios for explaining the formation of cores have been proposed (Gualandris & Merritt 2008).

In this paper, we revisit these problems by detailed $N$-body simulations using a range of more realistic models of merging galactic nuclei with varying mass ratio $q$ and central logarithmic slopes $\gamma$. We restrict the simulations to $q = M_{\text{gal},s}/M_{\text{gal},p} = M_\odot/M_\odot$, where $M_\text{gal},s$ and $M_\text{gal},p$ are the spin mass and the primary mass, respectively. We study the shape of the merger remnant of galaxies having different density profiles and different mass ratios. For each of our models, the hardening rate of the SMBH binary, the axis ratios of the merger remnant, and the resulting mass deficit are calculated. In Section 2, we describe our initial models and numerical methods used in our study. The evolution of SMBH binary in galaxy mergers is described in Section 3, as well as the time evolving densities and shapes of newly formed galaxies resulting from the galaxy merger. Section 4 describes the analysis and results of “mass deficits” resulting from the galaxy mergers. The coalescence timescales for binary SMBHs from the time of formation of binary until the full coalescence of SMBHs due to emission of GWs are presented in Section 5. Finally, Section 6 summarizes and discusses the results of the paper.

2. INITIAL CONDITIONS AND NUMERICAL METHODS

2.1. The Host Galaxies and Their SMBHs

Our aim is to study the dependence of the hardening rates with the mass ratios and central concentration of the nucleus, as well as the resulting imprint of the SMBH inspiral on the stellar distributions of post-coalescence nucleus. Closely following the initial setup of our previous numerical experiments (see Papers I and II), we represent the individual galaxies or galactic nuclei by spherically symmetric $N$-body models. The density profile of the models follows the Dehnen (1993) profile of the form

$$\rho_\text{D}(r) = \frac{(3-\gamma)M_{\text{gal}}}{4\pi r_0} \frac{r_0}{r^3(r + r_0)^{4-\gamma}},$$

(2)

where $M_{\text{gal}}$ is the total mass of the galaxy, $r_0$ is the scale radius, and $\gamma$ is the inner logarithmic slope. For $r \gg r_0$, the density declines with a logarithmic slope of $-4$, which, although not close to the exterior density of real galaxies, provides a convenient cutoff that avoids the waste of CPU time in following stars whose orbits are too far from the SMBH binary to have any significant impact on its long-term evolution. The resulting cumulative mass profile is given by

$$M_\text{D}(r) = M_{\text{gal}} \left( \frac{r}{r + r_0} \right)^{3-\gamma}.$$

(3)

In addition, we represent the SMBHs by massive particles with zero velocity placed at the center of both the primary and the secondary galaxies. The masses of the SMBHs are set to be 0.5% of their host galaxy. This value is few times higher than the currently accepted value of $M_{\text{gal}}/M_\odot \sim 0.001$ (Ferrarese & Merritt 2000; Gebhardt et al. 2000; Merritt & Ferrarese 2001; Haring & Rix 2004). However, some galaxies may have larger values of this ratio than others.

Thus, the two SMBHs in the simulations have the same mass ratio $q$ as their host galaxies. The Dehnen models have an isotropic distribution function of velocities (no net rotation), are spherically symmetric, and are initially set up to be in dynamical equilibrium with a gravitational potential given by $\Phi(r) = -GM_{\text{gal}}/r + \Phi_\odot(r)$, where $\Phi_\odot(r)$ is the gravitational potential due to the stars alone.

The galaxies size ratio scales with the corresponding mass ratio as $R_\text{e}/R_1 \propto \sqrt{M_2/M_1}$. By choosing four different values of $\gamma = 0.5, 1.0, 1.5$, and 1.75, the central part of our galaxy models represent the variety of observed density profiles in both early- and late-type galaxies.

For Dehnen models, the ratio of the effective radius (projected half-mass radius) $R_\text{e}$ to the half-mass radius $r_{1/2}$ is given approximately as

$$\frac{R_\text{e}}{r_{1/2}} \approx 0.75,$$

(4)

and shows only a weak dependence on $\gamma$. In the following, we refer to the more massive galaxy as the primary galaxy and the lighter one as the secondary. In our model units, we use for our primary galaxy a total mass and scale radius of $M_{\text{gal},p} = r_0,p = 1$. We also set the gravitational constant to $G = 1$.

In the absence of relativistic effects on the binary’s orbit, our models are scale free and can be a posteriori scaled to galaxies of different total mass and size. However, the inclusion of radiation reaction effects on the binary due to GW emission introduces an absolute scale associated to the universal value of the speed of light. Such scalings are described in detail in Section 5 when coalescence times are explicitly computed.

The effective radii of our models can be compared to the effective radii in observed galaxies to calculate $r_0$. By fixing galaxy mass $M_{\text{gal}}$, our models can be scaled to physical units using following relations:

$$[T] = \left( \frac{G M_{\text{gal},p}}{r_{0,p}^3} \right)^{-1/2} \approx 1.5 \text{ Myr} \left( \frac{M_{\text{gal},p}}{10^{11} M_\odot} \right)^{-1/2} \left( \frac{r_{0,p}}{\text{kpc}} \right)^{3/2},$$

(5)

$$[V] = \left( \frac{G M_{\text{gal},p}}{r_{0,p}} \right)^{1/2} \approx 655 \text{ km s}^{-1} \left( \frac{M_{\text{gal},p}}{10^{11} M_\odot} \right)^{1/2} \left( \frac{r_{0,p}}{\text{kpc}} \right)^{-1/2}. $$

(6)

We have performed a series of $N$-body experiments where we vary the mass ratio of the galaxies and SMBHs $q$ and the inner density slope $\gamma$. In Table 1, we summarize the model parameters and properties for the primary galaxies used in our simulations. The computational wallclock time increases with increasing cuspiness (higher $\gamma$) of the density profile because particles in the center need very small numerical time steps in order to resolve accurately the (locally) strongly varying gradients of
The Astrophysical Journal, 749:147 (14pp), 2012 April 20

Khan et al.

Table 1
Model Parameter and Properties of the Primary Galaxy

| Model | N    | γ    | r$_{1/2}$ | R$_{\text{eff}}$ |
|-------|------|------|-----------|------------------|
| A     | 128k | 0.5  | 3.13      | 2.35             |
| B     | 128k | 1.0  | 2.41      | 1.81             |
| C     | 128k | 1.5  | 1.69      | 1.27             |
| D     | 64k  | 1.75 | 1.35      | 1.01             |

Note. Columns from left to right show the model name, the particle number N, the half-mass radius r$_{1/2}$, and the effective radius R$_{\text{eff}}$, respectively, for our primary galaxy models.

the gravitational potential. In order to keep the computational time within reasonable limits, we built our model D (highest γ = 7/4) with only half as many particles as the other models.

The secondary galaxies used in our merger simulations have the same density profile, i.e., the same γ, as the primary galaxies, but have different masses $M_{\text{gal},i}$ and scale radii $r_{0,i}$. Both primary and secondary galaxy models have equal numbers of particles N.

The initial center of mass positions and velocities for the two galaxies are calculated from the Keplerian orbit of the equivalent two-body problem, with given apo- and pericenters, $r_a$ and $r_p$, respectively. The two stars start at the apocenter of their relative orbit. The initial separation between the center of mass of the two galaxy is $\Delta r = 15r_{0,p}$; and since the half-mass radius of each nucleus is $R_{1/2} \lesssim 2.5$, this choice ensures that the galaxies are initially well separated while, at the same time, we minimize computing time. The initial relative velocity of the two galaxies is chosen such that the SMBH separation at first pericenter passage is $\sim 2r_{0,p}$. In other words, the initial orbit of the binary galaxy has eccentricity $\sim 0.75$, which corresponds to a circularity parameter value of $\epsilon = L/L_c \sim 0.66$ in rough agreement with the values measured from full galaxy scale merger simulations (Kazantzidis 2010).

2.2 Numerical Methods and Hardware

The N-body integrations are carried out using the free GRAPE (Harfst et al. 2007), a parallel, direct-summation N-body code that uses general purpose hardware, namely, Graphic Processing Units (GPUs) cards to accelerate the computation of pairwise gravitational forces between all particles. Our updated version of the free GRAPE (Paral/Hermite Integration with GRAPE) code uses the fourth-order Hermite integration scheme to advance all particles with hierarchical individual block time steps, together with the parallel usage of GPU cards to calculate the acceleration a and its first time derivative $\dot{a}$ between all particles. For the force calculation, we use the APPORO library (Gaburov et al. 2009) which emulates on the GPU the standard GRAPE-6 library calls.

As free GRAPE does not include the regularization (Mikkola & Aarseth 1998) of close encounters or binaries, we have to use a gravitational (Plummer-)softening between all components. The softening length is chosen to be sufficiently small (i.e., $5 \times 10^{-5}$ in our model units) to keep the (dense) stellar system as collisional as possibleecessary. In the few runs in which we include the PN terms in the equations of motion of the SMBH binary, the softening between the SMBH particles is set equal to zero. Our implementation of the PN terms is identical to that used in Berentzen et al. (2009), except that we cutoff the PN expansion at the next order. We stress that the binary is, in all cases, integrated taking into account the gravitational field from the stellar cluster.

The N-body integrations were carried out on three computer clusters: “titan,” at the Astronomisches Rechen-Institut in Heidelberg, the GPU-enhanced cluster “kolo” at the University of Heidelberg, and “laohu” in Beijing.

For a detailed discussion on the N-body methods and choice of parameters, we refer the reader to Paper I.

3. SMBH Binary Evolution in Galaxy Mergers

Figure 1 shows the relative separation R between the two black holes during a galaxy merger. According to the equivalent two-body trajectory used to set up the initial galaxy orbital parameters, the galaxy centers, and hence the SMBHs, should reach a separation of $2r_{0,p}$ during the first pericenter passage. However, as Figure 1 shows, in case $q \gtrsim 1/2$, the SMBH separation shrinks well below $2r_{0,p}$ during the galaxy merger phase. As q decreases, and hence the mass of the secondary galaxy decreases as well, the galaxies reach the first pericenter passage after more closely following the corresponding two-body orbit. In the same figure, the color of the arrows signals the time T at which the two SMBHs in a particular model form a binary system. As q decreases, the arrows move from left to right signaling the longer interval of time between the start of merger and formation of the binary. Neither the time T nor $R_{\text{peri}}$ are as sensitive to γ as they are to q.

When the SMBHs become bound their separation is $R \approx r_h$. Once an SMBH binary is formed, dynamical friction and the ejection of stars by the gravitational slingshot act together to efficiently extract angular momentum from the massive binary and the binary separation shrinks rapidly. When the semi-major axis $a \approx a_0$, this initial rapid phase of binary hardening comes to an end as the pool of stars whose orbits cross the binary orbit gets depleted on the (local) dynamical timescale.$^9$

$^9$ The semi-major axis $a$ and eccentricity $\epsilon$ of the binary are defined here via the standard Keplerian relations, i.e., neglecting effects of the field stars.
The subsequent hardening of the binary occurs only if the loss cone orbits are replenished. In a spherical galaxy with no gas, the only dynamical mechanism for replenishing loss cone orbits is two-body relaxation, with the corresponding timescale for repopulating the loss cone orbits scaling as $\sim N/\ln N$ (Binney & Tremaine 2008), and the inspiral rate thus becomes strongly $N$-dependent (Makino & Funato 2004; Berczik et al. 2005). In Papers I and II, it was shown that the hardening phase, in the case of a mildly triaxial remnant nuclei resulting from a galaxy merger, is essentially $N$-independent, which allows our results to be scaled to real galaxies whose typical number of stars, $N_{\text{tar}}$, is always several orders of magnitude larger than the maximum number reachable with state-of-the-art direct $N$-body simulations. Furthermore, it was also shown in Papers I and II that the flux of stars into the loss cone in a triaxial remnant is also consistently higher than in the spherical case, meaning the binary evolution and coalescence can occur much faster than in spherical models.

Figure 2 shows the time evolution of the binary’s inverse semi-major axis $1/a$ from the time $T$ when the two black holes become bound. An initial phase with faster binary hardening becomes more evident for steep inner density profiles ($\gamma = 1.5, 1.75$) — presumably corresponding to the clearing of the original loss cone (Yu 2002). The binary hardening rates $s$ are calculated by fitting straight lines to $a^{-1}(t)$ in the late phase of binary evolution from $T = 60$ to 100. The results are shown in Figure 3. The hardening rates depend very weakly on the mass ratio of binary SMBHs, in contrast with the $\sim M_\bullet$ dependence found in equal-mass merger study (Papers I and II) and, for fixed $M_\bullet$, on the mass ratio $q$. On the other hand, the value of $s(t)$ increases significantly for higher values of $\gamma$. Note that for the single run with the smallest mass ratio in our sample, $q = 0.05$ and $\gamma = 1$, the hardening rates are significantly weaker than the remaining $\gamma = 1$ models.

Following Paper II, the SMBH binary hardening rate, which is a measure for the energy loss of the binary, is given by

$$s(t) \equiv \frac{d}{dt}\left(\frac{1}{a}\right) \approx \frac{2m^*\gamma(C)}{M_\bullet\sigma a} \int_0^{\infty} dE F_{\text{lc}}(E, t),$$

where $F_{\text{lc}}(E, t)$ is stellar flux into the loss cone, $(C) = O(1)$ is a dimensionless quantity obtained from three-body scattering experiments (Quinlan 1996), and $E = GM_\bullet/r + \Phi_\star(r) - 1/2 v^2$ is the (specific) energy of each star ($E > 0$ for bound stars). The behavior of the hardening rates can thus be qualitatively understood as follows. The flux $F_{\text{lc}}(E, t)$, and its time evolution, depends on the degree of symmetry in the underlying gravitational potential (or, equivalently, the orbit families supported by it)—including the radial density distribution of stars around the binary. On the other hand, the flux of stars into the loss cone is expected to peak around $r_h$ (Perets & Alexander 2008). For a given energy $E$ and fixed $M_\bullet$, $F_{\text{lc}}(E, t) \propto n(E, t)/\tau(E, t)$, where $n(E, t)$ is the number of stars of energy $E$ per unit (specific) energy and $\tau(E, t)$ is equal, in the spherical case, to $\tau_{\text{frx}} \propto \sigma^3/\rho \propto r^{-3/2}$ (Binney & Tremaine 2008) or, in the triaxial case, to the star’s orbital precession time $\tau_{\text{prec}}(E, t) \propto M_\bullet^{-1/2}(<r) \propto r^{(15/2)/\gamma}$, for $r \lesssim 2$ $r_h$ (Yu 2002). Note that, for typical nucleus parameters, $\tau_{\text{frx}} \ll \tau_{\text{prec}}$. As a result, and since for more concentrated nuclei (higher $\gamma$)

\[J \lesssim J_{\text{lc}} = \sqrt{GM_\bullet/r_{\text{frx}}}, \text{ where } f = O(1) \text{ (Lightman & Shapiro 1977).}\]
Evolution of SMBH binary semi-major axis, in $N$-body integrations of galaxy mergers for $\gamma = 0.5$, 1.0, 1.5, and 1.75 according to Table 2. $T$ measures the time since the instant when the SMBHs formed a bound pair.

(A color version of this figure is available in the online journal.)

Average hardening rates for $\gamma = 0.5$, 1.0, 1.5, and 1.75 according to Table 2. The average is measured between $T = 60$ and $T = 100$ in $N$-body units (see the text for details).

(A color version of this figure is available in the online journal.)

Ratio of intermediate to major ($b/a$, upper series of points) and minor to major ($c/a$, lower series of points) axes for $\gamma = 0.5$, 1.0, 1.5, and 1.75 according to Table 2.

(A color version of this figure is available in the online journal.)

$n(E_h, t)$ is higher and $\tau(E, t)$ is shorter, more concentrated nuclei experience higher $F_{lc}$ and higher $s(t)$.

The triaxiality also plays an important role in determining the hardening rate. This is because of its role in supporting centrophilic orbits; in fact, at fixed $n(E, t)$ only a fraction of the stars will be on centrophilic orbits, and this fraction is an increasing function of increasing triaxiality. As $q$ decreases, the triaxiality of the remnant also decreases as shown in Figure 4. This results presumably in a weaker hardening rate since the fraction of stars on centrophilic orbits is smaller. This effect partially explains why the dependence of the hardening rate on $M_*$ is much weaker than expected based on the results obtained in Papers I and II for equal-mass mergers.

We analyze the shape of the merger remnant by calculating principal axis ratios and density profiles for galaxy merger models. Figure 4 shows the principle axis ratios of the merger remnant; these were defined as the axis ratios of a homogeneous ellipsoid with the same inertia tensor. The departures from spherical symmetry become more apparent for mergers with larger $q$. For $q = 0.05$ and 0.1, the merger-induced triaxiality becomes residual and the remnant results in only a slightly
flattened system. The dynamical effect of this transition seems to be abrupt, as the hardening rates shown in Figure 3 are essentially independent of mass ratio down to \( q \sim 0.1 \). In such minor mergers, the secondary galaxy gets tidally disrupted within a few pericenter passages, and eventually the *naked* black hole of satellite spirals in due to dynamical friction. The evolution of SMBH binary in such a merger gives results very similar to those of binaries embedded in spherical galaxy models. Nevertheless, in reality situations such as these are not very likely to happen as both galaxies would surely retain some triaxiality or flattening from a previous major merger.

The density profiles of the newly formed galaxies as the result of the merging are shown in Figure 5. As a result of the galaxy merger and the ejection of stars from the central galaxy region by the inspiraling black holes, the merger remnant nuclei have significantly shallower inner slopes than those of its progenitors. The “damage” caused by infalling black holes increases with their mass. The values of \( \gamma \) are calculated by fitting Dehnen’s model to the merger remnant right after the time when phase of rapid binary hardening ends. Table 3 presents the final inner slopes for merged galaxies.

4. MASS DEFICITS

Observations of nearby galaxies show that Sérsic functions provide remarkably accurate fits for the major axis brightness surface profiles over the main bodies of bulges and early-type galaxies. This has been confirmed with increased accuracy and dynamic range over the last decade or so (Kormendy et al. 2009). Moreover, detailed state-of-the-art simulations of merging galaxies also generate profiles that are consistent with Sérsic functions (Hopkins et al. 2009a, 2009b). Even though there is not a complete theoretical understanding for the remarkable regularity of these profiles, their apparent generality and robustness led people to identify and interpret departures from Sérsic fits and use them as a probe of the physics underlying the coevolution of galaxies and their massive black holes.

Cores tend to be found in giant ellipticals and are loosely defined as the central region in a bulge or early-type galaxy where the surface brightness deviates and it is below the values that would result from the extrapolation of the Sérsic profile from the main body of the object down to its innermost region. Typically, they are associated to shallower cusps with \( \gamma \lesssim 0.5–1 \). One possible explanation of a central core in a gas-poor galaxy can be attributed to the ejection of stars by a hard SMBH binary. The destruction caused by the inspiraling massive black hole in the center of galaxy can be measured in terms of missing mass called “mass deficit.” Graham (2004) and Ferrarese et al. (2006) measured the mass deficits from the difference in flux between the observed galaxies light profiles at the center and the inward extrapolation from their outer Sérsic profiles. Estimated mass deficits from these studies yield \( M_{\text{def}} \sim (1–2) M_* \), and values as high as \( \sim (4–5) M_* \) were uncommon. More recent studies obtained larger values with an average value \( (M_{\text{def}}/M_*) \sim 11 \) and an error in the mean of about 18% (Kormendy et al. 2009).

The top panel of Figure 6 shows the cumulative mass profile for model C1 at different time of its evolution. The merger of a secondary galaxy with the primary causes some changes in the inner profile of the latter. This can be seen by comparing the profiles at times \( t = 0 \) and \( t = 100 \), between which there is a noticeable drop in the stellar density around \( r_h \) due to the merging of the two galaxies. Then, as the secondary is being tidally disrupted, after a few pericenter passages, the profile of the resulting merger remnant remains stable. Once, at time

**Figure 5.** Density profiles of merger remnant for \( \gamma = 0.5, 1.0, 1.5 \) and 1.75 (top to bottom) according to Table 2.

(A color version of this figure is available in the online journal.)
of the runs. The conclusion is that the derived mass deficit is quite insensitive to the exact final time of the runs, provided this occurs at a time when the binary’s semi-major axis already reached a separation that is orders of magnitude shorter than the influence radius $r_h$.

Ours and the (several) observational definitions of mass deficit are obviously not equivalent, but they are surely somehow related provided the assumption that the mass is ejected by the binary holds true. We will not attempt to dwell on the intricacies of a detailed comparison, but will instead use the mass deficits obtained from our simulations as qualitative indicators regarding the evolution of the galactic nuclei as they undergo minor/major dry mergers.

The bottom panel of Figure 6 shows the resulting mass deficits measured out to various radii for model C1. Mass deficits have a maximum value at around 2–3 $r_h$ for each model. We show the value of mass deficit at different radii in Table 3 for all the runs. First, we can see that the mass deficits are clearly larger for more concentrated (higher $\gamma$) models. They are also increasing with the mass ratio $q$. Note that what we are directly measuring from the runs is the effect of the inspiraling binary on the stellar distribution, not the total amount of mass ejected by the binary. The latter value should, by simple energetic arguments and given a fixed total binary mass, be on average the same regardless of the detailed properties of the surrounding nucleus. The “damage” incurred by the stellar cusps in more concentrated nuclei, however, is expected to be greater, as stars on loss cone orbits at small radii will represent a larger fraction of the (local) total stellar mass than those farther outside.

If galaxies undergo a series of mergers as suggested by the hierarchical galaxy formation scenario, then the mass deficits should add up because the time required to fill in the gap (removal of stars by inspiraling SMBH binary) is essentially of the order of relaxation time, which is several orders of magnitude greater than a Hubble time for bright elliptical galaxies $t \sim 202$, the two black holes become bound and form a binary, the stars inside the loss cone start to be cleared by the binary and a rapid evolution of the mass profile ensues till $t \sim 240$. This drop in the central density is due to the slingshot ejection of stars by the binary. As a result, the mass inside the loss cone drops rapidly, hence changing the mass profile there. Once this phase of rapid mass ejection ends, the density profile—notably inside $r_h$—remains almost constant as the binary continues its inspiral.

The gravitational potential is approximately spherical and thus the stars residing there can enter the loss cone only by diffusing in energy and angular momentum space through two-body encounters. The corresponding timescale is the relaxation time, which scales as $t_{\text{rel}} \sim \sigma^3 / \rho \sim \rho^{-3/2}$ (Spitzer 1987), since $\gamma \lesssim 3/2$ for all models at almost all times of interest, this implies that $t_{\text{rel}}$ increases toward the center. The timescale for such stars to enter the loss cone is thus much longer than the typical precession time needed for centripetal orbits of stars, resident at distances $r_h$, to come sufficiently close to interact with the binary. As a result, the density profiles well inside $r_h$ remain essentially constant during the remainder of the inspiral, so the damage impinged on the cusp by the SMBH inspiral following a merger is less severe than one might expect from naive spherical model scenarios. The conclusion is that it is not very likely that a hole in the stellar distribution will ever be created by such mergers.

We calculate the mass deficits for each of our galaxy mergers as the difference between the stellar mass enclosed at a given radius at the time when the binary first becomes bound and at the end of the runs. The end of the run is a priori a somewhat arbitrary choice since not all runs end at exactly the same point in the binary’s inspiral. However, when we compare the four cases (A2, A4, B3, and B4) for which we followed the binary up to final coalescence, we realize that most of the change in the density profiles happens at a relatively early stage in the evolution of the binary, and this is mostly covered in almost all

| Run | $r_h$ | $\gamma$ | $\sigma_{\text{inl}}$ | $T_{\text{f}}$ | $M_{\text{def, 1}}$ | $M_{\text{def, 1.5}}$ | $M_{\text{def, 2}}$ | $M_{\text{def, 2.5}}$ | $M_{\text{def, 3}}$ |
|-----|-------|--------|-----------------|-------|----------------|----------------|----------------|----------------|----------------|
| A1  | 0.19  | 0.34   | $6.4 \times 10^{-4}$ | 180   | 0.51           | 0.76           | 0.84           | 0.93           | 0.72           |
| A2  | 0.19  | 0.26   | $5.5 \times 10^{-4}$ | 180   | 0.50           | 0.94           | 1.20           | 1.20           | 1.05           |
| A3  | 0.19  | 0.22   | $5.9 \times 10^{-4}$ | 210   | 0.68           | 1.21           | 1.49           | 1.40           | 1.34           |
| A4  | 0.22  | 0.18   | $9.7 \times 10^{-4}$ | 220   | 0.58           | 0.76           | 1.05           | 1.50           | 1.86           |
| A4(\$P/N\$) | 0.22 | 0.15 | Merged | 204   | 0.59           | 0.74           | 1.06           | 1.47           | 1.81           |
| B1  | 0.155 | 0.61   | $2.9 \times 10^{-4}$ | 140   | 0.50           | 1.18           | 1.31           | 1.38           | 1.64           |
| B2  | 0.14  | 0.57   | $3.0 \times 10^{-4}$ | 110   | 0.83           | 1.19           | 1.35           | 1.12           | 1.35           |
| B3  | 0.115 | 0.51   | $4.0 \times 10^{-4}$ | 115   | 1.33           | 1.73           | 2.0            | 2.13           | 1.87           |
| B3(\$P/N\$) | 0.115 | 0.45   | Merged | 168   | 1.69           | 1.75           | 2.10           | 2.30           | 1.91           |
| B4  | 0.13  | 0.48   | $3.9 \times 10^{-4}$ | 230   | 0.80           | 2.0            | 2.9            | 3.50           | 3.26           |
| B4(\$P/N\$) | 0.13 | 0.43 | Merged | 189   | 1.05           | 2.05           | 2.80           | 3.34           | 3.21           |
| C1  | 0.066 | 0.86   | $1.25 \times 10^{-4}$ | 090   | 0.71           | 1.18           | 1.42           | 1.51           | 1.29           |
| C2  | 0.066 | 0.80   | $1.2 \times 10^{-4}$ | 090   | 0.96           | 2.50           | 2.88           | 2.60           | 2.40           |
| C3  | 0.067 | 0.77   | $1.4 \times 10^{-4}$ | 080   | 1.03           | 1.52           | 2.26           | 2.40           | 2.26           |
| C4  | 0.069 | 0.70   | $1.02 \times 10^{-4}$ | 140   | 1.10           | 2.3            | 3.9            | 4.80           | 5.30           |
| D1  | 0.045 | 1.15   | $8.3 \times 10^{-5}$ | 080   | 0.91           | 1.27           | 1.82           | 2.01           | 1.89           |
| D2  | 0.046 | 1.06   | $7.1 \times 10^{-5}$ | 100   | 1.15           | 1.85           | 2.56           | 3.04           | 3.28           |
| D3  | 0.046 | 1.04   | $7.0 \times 10^{-5}$ | 112   | 1.39           | 2.36           | 3.14           | 3.81           | 3.97           |
| D4  | 0.047 | 1.01   | $7.9 \times 10^{-5}$ | 130   | 1.50           | 2.84           | 3.88           | 4.96           | 5.62           |

Note. The columns from left to right represent: (1) galaxy merger model, (2) the influence radius, (3) the inner density slope $\gamma$ for merger remnant, (4) the separation of the two black holes at the end of the run, (5) time at the end of the run in N-body units, and (6–10) mass deficits $M_{\text{def}}$, in units of the mass of the binary $M_\star$ measured within $n \times r_h^2$, $r_h$ is calculated at the time of the formation of the SMBH binary.
In this section, we derive an estimate for the coalescence times of the SMBH binaries in our simulations, similar to what has been done in Berentzen et al. (2009). In the latter work, it has been shown that the predictions of coalescence times for SMBH binaries in rotating, triaxial galactic nuclei agrees well with the results of PN N-body simulations, up to order 2.5PN. To test the accuracy of our estimates made here, we repeat four of our N-body simulations of merging galaxies—this time including the PN equations of motion for the SMBHs up to order 3.5PN (see Section 5.1).

In Figure 2, one can see the rate of change of the binary’s inverse semi-major axis during the hard binary phase is approximately independent of time, so we can write

$$s_{\text{NB}} = -\frac{1}{a^2} \left( \frac{da}{dt} \right)_{\text{NB}} \approx \text{const.}$$  \hspace{1cm} (8)

a value which we can measure directly from each run.

At some time in the simulation, the semi-major axis $a$ will have fallen to such a small value that binary shrinking due to GW emission takes over. At such small separations—depending on mass and eccentricity of the binary—the GW emission becomes so efficient in extracting the angular momentum and energy from the binary that the latter becomes effectively decoupled from the rest of the stellar system and the SMBHs are led eventually to coalescence as if they were an isolated binary.

The orbit-averaged expressions—including the lowest order 2.5PN dissipative terms—for the rates of change of a binary’s semi-major axis and eccentricity due to GW emission are given by Peters (1964):

$$\left( \frac{da}{dt} \right)_{\text{GW}} = -\frac{64}{5} G^3 M_1 M_2 M_* \frac{1}{a^5 c^3} (1 - e^2)^{7/2} \left( 1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \right),$$  \hspace{1cm} (9)

$$\left( \frac{de}{dt} \right)_{\text{GW}} = -\frac{304}{15} G^3 M_1 M_2 M_* \frac{1}{a^5 c^3} (1 - e^2)^{5/2} \left( 1 + \frac{121}{304} e^2 \right).$$  \hspace{1cm} (10)

The corresponding hardening rate due to GW emission alone is thus given as

$$s_{\text{GW}} = \frac{64}{5} G^3 M_1 M_2 M_* \frac{1}{a^5 c^3} (1 - e^2)^{7/2} \left( 1 + \frac{73}{24} e^2 + \frac{37}{96} e^4 \right).$$  \hspace{1cm} (11)

To calculate the coalescence time in our simulations, we divide the evolution into two distinct regimes (see, e.g., Preto et al. 2009): (1) the classical regime, in which the hardening is driven by stellar dynamical effects and (2) the relativistic regime in which the GW emission is dominant. We define this latter regime, starting from a time $t_0$ when

$$s_{\text{NB}} = s_{\text{GW}}.$$  \hspace{1cm} (12)

The semi-major axis $a_0$ at time $t_0$ can be calculated from Equation (12) by assuming that both $s_{\text{NB}}$ and eccentricity $e(t)$ remain roughly constant\footnote{Note that if $e(t)$ (slightly) increases as suggested by scattering experiments by Quinlan (1996) and Paper II, then the full time to coalescence shall be smaller than our estimated $t_{\text{coal}}$. Therefore, our results here should be interpreted as upper bounds to the true coalescence times.} during the stellar dynamical hardening phase, as supported by our simulations.

We find that $a_0$ is typically smaller than $a_{\text{final}}$, the semi-major axis at the time $t_{\text{final}}$ at which our simulations end. Therefore, $t_0$ usually exceeds $t_{\text{final}}$ and the time interval $\Delta t$ between $t_{\text{final}}$ and $t_0$ can be derived as

$$\Delta t = s^{-1}_{\text{NB}} \left( \frac{1}{a_0} - \frac{1}{a_{\text{final}}} \right).$$  \hspace{1cm} (13)
simulations obtained through the procedure described above are very shallow inner slopes and faint ellipticals have steep cusps. D’s are scaled with M32. The choice of galaxies for our models to other galaxy masses by assuming the validity of the models to that of the Galactic center and then extrapolate it to physical units. In Paper I, we equated the effective radii of our models to the effective radius of the Galactic center (or similarly compact nuclei) by assuming the validity of the relation and a stellar distribution \( \rho \). We use the observed mass of the SMBH and its influence radius and compare it to the mass and influence radius of the primary galaxy in our models to fit \( M_{\text{gal}} \) and \( r_0 \). Model A’s, which have a very shallow slope, are scaled with M87, model B’s with NGC4486A, and model D’s are scaled with M32. The choice of galaxies for our models scaling is consistent with the fact that bright ellipticals have very shallow inner slopes and faint ellipticals have steep cusps. The estimated coalescence times for SMBH binaries in our simulations obtained through the procedure described above are presented in Table 5. We find that the time intervals spend in the Newtonian and in the relativistic regime, respectively, are always comparable in almost all our simulations (see Column 7 in Table 5). This can be understood as follows: Binaries with high eccentricities reach the relativistic regimes earlier than those with low eccentricity, thus shortening the stellar dynamical phase. Also the higher eccentricities mean that the GW emission becomes more efficient, which then shortens the relativistic inspiral phase. The black hole of M32 with a mass \( \sim 3 \times 10^6 M_{\odot} \) corresponds to possible sources detectable by \( LISA \). The coalescence timescales for low-mass black hole binaries are less than a Gyr (see Table 5), which suggests that prompt coalescence of binary SMBH detectable by \( LISA \) should be very common at high redshifts. Even at the high-mass end, our models suggest coalescence timescales \( \sim \) Gyr or even less depending on the eccentricity of the binary. These timescales are short enough that SMBH binaries in these galaxies should achieve full coalescence before a subsequent galaxy merger occurs. The coalescence times for SMBHs at the low-mass end are significantly \( \sim 10 \times \) shorter than the ones presented in Paper I and more in line with those obtained in Paper II in which the scaling of lower-mass nuclei was set up by matching the total mass in stars to that observed in the inner \( \sim 2.5 \text{ pc} \) of the Galactic center. The reason for this discrepancy is the weak dependence of effective radii with the galaxy luminosity relation adopted there, which was optimized for nuclei with \( \gtrsim 10^8 M_{\odot} \) black holes, but would lead to an overestimate of the influence radius of the Galactic center (or similarly compact nuclei) by  

### Table 4

| Model | Galaxy | \( M_\bullet \) (\( M_{\odot} \)) | \( r_0 \) (pc) | \( T \) (Myr) | \( L \) (kpc) | \( M \) (\( M_{\odot} \)) | Speed of Light (c) |
|-------|--------|-----------------|---------------|-------------|-------------|-----------------|-----------------|
| A     | M87    | \( 3.6 \times 10^9 \) | 460           | 1.9         | 2.3         | \( 7.2 \times 10^{11} \) | 257             |
| B     | N4486A | \( 1.3 \times 10^7 \) | 31            | 1.4         | 0.3         | \( 2.6 \times 10^9 \)  | 1550            |
| D     | M32    | \( 3.1 \times 10^6 \) | 3             | 0.75        | 0.12        | \( 6.2 \times 10^8 \)  | 2011            |

**Note.** Columns from left to right: (1) galaxy mergers model, (2) semi-major axis (in model units) at the end of simulation \( t_{\text{final}} \), (3) hardening rate (model units), (4) eccentricity at \( t_{\text{final}} \), (5) semi-major axis of the binary at which the stellar dynamical hardening becomes equal to the hardening due to GWs, (6) lifetime of binary in classical stellar dynamical hardening phase, (7) ratio between the time spent in the classical regime and the time spent in the GW regime, and (8) full time to coalescence of SMBHs.

### Table 5

| Run | \( a_{\text{final}} \) (pc) | \( t_{\text{final}} \) (Gyr) | \( \epsilon_0 \) | \( a_0 \) (pc) | \( t_0 \) (Gyr) | \( t_0/ t_{GW} \) | \( t_{\text{coal}} \) (Gyr) |
|-----|-----------------|-----------------|-------------|-------------|-------------|----------------|-----------------|
| A1  | \( 6.4 \times 10^{-4} \) | 9.10            | 0.50        | 3.5 \times 10^{-1} | 1.30        | 2.1            | 1.89            |
| A2  | \( 5.5 \times 10^{-4} \) | 10.8            | 0.98        | 3.9 \times 10^{0}  | 0.12        | 1.1            | 0.23            |
| A3  | \( 5.9 \times 10^{-4} \) | 10.3            | 0.70        | 6.9 \times 10^{-1} | 0.63        | 1.2            | 1.15            |
| A4  | \( 9.7 \times 10^{-4} \) | 9.60            | 0.88        | 1.6 \times 10^{0}  | 0.30        | 1.1            | 0.57            |
| B1  | \( 2.9 \times 10^{-4} \) | 23.3            | 0.62        | 7.4 \times 10^{-3} | 2.10        | 1.3            | 3.70            |
| B2  | \( 3.0 \times 10^{-4} \) | 21.9            | 0.98        | 7.1 \times 10^{-2} | 0.24        | 0.5            | 0.77            |
| B3  | \( 4.0 \times 10^{-4} \) | 20.4            | 0.95        | 4.5 \times 10^{-2} | 0.39        | 1.4            | 0.66            |
| B4  | \( 3.9 \times 10^{-4} \) | 22.2            | 0.96        | 6.3 \times 10^{-2} | 0.27        | 0.7            | 0.64            |
| D1  | \( 9.2 \times 10^{-5} \) | 75.7            | 0.69        | 2.1 \times 10^{-3} | 0.48        | 1.0            | 0.98            |
| D2  | \( 7.1 \times 10^{-5} \) | 69.8            | 0.66        | 2.4 \times 10^{-3} | 0.43        | 1.0            | 0.82            |
| D3  | \( 7.6 \times 10^{-5} \) | 69.5            | 0.61        | 2.6 \times 10^{-3} | 0.41        | 1.4            | 0.70            |
| D4  | \( 7.5 \times 10^{-5} \) | 59.9            | 0.60        | 3.3 \times 10^{-3} | 0.38        | 1.3            | 0.67            |

**Notes.** Columns from left to right: (1) galaxy mergers model, (2) semi-major axis (in model units) at the end of simulation \( t_{\text{final}} \), (3) hardening rate (model units), (4) eccentricity at \( t_{\text{final}} \), (5) semi-major axis of the binary at which the GW regime \( t_0 \) and (2) the binary’s lifetime \( t_{GW} \) until the final coalescence. The lifetime \( t_{GW} \) of an isolated relativistic binary can be calculated from Equations (9) and (10) for a given semi-major axis \( a_0 \) and corresponding eccentricity \( \epsilon_0 \), respectively (see Equation (5.14) in Peters 1964). In order to compute the GW evolution rates, one is required to adopt some physical units, e.g., of length and mass. In Paper I, we equated the effective radii of our models to the effective radii of observed galaxies to fix the mass (\( M_{\text{gal}} \)) and the length (\( r_0 \)) units. In Paper II, we equate the influence radii of our models to that of the Galactic center and then extrapolate it to other galaxy masses by assuming the validity of the \( M_{\bullet} - \sigma \) relation and a stellar distribution \( \rho(r) \propto r^{-7/4} \) at the center. Here, to compute the coalescence times, we select three different elliptical galaxies (M32, M87, NGC4486A) to scale our models to physical units (Table 4). We use the observed mass of the SMBH and its influence radius and compare it to the mass and influence radius of the primary galaxy in our models to fix \( M_{\text{gal}} \) and \( r_0 \). Model A’s, which have a very shallow slope, are scaled with M87, model B’s with NGC4486A, and model D’s are scaled with M32. The choice of galaxies for our models scaling is consistent with the fact that bright ellipticals have very shallow inner slopes and faint ellipticals have steep cusps.
an order of magnitude. Naturally, the coalescence times at high-mass end of binary SMBHs are in nice agreement between both studies.

5.1. Post-Newtonian Simulations

In order to verify the accuracy of our estimate for the coalescence times given in Table 5, we select the two cases A2 and A4 from Table 5 (plus two other cases, B3 and B4, not shown) and re-started these runs with exactly same initial conditions. The choice of these runs is motivated by their relative short coalescence times compared to other runs.

We have implemented the relativistic effects to the MBH binary only by using the $P\mathcal{N}$ equations of motion written in the inertial frame of binary center of mass including all the terms up to $3.5P\mathcal{N}$ order

$$\frac{dv}{dt} = -\frac{GM_\bullet}{r^2} [(1 + A)\mathbf{n}_{12} + B\mathbf{v}_{12}] + \mathcal{O}(1/c^5),$$

(14)

where $\mathbf{n} = r/r$ and the coefficients $A$ and $B$ are complicated expressions of the binary’s relative separation and velocity (Blanchet 2006). The $P\mathcal{N}$ approximation is a power series expansion in $1/c$: the 0th-order term corresponds to the dominant Newtonian acceleration. The $1P\mathcal{N}$-, $2P\mathcal{N}$-, and $3P\mathcal{N}$-order terms are conservative and proportional to $c^{-2}$, $c^{-4}$, and $c^{-6}$. The dissipative $2.5P\mathcal{N}$ and $3.5P\mathcal{N}$ terms, which are proportional to $c^{-2}$ and $c^{-7}$, cause the loss of orbital energy and of angular momentum due to the GW radiation reaction. We treat the SMBHs as point particles and thus we neglect any spin–orbit or spin–spin coupling, which in general is taken into account in the $1.5P\mathcal{N}$ (spin–orbit), $2P\mathcal{N}$ (spin–spin), and $3.5P\mathcal{N}$ (spin–orbit) terms in our common $P\mathcal{N}$ implementation.

Figure 7 shows the evolution of the inverse semi-major axis $1/a$ and eccentricity $e$ of the SMBH binary for run A2 with and without $P\mathcal{N}$ corrections. Hardening of the binary due to GWs emission starts to dominate at $1/a \sim 1000$, which is slightly larger than the estimated $1/a_0$ (all in model units). This is due to the slightly higher value of eccentricity reached during the run without $P\mathcal{N}$, which is then used to estimate $a_0$. For the correct value of eccentricity obtained from run with $P\mathcal{N}$, the estimated $1/a_0$ matches accurately with the value where hardening due to GWs starts to dominate. The full coalescence time for the binary in our $P\mathcal{N}$ simulations is 0.25 Gyr, which is very close to estimated $t_{\text{coal}} \sim 0.23$ Gyr. The inverse semi-major axis and eccentricity evolution for model A4 are shown in Figure 8. If we look at the evolution of the inverse semi-major axis for this run then we see that hardening due to GW becomes important at estimated $1/a_0$, as the average eccentricity evolution matches almost perfectly between the $P\mathcal{N}$ and non-$P\mathcal{N}$ runs.

As in the (isolated) models of Berentzen et al. (2009), we find very good agreement between our estimated coalescence times and those found using full $P\mathcal{N}$ terms in our merger runs. We should note that the coalescence times for models D are all in the upper range of the coalescence times estimated in Paper II for $10^9 M_\odot$ LISA binaries. This can be easily understood if we note that in the previous paper, the nuclei were all modeled with $\gamma = 1$, whereas here we choose $\gamma = 7/4$, more in line with the expected properties of compact nuclei. We can see both from Figure 1 and Table 5 that the binary eccentricities tend to be lower ($e \sim 0.6–0.65$) for $\gamma = 7/4$, and indeed in Papers I and II, the binaries very often reached values $e \gtrsim 0.9$ (in accordance with values obtained in this paper for $\gamma = 1$). The dependence on eccentricity of the coalescence time under GW emission, $T_{\text{coal}, \text{GW}} \sim (1 - e^2)^{3/2}$, could easily account for a decrease of an order of magnitude or so when the binaries are very eccentric. It will be accordingly very important to further investigate the dependence of the eccentricity evolution under different values of $\gamma$ and $q$, and this is the subject of a forthcoming paper (M. Preto et al. 2012, in preparation).

6. SUMMARY AND CONCLUSIONS

Direct $N$-body simulation of mergers of spherically symmetric galaxies with different mass ratios were performed to investigate the evolution of binary SMBHs from the onset of merger through the stellar hardening phase until the eventual relativistic coalescence. The merging galaxies have different initial density profiles, varying from shallow ($\gamma = 1/2, 1$) to steep cusps ($\gamma = 3/2, 7/4$), thus covering the range of stellar distribution typically observed in the centers of bulges or early-type galaxies. In Papers I and II, we have shown that merger-induced triaxiality could support a purely stellar dynamical solution to the FPP in the case of equal-mass mergers. In order to assess the prospects for such a solution in the more general unequal-mass case, we measured both the hardening rate and merger-induced triaxiality for mass ratios in the range $q \in [0.05, 1]$. The merger-induced
How the shape of the progenitor galaxies affects the shape of the merger remnant and the resulting SMBH hardening rates is yet unclear. Nevertheless, it is important to pursue increasingly realistic stellar dynamical models of galactic nuclei in order to refine our understanding of binary SMBH evolution tracks and associated timescales. A detailed study of the $N$-independence and merger-induced triaxiality in unequal-mass mergers—including nonspherical progenitors resulting from previous merger events—is currently being pursued (M. Preto et al. 2012, in preparation).

We have produced estimates of coalescence times using a simplified prescription for the late relativistic phase of the inspiral—i.e., adopting the Peters (1964) equations for the evolution of an isolated binary and ignoring the late stellar-driven eccentricity evolution. We assessed the accuracy of these approximations by including $\mathcal{PN}$ terms (up to $3.5\mathcal{PN}$ order) to the equations of motion of the binary in a few representative cases of our sample. At least in these few selected cases, the agreement is remarkably good, and this is mainly due to the fact that the eccentricity evolution is almost identical in the $\mathcal{PN}$ and non-$\mathcal{PN}$ runs. One should add a little note of caution though. The eccentricity evolution may be a quite sensitive function of the properties of the surrounding stellar cluster (slope $\gamma$, mass ratio $q$, amount of net rotation, etc.) and on the initial conditions of the binary. Previous studies indeed indicate that one could expect significant evolution of the eccentricity during the hardening phase, even though not always of the same sign (Sesana 2010; Preto et al. 2011; Sesana et al. 2011). The coalescence times resulting from our calculations are consistently shorter than a Hubble time at a given redshift (up to $z \sim 2$–3 for all masses and up to $z \sim 6$–10 for those in the LISA mass range). Coalescence times for binaries of $\sim 10^6 M_\odot$ are all shorter than 1 Gyr; for the most massive $\sim 10^9 M_\odot$ they range between hundreds of millions of years (highly eccentric ones) to $\sim 1$–2 Gyr (less eccentric ones). SMBH binaries are therefore very promising sources of GWs for LISA both at low and high redshifts. The coalescence times obtained here—especially those from models D—may be effectively an upper bound if the mild eccentricity growth observed in these runs is not generic (we have only four runs). In Paper II, we used models with $\gamma = 1$ to compute the evolution of binaries in the LISA mass range and we obtained faster coalescences; this was partly because the binaries were significantly more eccentric. Therefore, the detailed properties of the distribution of coalescence times hinges in part on the dependence of the eccentricity evolution on the properties of the surrounding cluster. The question of whether eccentricity grows, decays, or remains approximately constant during the hardening phase still remains uncertain and we are currently pursuing it (M. Preto et al. 2012, in preparation). Interestingly, the duration of how long the binaries remain in the stellar dynamical hardening phase and in the relativistic one are roughly equal by an eccentricity regulated process. This means that SMBH binaries on highly eccentric orbits have less probability to be detected in the classical regime but at the same time are sources of strong GW signals to be detected. On the other hand, binaries with low eccentricities remain longer in the Newtonian regime and therefore are more likely to be observed in this phase, but may only be luminous GW sources in the final phase of coalescence.

With our results we moved another step toward a consistent stellar dynamical solution to the FPP, and thus of providing a solid dynamical substantiation to cosmological scenarios.

![Figure 8](image_url). Evolution of inverse semi-major axis $1/a$ (top) and eccentricity $e$ (bottom) for model A4 with and without $\mathcal{PN}$ terms, $a_0$ is the semi-major axis of the SMBH binary for which stellar dynamical hardening is equal to $\mathcal{PN}$ hardening.

(A color version of this figure is available in the online journal.)
where prompt coalescences are the norm during SMBH–galaxy coevolution (Sesana et al. 2007; Volonteri 2010). If our results hold true in real galaxies, the bottleneck to SMBH coalescences—if any—is likely to be associated to the long timescales needed for SMBHs to become a bound pair in galaxy mergers of unequal mass, especially in case timescales needed for SMBHs to become a bound pair in galaxy coalescences—if any—is likely to be associated to the long time delays (Preto & Amaro-Seoane 2010).

What about the role of gas in hardening the SMBH binaries at sub-parsec scales? Numerical simulations of merging disk galaxies with a substantial gas fraction have shown that gas is driven toward the central region of the merger remnant where it accumulates. As a consequence, the remnant becomes more axisymmetric (i.e., losing triaxiality) and the orbital structure changes noticeably as compared with the purely stellar case (Naab et al. 2006). Thus, we can expect that the fraction of centrophilic orbit is affected in the same manner by the presence of the gas. In which direction it affects the hardening rate is still an unresolved issue.

The gas that is funneled to the center of the remnant may undergo strong starbursts, which may lead to the formation of a steep stellar density cusp—even before the binary has formed—as shown by Callegari et al. (2009), which then in turn increases the binary’s hardening rate as shown in this work.

The gas could certainly also assist the inspiral if it is cold enough to settle into a circumbinary thin disk (Armitage & Natarajan 2005), even though the uncertainties associated with its long-term dynamical behavior are potentially more severe than those related to stellar dynamics. For instance, gaseous disks are susceptible to fragmenting and forming stars. In order for the disk to be at worst marginally stable against fragmentation at every radius, its total mass is constrained to be \( M_{\text{disk}} \sim (0.1–0.2) M_\odot \) (Cuadra et al. 2009; Lodato et al. 2009). This is not likely to be enough to drive the binary to coalescence.

The “mass deficits” induced to the centers of galaxies with no gas by inspiraling SMBH binaries are found to be of order \( \sim (1–5) M_\odot \), depending on the slope \( \gamma \) of their central stellar distribution and on their mass ratio \( q \). This is consistent with the picture that cores in giant ellipticals are scoured by SMBH binary inspirals during successive merger events—assuming only that relaxation times are long enough in those systems for the results from different mergers to be cumulative. For galaxies in the mass range of the Milky Way, the picture is basically distinct. Since the amount of mass depletion in a single merger event only partially destroys a steep inner cusp, the result is that even if a major merger were to have happened in the Galactic center at redshift \( z = 1 \) or larger, it would have been enough time to regrow a mass-segregated Bahcall & Wolf cusp since then (Preto & Amaro-Seoane 2010).

We thank Gabor Kupi and Patrick Brem for their help in testing and implementing the \( P^3N \) code subroutine for the binary SMBHs, which we use in our phi-GRAPE+GPU code. We also thank the referee for his/her valuable comments.

F.K. was supported by a grant from the Higher Education Commission (HEC) of Pakistan administrated by the Deutscher Akademischer Austauschdienst (DAAD). M.P. and R.S. acknowledge support by the Deutsches Zentrum für Luft-und Raumfahrt (through \( LISA \) Germany project). We acknowledge support by the Chinese Academy of Sciences Visiting Professorship for Senior International Scientists, grant number 2009S1-5 (The Silk Road Project; R.S. and P.B.). P.B. and I.B. both acknowledge financial support by the Deutsche Forschungsgemeinschaft (DFG) through SFB 881 “The Milky Way System” (sub-projects Z2 and A1) at the Ruprecht-Karls-Universität Heidelberg.

P.B. acknowledges the special support by the NASU under the Main Astronomical Observatory GRID/GPU computing cluster project (“graffias” cluster). P.B.’s studies are also partially supported by the program Cosmomicrophysics of NASU.

I.B. acknowledges funding through a Frontier Innovation grant of the University of Heidelberg sponsored by the German Excellence Initiative.

We used the following GPU clusters to run our simulations: “laohu” at the Center of Information and Computing at National Astronomical Observatories, Chinese Academy of Sciences, funded by Ministry of Finance of People’s Republic of China under the grant ZDY2008-2; “ko1ob,” which has been funded through a Frontier project supported by the excellence scheme of the University of Heidelberg and by the DFG; and “titian,” funded under the grants I/80 041-043 and I/81 396 of the Volkswagen Foundation and grants 823.219-439/30 and /36 of the Ministry of Science, Research and the Arts of Baden-Württemberg, Germany.

REFERENCES

Alexander, T., & Hopman, C. 2009, ApJ, 697, 1861
Armitage, P., & Natarajan, P. 2005, ApJ, 634, 921
Bahcall, J. N., & Wolf, R. A. 1976, ApJ, 209, 214
Bahcall, J. N., & Wolf, R. A. 1977, ApJ, 216, 883
Barack, L., & Cutler, C. 2004, Phys. Rev. D, 69, 082005
Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, Nature, 287, 307
Berczik, P., Merritt, D., & Spurzem, R. 2005, ApJ, 633, 680
Berczik, P., Merritt, D., Spurzem, R., & Binney, J., 2006, ApJ, 642, L21
Berentzen, I., Preto, M., Berczik, P., Merritt, D., & Spurzem, R. 2009, ApJ, 695, 455
Binney, J., & Tremaine, S. 2008, Galactic Dynamics (Princeton, NJ: Princeton Univ. Press)
Blanchet, L. 2006, Living Rev. Rel., 9, 4
Buchholz, R., Schödel, R., & Eckart, A. 2009, A&A, 499, 483
Callegari, S., Kazantzidis, S., Mayer, L., et al., 2011, ApJ, 729, 85
Callegari, S., Mayer, L., Kazantzidis, S., et al., 2009, ApJ, 696, L89
Cuadra, J., Armitage, P. J., Alexander, R. D., & Begelman, M. C. 2009, MNRAS, 393, 1423
Danzmann, K., et al., 2011, ESA LISA Yellow Book, http://sci.esa.int/science-e/www/object/index.cfm?fobjectid=48364
Dehnen, W. 1993, MNRAS, 265, 250
Do, T., Ghez, A. M., Morris, M. R., et al., 2009, ApJ, 703, 1323
Emel'chenko, E., Cappellari, M., Krajnović, D., et al. 2007, MNRAS, 379, 401
Faber, S. M., Tremaine, S., Ajhar, E. A., et al. 1997, AJ, 114, 1771
Ferrarese, L., Côte, P., Jordán, A., et al., 2006, ApJS, 164, 334
Ferrarese, L., & Ford, H. 2005, Space Sci. Rev., 116, 523
Ferrarese, L., & Merritt, D. 2000, ApJ, 539, L9
Gaburov, E., Harfst, S., & Portegies Zwart, S. 2009, New Astron., 14, 630
Gebhardt, K., Bender, R., Bower, G., et al. 2000, ApJ, 539, L13
Graham, A. W. 2004, ApJ, 613, L33
Gualandris, A., & Merritt, D. 2008, ApJ, 678, 780
Gualandris, A., & Merritt, D. 2012, ApJ, 744, 74
Harfst, S., Gualandris, A., Merritt, D., et al. 2007, New Astron., 12, 357
Häming, N., & Rix, H. 2004, ApJ, 604, L89
Hopkins, P. F., Cox, T. J., Dutta, S. N., et al. 2009a, ApJS, 181, 135
Hopkins, P. F., Lauer, T. R., Cox, T. J., Hernquist, L., & Kormendy, J. 2009b, ApJS, 181, 486
Hughes, S. A. 2003, Ann. Phys., 303, 142
Kormendy, J., & Bender, R. 1996, ApJ, 464, L119
Kormendy, J., Fisher, D. B., Cornell, M. E., & Bender, R. 2009, ApJS, 182, 216
Lightman, A. P., & Shapiro, S. L. 1977, ApJ, 211, 244
Lodato, G., Nayakshin, S., King, A. R., & Pringle, J. E. 2009, MNRAS, 398, 1392
Makino, J., & Ebisuzaki, T. 1996, ApJ, 465, 527
Makino, J., & Funato, Y. 2004, ApJ, 602, 93
Merritt, D. 2006a, ApJ, 648, 976
Merritt, D. 2006b, Rep. Prog. Phys., 69, 2513
Merritt, D. 2009, ApJ, 694, 959
Merritt, D. 2010, ApJ, 718, 739
Merritt, D., & Ekers, R. D. 2002, Science, 297, 1310
Merritt, D., & Ferrarese, L. 2001, MNRAS, 320, L30
Merritt, D., Mikkola, S., & Szell, A. 2007, ApJ, 671, 53
Merritt, D., & Milosavljević, M. 2005, Living Rev. Relativ., 8, 8
Merritt, D., & Poon, M. Y. 2004, ApJ, 606, 788
Merritt, D., & Vasiliev, E. 2011, ApJ, 726, 61
Mikkola, S., & Aarseth, S. 1998, New Astron., 3, 309
Milosavljević, M., & Merritt, D. 2003, ApJ, 596, 860
Naab, T., Jesseit, R., & Burkert, A. 2006, MNRAS, 372, 839
Perets, H. B., & Alexander, T. 2008, ApJ, 677, 146

Peters, P. C. 1964, Phys. Rev. B, 136, 1224
Poon, M. Y., & Merritt, D. 2001, ApJ, 549, 192
Preto, M., & Amaro-Seoane, P. 2010, ApJ, 708, L42
Preto, M., Berentzen, I., Berczik, P., Merritt, D., & Spurzem, R. 2009, J. Phys.: Conf. Ser., 154, 012049
Preto, M., Berentzen, I., Berczik, P., & Spurzem, R. 2011, ApJ, 732, L26
Preto, M., Merritt, D., & Spurzem, R. 2004, ApJ, 613, L109
Quinlan, G. D. 1996, New Astron., 1, 35
Roos, N., Kaastra, J. S., & Hummel, C. A. 1993, ApJ, 409, 130
Sesana, A. 2010, ApJ, 719, 851
Sesana, A., Guandaliris, A., & Dotti, M. 2011, MNRAS, 415, L35
Sesana, A., Volonteri, M., & Haardt, F. 2007, MNRAS, 377, 1711
Spitzer, L. 1987, Dynamical Evolution of Globular Clusters (Princeton, NJ: Princeton Univ. Press)
Sridhar, S., & Touma, J. 1999, MNRAS, 303, 483
Volonteri, M., J., Lehto, H. J., Nilsson, K., et al. 2008, Nature, 851, 452
Volonteri, M. 2010, A&AR, 18, 279
Yu, Q. 2002, MNRAS, 331, 935