Accretion into the central cavity of a circumbinary disc

Daniel J. D’Orazio,1* Zoltán Haiman1* and Andrew MacFadyen2*

1Department of Astronomy, Columbia University, 550 West 120th Street, New York, NY 10027, USA
2Center for Cosmology and Particle Physics, Physics Department, New York University, New York, NY 10003, USA

ABSTRACT
A near-equal-mass binary black hole (BH) can clear a central cavity in a circumbinary accretion disc; however, previous works have revealed accretion streams entering this cavity. Here we use 2D hydrodynamical simulations to study the accretion streams and their periodic behaviour. In particular, we perform a suite of simulations, covering different binary mass ratios \( q = M_2/M_1 \) in the range 0.003 \( \leq q \leq 1 \). In each case, we follow the system for several thousand binary orbits, until it relaxes to a stable accretion pattern. We find the following results: (i) the binary is efficient in maintaining a low-density cavity. However, the time-averaged mass accretion rate into the cavity, through narrow coherent accretion streams, is suppressed by at most a factor of a few compared to a disc with a single BH with the same mass; (ii) for \( q \gtrsim 0.05 \), the accretion rate is strongly modulated by the binary, and depending on the precise value of \( q \), the power spectrum of the accretion rate shows either one, two or three distinct periods; and (iii) for \( q \lesssim 0.05 \), the accretion rate becomes steady, with no time variations. Most binaries produced in galactic mergers are expected to have \( q \gtrsim 0.05 \). If the luminosity of these binaries tracks their accretion rate, then a periodogram of their light curve could help in their identification, and to constrain their mass ratio and disc properties.

Key words: accretion, accretion discs – black hole physics – gravitational waves – galaxies: active.

1 INTRODUCTION
Massive black holes (MBHs) appear to reside in the nuclei of most nearby galaxies (see, e.g. reviews by Kormendy & Richstone 1995; Ferrarese & Ford 2005). In hierarchical structure formation models, galaxies are built up by mergers between lower mass progenitors, which deliver nuclear MBHs (e.g. Springel, Di Matteo & Hernquist 2005; Robertson et al. 2006), along with a significant amount of gas (Barnes & Hernquist 1992), to the central region of the newly born post-merger galaxy. Since mergers are common (e.g. Haehnelt & Kauffmann 2002), it follows that massive black hole binaries (MBHBs) should also be common in galactic nuclei.

Despite this expectation, observational evidence for MBHBs remains scarce (see, e.g. Komossa 2006; Tsalmantza et al. 2011; Eracleous et al. 2012). The dearth of MBHBs could be attributed to several factors: it is possible that typically only one of the two black holes (BHs) is active at spatially resolvable separations; binaries may also lose their angular momentum efficiently due to the surrounding stars and gas and quickly move to spatially unresolvable orbital separations. Another possible hindrance, which we address in this paper, is that the outward gravitational torques from the binary can balance the inward viscous torques and pressure forces, clearing a central cavity in a putative circumbinary gas disc (Artymowicz & Lubow 1994), possibly rendering the system too dim for detection. Overall, identifying MBHBs is difficult, and a better understanding of their expected observational signatures, especially those based on time variability (Haiman et al. 2009a), is needed. Merging MBHBs should be unambiguously identifiable by gravitational wave (GW) detectors, such as evolving Laser Interferometer Space Antenna (sLISA; Amaro-Seoane et al. 2013) or ongoing Pulsar Timing Arrays (PTAs; e.g. Lommen 2012). Identifying the electromagnetic (EM) counterparts of these GW sources (among the many false candidate galaxies in the GW error box) will, however, likewise require an understanding of their observational signatures.

Recent studies have explored the gas dynamics of circumbinary accretion discs around near-equal-mass binaries in some detail. Since the system is not axisymmetric, this requires a 2D or 3D treatment. (MacFadyen & Milosavljević 2008, hereafter MM08) have run 2D hydrodynamical simulations for an equal-mass binary. 2D smoothed particle hydrodynamical (SPH) simulations have been carried out for equal-mass and 2:1 mass-ratio binaries by Hayasaki, Mineshige & Sudou (2007) and for a 3:1 mass-ratio binary by Cuadra et al. (2009) and Roedig et al. (2012). Shi et al. (2012) have followed up on the work of MM08 for an equal-mass binary by running 3D magnetohydrodynamical (MHD) simulations, and...
Noble et al. (2012) have further added a post-Newtonian treatment of general relativistic (GR) effects, and followed the disc through the late stages of orbital inspiral (from orbital separation $r = 20$ to 8M). Farris et al. (2012a) followed the merger of an equal-mass binary and a surrounding disc through merger in full 3D general relativistic MHD (GRMHD), starting from 10M. Finally, Farris et al. (2012b) have added gas cooling to GRMHD simulations of an equal-mass binary prior to decoupling, through decoupling and to merger starting from 10M. A generic result of all of these studies is that a low-density cavity is carved out by the binary torque, but gas leaks into the cavity through non-axisymmetric streams (as first discussed in the SPH simulations of Artymowicz & Lubow 1996). These streams can power significant accretion on to the binary components, and should lead to bright EM emission.

A particularly promising feature is that the accretion rate on to the BHs can be both high and strongly variable, modulated by the binary’s orbital motion. This could allow a detection of subparsec binaries by looking for periodic variations in the luminosity of AGN-like objects (Haiman, Kocsis & Menou 2009b) or periodic shifts and intensity variations of spectral lines (e.g. Haiman et al. 2009b; Shen & Loeb 2010; Eracleous et al. 2012, and references therein). If the accretion remains significant and periodic down to $\ll$parsec separations, then it could also enable the identification of EM counterparts of GW sources: either for precursors to eLISA sources in the $M = 10^7-10^8$ M$\odot$ range (Kocsis et al. 2006; Kocsis, Haiman & Menou 2008) or by detecting periodic modulations of more massive $M = 10^9-10^{10}$ M$\odot$ binaries discovered by PTAs (Sesana et al. 2012; Tanaka, Menou & Haiman 2012).

Although existing studies have focused on near-equal-mass MBHBs, in reality, coalescing MBHBs should have a distribution of mass ratios $q = M_2/M_1$. Mergers occur between galaxies over a wide range of sizes, harbouring central BHs of different masses, so that MBHBs resulting from galactic mergers should have a correspondingly wide range of mass ratios. Studies based on Monte Carlo realizations of dark matter merger trees indeed find broad distributions between $10^{-2} \lesssim q < 1$, generally peaking in the range $q \approx 0.1-1$ (e.g. Volonteri, Haardt & Madau 2003; Sesana et al. 2005, 2012; Gergely & Biermann 2012). However, the predictions depend on the occupation fraction of MBHBs, the redshift evolution of the correlation between the masses of MBHBs and their host galaxies, as well as on the limit on the mass ratio of host galaxies whose nuclear MBHBs can coalesce; $q < 0.1$ mergers could in fact be most common (e.g. Lippai, Frei & Haiman 2009).

Here we follow up on the earlier work of MM08 and move beyond the near-equal-mass binary case. We study the periodicity and the time-averaged rate of accretion across the central cavity, by running 2D hydrodynamical simulations of a circumbinary disc for 10 different binary mass ratios ranging from $q = 0.003$ to 1. Clearly, one expects that in the limit $q \to 1$, the accretion rate approaches that of an accretion disc around a single BH, and will no longer be time variable. The main goal in this paper is to answer the following basic questions: How does the mean accretion rate, and its fluctuations, depend on the mass ratio? In particular, down to what mass ratio is the mean accretion and/or its variability significantly affected by the binary torques? We address these questions with the caveat that, throughout this paper, accretion is defined as the mass crossing the inner boundary of the simulation domain and not necessarily that accreted by either BH.

The rest of this paper is organized as follows. In Section 2, we describe the setup of our numerical simulations, including changes we made to the public version of the Eulerian grid code FLASH and the initial and boundary conditions we adopted. In Section 3, we present our main results, namely that we find four distinct patterns for the time variability of the accretion rate as a function of the mass ratio $q$. In Section 4, we compare our findings with that of MM08 as well as investigate the dependence of our results on the magnitude of viscosity on the resolution. We also discuss scaling of the simulations to physical parameters, such as BH mass and orbital separation, and discuss the corresponding orbital and residence times, as well as some caveats. Finally, in Section 5, we conclude by briefly summarizing our main results and their implications. The Appendix details our implementation of viscosity in polar coordinates, an important addition to FLASH.

2 DETAILS OF NUMERICAL SIMULATIONS

To simulate a gas disc in the gravitational field of a binary, we use the Eulerian grid-based hydrodynamical code FLASH (version 3.2; Fryxell et al. 2000).$^2$ FLASH solves the volume-integrated fluid equations by solving the Riemann problem at each cell boundary. A piecewise parabolic representation of the fluid variables is used to interpolate between cells, i.e. FLASH is a Piecewise Parabolic Method (PPM) code, accurate to second order in both space and time. FLASH uses a monotonicity constraint, rather than artificial viscosity, to control oscillations near discontinuities. This makes it well suited for following supersonic fluid dynamics in the inner regions of circumbinary disc. FLASH also supports polar coordinates, which is convenient for simulating discs.

2.1 Numerical implementation and assumptions

We assume a geometrically thin accretion flow with angular momentum aligned with that of the binary. This permits a decoupling of the fluid equations in the $z$-direction, perpendicular to the plane of the disc, so that we can define height-integrated fluid variables and set up simulations in two dimensions. In follow-up studies, we plan to extend these simulations to the full three dimensions, which we expect will be important in determining the amount of inflow into a putative circumbinary cavity (for recent 3D grid-based simulations, see Noble et al. 2012; Shi et al. 2012). In the present study, we choose 2D polar coordinates $(r, \phi)$ and employ FLASH to solve the following standard set of 2D hydrodynamical equations:

$$\frac{\partial \Sigma}{\partial t} + \nabla \cdot (\Sigma \mathbf{v}) = 0$$
$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = -\frac{1}{\Sigma} \nabla P - \nabla \Phi_{\text{bin}} + \nabla \cdot \nu \nabla \mathbf{v} + \nabla \left( \frac{1}{2} \mathbf{v} \cdot \nabla \mathbf{v} \right)$$
$$\frac{\partial (\Sigma E)}{\partial t} + \nabla \cdot [(\Sigma E + P) \mathbf{v}] = \Sigma \mathbf{v} \cdot (-\nabla \Phi_{\text{bin}}).$$

Here, $\Sigma$ is the vertically integrated disc surface density, $\mathbf{v} = v_r \mathbf{\hat{r}} + v_\phi \mathbf{\hat{\phi}}$ is the fluid velocity, $P$ is the pressure, $\nu$ is the coefficient of kinematic viscosity, $E$ is the total internal plus kinetic energy of the fluid, $E = \epsilon + \frac{1}{2}|\mathbf{v}|^2$ and $\Phi_{\text{bin}}$ is the gravitational potential of the

$^1$ Nixon et al. (2011a) explored a range of mass ratios, using 3D SPH simulations, but restricted their study to retrograde discs.

$^2$ We note that MM08 used an earlier release, version 2, of the same code.
binary. The gravitational potential is inserted into the simulation by hand, and is given by

$$\Phi_{\text{bin}}(r, \phi) = -\frac{GM(1 + q)^{-1}}{r^2 + \left(\frac{a}{1 - q}\right)^2 - \frac{2a}{1 - q} \cos(\phi - \Omega_{\text{bin}} t)}^{1/2}$$

$$-\frac{GM(1 + q^{-1})^{-1}}{r^2 + \left(\frac{a}{1 - q}\right)^2 + \frac{2a}{1 - q} \cos(\phi - \Omega_{\text{bin}} t)}^{1/2}. \quad (2)$$

Here, $$\Omega_{\text{bin}} = \left(\frac{GM}{a^3}\right)^{1/2}$$ is the binary’s orbital frequency, $$a$$ is the separation of the primary, $$M_p > M_s$$ are the masses of the primary and the secondary, $$M = M_p + M_s$$ is the total mass and $$q = M_s/M_p \leq 1$$ is the mass ratio. The origin of the coordinate system is chosen to coincide with the binary’s centre of mass. In the case of a single point mass, we use the limit of equation (2) as $$q, a \to 0, \Phi_{\text{bin}} = -GM/r$$.

Note that we do not evolve the orbital parameters of the binary nor do we allow its centre of mass to wander; these simulations are numerical experiments which are physically motivated in the limit of small disc mass (this assumption is justified for our physical parameter choices; see discussion in Section 4.4 below).

We neglect self-gravity of the disc. Given the local sound speed $$c_s$$, the Toomre parameter $$Q = c_s/\Omega_p/(\pi G \Sigma)$$ can be written as $$Q \sim (H/r)(M/M_d)$$, for a disc with mass $$M_d$$ and vertical scaleheight $$H$$ in hydrostatic equilibrium. For a thin disc $$H/r \lesssim 0.1$$, but in all of our simulations, we choose $$M \gg M_d$$ and thus $$Q \gg 1$$, making the disc stable to gravitational fragmentation. This is justified for standard Shakura–Sunyaev discs, when the simulations are scaled to physical BH masses and sufficiently small separations (see Section 4.4 below). For comparison, we note that in their SPH simulations, Cuadra et al. (2009) and Roedig et al. (2012) studied more massive discs, with $$M_d \sim 0.2 M$$, making self-gravity important.

The pressure is given, as in MM08, by a locally isothermal equation of state,

$$P = c_s^2(r) \Sigma, \quad (3)$$

where the sound speed is assumed to scale with radius as $$c_s \propto r^{-1/2}$$. For a Keplerian potential, this corresponds to a constant disc scaleheight to radius ratio, $$H/r = c_s/v_g \equiv M^{-1}$$, where $$v_g$$ is the orbital velocity in the disc and $$M$$ is the corresponding Mach number. Throughout our simulations, we choose the disc sound speed such that, for a Keplerian azimuthal velocity, $$H/r = 0.1$$ (or $$M = 10$$) everywhere.\(^3\)

To incorporate viscosity, FLASH calculates the momentum flux across cell boundaries due to viscous dissipation (the last two terms in the second of equations 1). To compute $$v$$ we adopt the $$\alpha$$ prescription, $$v = \alpha c_s H$$ (Shakura & Sunyaev 1973), where $$\alpha$$ is a dimensionless parameter indicating the scale of turbulent cells, and the scaleheight is computed from $$H = c_s/\sqrt{v_g}$$ with $$v_g$$ being the Keplerian value. Following MM08, we choose a fiducial, constant $$\alpha = 0.01$$ (although we explore the effects of increasing $$\alpha$$ in Section 4.2). Since the FLASH viscosity implementation is not fully supported in polar coordinates, we made adjustments to the routines that compute the momentum flux and viscous diffusion time from $$v$$, originally in Cartesian coordinates. These modifications and tests of our polar viscosity implementation are detailed in the Appendix.

\(^3\) Since we add the binary’s quadrupole potential, and the non-zero pressure makes the azimuthal velocities slightly sub-Keplerian, in practice $$M$$ in the simulations approaches ~18 near $$r = a$$ and becomes a constant $$M \sim 10$$ at $$r > 3a$$.

### 2.2 Numerical parameter choices

The inner edge of the computational domain is chosen at $$r_{\text{min}} = a$$ and the outer edge at $$r_{\text{max}} = 100a$$. Although we are only interested in the inner few $$r/a$$ of the disc, extending the computations to larger radii acts as a buffer for small initial numerical transients, and also provides a potential reservoir of gas from which the inner regions can be fed. As in other Eulerian codes, FLASH uses a boundary zone of ‘guard cells’ which enforces boundary conditions. We choose a diode-type inner boundary condition: the values of fluid variables in cells bordering the guard cells are copied into the guard cells, with the restriction that no fluid enter the domain, $$v_r(r_{\text{min}}) \leq 0$$. We adopt an ‘outflow’, outer boundary condition: this is identical to the diode version, except that flow is allowed both into and out of the domain. In practice, with our initial conditions in Section 2.3, we find an outflow at the outer boundary of our disc. However, the simulation does not run for a significant fraction of a viscous time at the outer radius (see below) and we expect inflow at the outer boundary to establish itself eventually.

Unless specified otherwise, the spatial resolution is fixed throughout the grid, with the default values set at $$[\Delta r/a, \Delta \phi/2\pi] \simeq [0.024, 0.0078]$$, corresponding to a grid of $$\sim 4096 \times 128$$ cells. The wavelength of a density wave due to a Lindblad resonance is of order $$2\pi H \sim 0.6a$$ (e.g. Dong, Rafikov & Stone 2011b; Duffell & MacFadyen 2012); as in MM08, the radial resolution is chosen to resolve this length-scale with many cells. Note that our fiducial resolution results in cell aspect ratios which are approximately square at $$r \sim 3a$$. To test numerical convergence (see Section 4.3 below), we performed several additional runs, increasing the radial and azimuthal resolutions, by factors of 1.42 and 1.42\(^2\) ~ 2.0, from the lowest resolution runs. The time resolution is set to be half of the shortest propagation time (viscous or dynamical) across a cell. The effects of changing the resolution are discussed in Section 4.3.

We run our simulations for between $$4 \times 10^3$$ and $$10^5$$ binary orbits. For reference, we note that the viscous time can be related to the orbital time as

$$t_{\text{visc}} = \frac{2r^2}{3 \nu} = \frac{M^2}{2\pi a} \simeq 1060 \left(\frac{M^2}{100}\right) \left(\frac{0.01}{\alpha}\right) t_{\text{orb}}. \quad (4)$$

Thus, our typical run of 4000 binary orbits corresponds to ~4 viscous times at the innermost regions, but less than one viscous time at $$r \gtrsim 3a$$.

We have performed 34 runs altogether, for 10 different binary mass ratios between $$q = 0.003$$ and 1.0 (including control runs for $$q = 0$$, i.e. a single BH). The mass ratio, resolution and the number of binary orbits followed in each of our simulation runs are summarized in Table 1.

### 2.3 Initial conditions

In our initial conditions, we insert a central cavity around a binary, and also include a density pile-up just outside the cavity wall. Such a surface density profile is expected to develop during the inward migration of the secondary. When the secondary arrives at the radius where the local disc mass is too small to absorb the secondary’s angular momentum, its migration stalls, the inner disc drains on to the primary, and continued accretion from larger radii causes a pile-up of gas outside the secondary’s orbit (Syer & Clarke 1995; Ivanov, Papaloizou & Polnarev 1999; Milosavljević & Phinney 2005; Chang et al. 2010; Rafikov 2012). As emphasized recently by Kocsis, Haiman & Loeb (2012a), the details of this process are still uncertain, as the coupled time-dependent migration, cavity...
formation and pile-up, has not been modelled self-consistently, even in 1D calculations. However, using self-consistent steady-state solutions, Kocsis, Haiman & Loeb (2012b) showed that in many cases, the pile-up can cause overflow already at large binary separations.

For simplicity, and for ease of comparison, we adopt the same initial surface density profile as in MM08. This profile is motivated by the earlier results of Milosavljević & Phinney (2005); it has very little gas inside of \( r \approx 3a \), and peaks at \( \sim 8a \),

\[
\Sigma(r, t_0) = \Sigma_0 \left( \frac{r}{r_0} \right)^3 \exp \left[ - \left( \frac{r}{r_0} \right)^2 \right].
\]

Here, \( \Sigma_0 \) is an arbitrary constant and \( r_0 = 10a \). The initial profile is shown by the (blue) dashed curves in Fig. 4 below.

We also follow MM08, and incorporate pressure gradients and the quadrupole contribution of the binary’s potential into the initial azimuthal velocity

\[
\Omega_z^2 = \Omega_*^2 \left[ 1 + \frac{3}{4} \left( \frac{\alpha}{r_0} \right)^2 \frac{q}{1 + q} \right]^2 + \frac{1}{r \Sigma} \frac{dP}{dr}
\]

and account for viscous drift in the initial radial velocities

\[
v_r = \frac{d}{dr} \left( r^3 \nu \Sigma \frac{d\Sigma}{dr} \right) \left[ r \Sigma \frac{d}{dr} \left( r^2 \Omega_z \right) \right]^{-1}.
\]

We emphasize that with these initial conditions, the disc is not initially in equilibrium, and material diffuses away from the peak of the surface density, both inward and outward, due to pressure gradients. However, after running the simulations for several thousand orbits, the system relaxes to a steady pattern of accretion, and we do not expect the initial profile to significantly influence our conclusions.

### 2.4 Disc parameters and accretion rate

Our primary goal is to quantify the magnitude and variability of accretion across the central cavity. To this end, we compute the time-dependent accretion rate at the inner edge of the simulation domain, \( r_{\text{min}} = a \)

\[
\dot{M}(t) = \int_0^{2\pi} \Sigma(r_{\text{min}}) v_r(r_{\text{min}}) r_{\text{min}} d\phi.
\]

Since we are unable to track the fate of the gas at smaller radii, nor do we allow the masses of the BHs to increase, this mass is effectively lost from the simulation. In practice, the total mass that is lost is a small fraction of the total initial disc mass (at most a few per cent by the end of each run).

Because we neglect the self-gravity of the disc, equations (1) are independent of \( \Sigma_0 \), and there is no unique way to assign a physical normalization to \( \dot{M}_{\text{sim}} \) without appealing to a disc model which includes additional physics. Instead, MM08 compare the average accretion rate at the edge of the integration domain (\( r_{\text{min}} = a \) to that in a disc in a Keplerian potential with the same surface density at \( r \approx 3a \),

\[
\frac{\dot{M}_{\text{min}}}{\dot{M}_{\text{free}}} \approx \frac{\langle \dot{M}_{\text{bin}}(a) \rangle}{6\pi \alpha} \frac{H}{r} \left( \frac{GM_r}{r} \right)^{-2} (\text{GM})^{-1/2} \Sigma^{-1}(3a).
\]

MM08 find \( \dot{M}_{\text{bin}}/\dot{M}_{\text{free}} \approx 0.2 \). To make this comparison meaningful, one must assume that \( \dot{M}_{\text{free}}(3a) \approx \dot{M}_{\text{free}}(a) \), i.e. that the reference, fictitious point-mass disc is in steady state. However, for a steady-state disc, specifying \( \alpha \) and \( H/r \) (along with the dominant source of opacity) sets the physical value \( \dot{M}_{\text{free}} \). For our fiducial values of \( M = 10^7 M_\odot \) and \( a = 10^3 r_S \) (and with \( \alpha = 0.01 \) and \( H/r = 0.1 \)), we find the unphysically large value of \( \dot{M}_{\text{free}} \approx 10^3 \dot{M}_{\text{Edd}} \), a result of the above choices (where \( \dot{M}_{\text{Edd}} \) is the accretion rate that would produce the Eddington luminosity, with a radiative efficiency of 10 per cent). If we instead require \( \dot{M}_{\text{free}} \approx \dot{M}_{\text{Edd}} \), then this would translate to a much thinner disc, with \( H/r \approx 4 \times 10^{-3} \). In such a thin/cold disc, resolving density waves with the same number of cells as in MM08 would require a radial grid resolution ~17 times higher than in our highest resolution run, and would be impractical.

Rather than attempting to compare our binary simulations to a hypothetical steady-state point-mass disc, we choose to leave the surface density normalization \( \Sigma_0 \) essentially arbitrary, and instead perform explicit reference simulations with a single BH (i.e. \( q \to 0 \)), with the same initial conditions as the binary runs. This approach has the advantage of explicitly isolating the effect of turning on/off the binary. Note, however, that our point-mass runs should not be expected to produce the steady-state accretion rate for a corresponding single-BH system. This is because our initial conditions are far from this state, and we do not run the simulations long enough (i.e. for a few viscous time at the outer edge) to allow it to settle to the correct steady state.

---

**Table 1.** Summary of our simulation runs. Low, medium, and high radial and azimuthal resolutions are denoted by ‘Lo\(\Delta r\)’; ‘Mid\(\Delta r\)’; ‘Hi\(\Delta r\)’ and ‘Lo\(\Delta \phi\)’; ‘Mid\(\Delta \phi\)’; ‘Hi\(\Delta \phi\)’ and correspond to \( \Delta r/a = 0.035, 0.24, 0.017 \) and \( \Delta \phi/2\pi = 0.0078, 0.0052, 0.0039 \), respectively. The viscosity parameter \( \alpha \) is set to 0.01 unless otherwise specified.

| Mass ratio \( q \) | Spatial resolution \([\Delta r, \Delta \phi]\) | No. of orbits \((N_{\text{orb}} = t_{\text{max}} \Omega_{\text{min}}/2\pi)\) |
|------------------|----------------------|------------------|
| 1.0 \((\alpha = 0.02)\) | \([\text{Lo}\Delta r, \text{Lo}\Delta \phi]\), \([\text{Mid}\Delta r, \text{Lo}\Delta \phi]\) | 5500, 6800, 4500, 4200 |
| 1.0 \((\alpha = 0.04)\) | \([\text{Mid}\Delta r, \text{Lo}\Delta \phi]\) | 5200 |
| 1.0 \((\alpha = 0.1)\) | \([\text{Hi}\Delta r, \text{Hi}\Delta \phi]\) | 1400 |
| 0.75 | \([\text{Lo}\Delta r, \text{Lo}\Delta \phi]\), \([\text{Hi}\Delta r, \text{Hi}\Delta \phi]\) | 6500, 4500 |
| 0.5 | \([\text{Mid}\Delta r, \text{Lo}\Delta \phi]\) | 6000, 4500 |
| 0.25 | \([\text{Hi}\Delta r, \text{Hi}\Delta \phi]\) | 5700, 4500 |
| 0.1 | \([\text{Lo}\Delta r, \text{Lo}\Delta \phi]\), \([\text{Mid}\Delta r, \text{Mid}\Delta \phi]\), \([\text{Hi}\Delta r, \text{Hi}\Delta \phi]\) | 7000, 8000, 4500, 4100 |
| 0.075 | \([\text{Mid}\Delta r, \text{Lo}\Delta \phi]\) | 4500, 5200 |
| 0.05 | \([\text{Hi}\Delta r, \text{Hi}\Delta \phi]\) | 4500, 7600, 4500, 4200 |
| 0.025 | \([\text{Mid}\Delta r, \text{Lo}\Delta \phi]\) | 4500 |
| 0.01 | \([\text{Hi}\Delta r, \text{Hi}\Delta \phi]\) | 5200, 5500 |
| 0.003 | \([\text{Mid}\Delta r, \text{Lo}\Delta \phi]\) | 5500 |
| 0 \((\text{All } \alpha)\) | \([\text{Lo}\Delta r, \text{Lo}\Delta \phi]\), \([\text{Mid}\Delta r, \text{Lo}\Delta \phi]\), \([\text{Hi}\Delta r, \text{Hi}\Delta \phi]\) | 6000, 8000, 4500, 4200 |
2.5 Tests of code implementation

To ensure that non-axisymmetric perturbations are not induced artificially in the disc, we simulate a disc around a single BH for $10^4$ binary orbits. To keep axisymmetry during these runs, we re-implemented the FLASH routine UNBIASED-GEOMETRY (UBG)\(^4\) into the FLASH3 routines. UBG cleans up round-off errors in the cell boundary positions, and forces the polar grid to keep cell sizes uniform in the azimuthal direction. Without UBG, small perturbations in the azimuthal direction grow to significant size after a few hundred orbits. We have verified that after re-implementing UBG, axisymmetry is preserved for $10^4$ binary orbits.

A detailed description as well as multiple tests of our viscosity implementation are laid in out in the Appendix.

3 RESULTS

3.1 Equal-mass binary

We begin by describing the results of our equal-mass binary runs in some detail. Although these are very similar to those of MM08, this will serve as a useful point of comparison for our unequal-mass runs.

3.1.1 A toy model with massless particles

Before showing the results from our simulations, we consider a simple toy model, based on the orbits of non-interacting massless test particles around a binary. We populate a 2D disc with test particles, centred on the binary’s centre of mass, and leave a central cavity. We assign initial velocities equal to the Keplerian velocities around a single point mass $M = M_p + M_B$. We then follow the orbit of each test particle in the rotating binary potential, by numerically solving the restricted three-body problem for each individual particle (for $10^5$ particles in practice, using equations 3.16 and 3.17 in Murray & Dermott 2000).

The results of this simple exercise are displayed in Fig. 1, which shows the locus of the test particles initially, as well as after 0.25, 0.5 and 0.75 binary orbits (in a frame corotating with the binary). As this figure demonstrates, there is a tendency for the binary to pull streams of particles into the cavity. This, of course, is purely a gravitational effect. As Artymowicz & Lubow (1996) have pointed out, such mass flows occur near unstable co-rotation equilibrium points in the binary potential.

The toy model cannot be pushed much further in time, since after ~0.75 binary orbits, the trajectories of the test particles cross – this necessitates a hydrodynamical treatment. Nevertheless, the figure does suggest that an empty cavity cannot be maintained by an equal-mass binary, even in the absence of pressure or viscosity. Furthermore, as we will see below, the simulations show accretion streams with morphologies quite similar to those in the bottom-right panel of Fig. 1.

3.1.2 Hydrodynamical evolution: reaching steady state

We next present the results from our equal-mass binary simulations. The disc evolves through two distinct stages. As explained above, the disc is set up to be out of equilibrium, and we observe an initial transient state, which lasts for ~2500 orbits. We expect the details of this state to depend on the initial conditions. The disc then settles to a quasi-steady state, which persists for the rest of the simulation. In this quasi-steady state, the disc exhibits significant accretion, which varies periodically on two time-scales $(1/2)\Omega_{\text{bin}}$ and $(5–6)\Omega_{\text{bin}}$. We expect these latter features to be robust and insensitive to initial conditions.

Transient state. Initially, pressure forces and viscous stresses act to move material inside of the initial density peak at $r \sim 8a$ inward while pressure gradients move material outside of the density peak outward. Once the inner disc material reaches $r \sim 2a$, its surface density becomes strongly perturbed and highly non-axisymmetric. Ruminicent of the evolution of the toy model in Section 3.1.1, two narrow point-symmetric streams develop, in which the gas flow becomes nearly radial. About 3 per cent of the material in these streams exits the integration domain at $r_{\text{bin}} = a$. The rest gains angular momentum from the faster moving BHs and is flung back towards the bulk of the disc material at $r \sim 2a$. This process maintains a central cavity within the disc, with gas streams being pulled in and pushed out on a period of $\sim 1.5\Omega_{\text{bin}}$.

As more disc matter flows in from the initial density peak to the inner $r \sim 2a$ region, the streams become more dense. When these streams are flung back out and hit the opposing cavity wall, they generate noticeable overdensities, and deform the circular shape of the cavity to become eccentric, though still point symmetric. These overdensities then rotate at the disc’s orbital velocity. They spread out and propagate into the disc as differential rotation causes them to wind up, creating a point-symmetric ($m = 2$), rotating spiral pattern.

---

\(^4\) This routine is titled clean_last_bits in the FLASH2 download.
Quasi-steady state. After the initial $\sim 2500$ orbits, the point symmetry of the transient state breaks down as stream generation becomes preferentially stronger on one side of the cavity. This causes more mass, in the form of a stream, to be driven into the opposite cavity wall, pushing that side of the wall farther away from the binary. This lopsided state can grow from a small initial asymmetry, through a genuine physical instability, as follows.

Initially, in the transient point-symmetric state, the central cavity has an elliptical (but still point-symmetric) shape, which rotates along with the disc, its inner edge completing a rotation once every three binary orbits. Streams are simultaneously pulled from the two near sides of this elliptical cavity. After the streams form, they are flung across the cavity, and hit the region approximately diagonally across, close to the azimuth where the opposite stream formed. The above process makes the cavity more eccentric over time, since when the outward-going stream material hits the cavity wall, it pushes it outward, further away from the binary’s centre of mass.

Fig. 2 shows the above, via three snapshots of the disc in the transient, point-symmetric state, in time order from left to right. The first panel shows the initial formation of the two opposing streams, demonstrating that the streams originate at locations where the cavity wall is closest to the binary’s centre of mass. The third panel shows the collision of the outward-going streams with the opposing cavity wall, demonstrating that the streams collide with the cavity wall approximately diagonally across their initiation point. The second panel shows the morphology of the streams half-way between these time steps, for the sake of completeness.

Next, imagine that due to numerical noise, one stream (say, from side ‘A’ of the cavity) carries slightly more momentum than the spatially opposite stream emanating from side ‘B’. This can happen due to a small initial lopsidedness in the shape of the cavity, with side ‘A’ being closer to the binary than side ‘B’ (or due to an asymmetry in the density or velocity field). In our simulations, this can only be due to small numerical noise, but in reality, discs will obviously not be perfectly symmetric, either. The stronger stream will hit side ‘B’ of the cavity as before, but will push the cavity wall farther away from the binary than the comparatively smaller counterpart stream hitting side ‘A’. It is easy to see that this can lead to a runaway behaviour: side ‘A’ of the cavity will have absorbed less momentum, and will now be even closer to the binary’s centre of mass, relative to side ‘B’ — causing a larger asymmetry in the next pair of streams, which further increases the lopsidedness of the cavity, and so on.

This reinforcement-feedback process continues for a period of $\gtrsim 200–300$ orbits, after which the weaker stream, and its effect on the cavity wall structure disappears entirely. The central cavity takes on a lopsided shape, with a near-side where streams are pulled from the cavity wall by each passage of the holes, and a far-side where non-accreted material from the streams is flung back and crashes into the cavity wall. At the azimuthal locations of these crashes, the far-side of the wall develops very strong shocks, with Mach numbers up to $M \sim 15$. (However, since our disc is locally isothermal, this is likely an upper limit for the shock strength.)

The lopsided cavity precesses in the frame at rest with respect to the binary centre of mass, completing a rotation once every $\sim 400$ binary orbits. Excitation of a similar lopsided cavity is observed in the 3D MHD, as well as 2D hydrodynamical simulations of Shi et al. (2012), Noble et al. (2012) and MM08. Shi et al. (2012) explore the possible generation mechanisms of this mode. They find its growth to be consistent with being caused by asymmetric stream impacts described above.

We show examples of the 2D surface density distributions in Fig. 3. The top row of this figure shows snapshots at $\sim 1000$ binary orbits (left) and at $\sim 4000$ binary orbits (right), of the inner 6 per cent of the simulated disc (i.e. $\pm 6r/a$ are shown in both directions). The solid circle marks the inner boundary of the simulation domain at $r = r_{\text{min}} = a$. The larger dotted circle at $r \simeq 2.08a$ is the position of the $(m, l) = (2, 1)$ outer Lindblad resonance (not present, but shown for reference). The left of these two panels illustrates the point-symmetric, transient state. The zoomed-in inset of this panel shows two weak point-symmetric streams reminiscent of the streams seen in the toy model of Fig. 1. The right-hand panel shows the disc after it has settled to its quasi-steady, lopsided state, with a single stream. The top-left panel of Fig. 4 shows snapshots of the azimuthally averaged surface density profile of the equal-mass binary disc at three different times, after 0, 2000 and 4000 orbits. For comparison,
Figure 3. Top row: surface density distributions for the equal-mass ratio ($q = 1.0$) binary during a transient, point-symmetric state after $\sim 1000$ binary orbits (left) and during the quasi-steady asymmetric state after $\sim 4000$ binary orbits (right). The inset in the top-left panel zooms into the inner $\pm 2.5r/a$ of the disc in order to show the stream morphology. Bottom two rows: snapshots at $\sim 4000$ binary orbits, during the quasi-steady-state phase, for mass ratios $q = 0.5, 0.1, 0.075$ and $0.01$, as labelled. Each panel shows the inner $\sim 6$ per cent of the simulated disc, extending $\pm 6r/a$ in both directions. The solid circles mark the inner boundary of the simulation at $r = r_{\text{min}} = a$. The larger dotted circle at $r \approx 2.08a$ is the position of the $(m, l) = (2, 1)$ outer Lindblad resonance (shown only for reference). Surface densities are plotted with the same linear grey-scale in each panel, with the darkest regions corresponding to a maximum density of $0.8\Sigma_0$ ($0.4\Sigma_0$ for the top-left panel). Orbital motion is in the clockwise direction.
Figure 4. Snapshots of the azimuthally averaged disc surface density at different times (shown by different curves in each panel, from 0 to 4000 orbits, as labelled), and for different mass ratios (shown in different panels, from $q = 1.0$ to $0.01$, as labelled). In each panel, the solid (black) curve shows, for reference, the density profile in the point-mass ($q = 0$) case after 4000 orbits. The vertical dotted lines mark the radius where binary and viscous torques balance (from Fig. 5); these lie close to the observed cavity edges. The vertical solid lines mark the inner edge of the integration domain ($r = r_{\text{min}} = a$). In each case, the inner circumbinary disc spreads inward with time, but the density profile remains sharply truncated, with a low-density central cavity inside $r \lesssim 2a$.

The density profile is also shown for the single point-mass ($q = 0$) case, after 4000 orbits. As the figure shows, the inner circumbinary disc spreads inward with time. By 4000 orbits, the disc structure at $r/a \gtrsim 5$, where the effect of the binary is relatively small, closely follows the $q = 0$ case. However, the density profile remains sharply truncated inside $r \lesssim 2a$ (i.e. below the point-mass case, even at 4000 orbits). As a result of the initial density profile, the peak density first decreases as matter drains inward towards the holes and also outward towards the outer boundary. With time, despite the leakage of streams to the binary, inward viscous diffusion causes a gas pileup behind the cavity wall. The figure also shows that the position of the cavity wall moves slightly outward between the $t = 2000$ and 4000 snapshots. This is because as the disc settles to its lopsided quasi-steady state, the cavity size grows in the azimuthally averaged sense.

3.1.3 Torque balance and the size of the central cavity

As (the top-left panel of) Fig. 4 shows, the central cavity around the equal-mass binary extends to $r \sim 2a$. Here, we take the cavity edge $r_{\text{ce}}$ to be the radius where the negative viscous torque density matches the binary torque density (in an azimuthally averaged sense),

$$\left[ \frac{dT}{dr} \right]_{\text{bin}} + \left[ \frac{dT}{dr} \right]_{\text{visc}} = 0.$$  \hfill (10)

To gain insight into the transport of angular momentum and the clearing of the central cavity, we therefore compute the time- and azimuthally averaged torque densities from the binary potential,

$$\left\langle \frac{dT}{dr} \right\rangle_{\text{bin}} = \left\langle \frac{1}{2\pi} \int_0^{2\pi} \left( \frac{\partial}{\partial r} \Sigma(r, \phi) \frac{\partial \Omega}{\partial r} r \phi \right) \right\rangle_{t},$$  \hfill (11)

and from viscous stresses

$$\left\langle \frac{dT}{dr} \right\rangle_{\text{visc}} = 2\pi \left\langle \frac{d}{dr} \left( r^2 \left\langle \frac{\partial \Omega}{\partial r} \phi \right\rangle \right) \right\rangle_{t}.$$  \hfill (12)

The outer derivative in equation (12) is taken numerically, and all of the above values are measured directly from the simulation outputs, except for the binary potential derivative in equation (11) which is
given analytically. The time averages are taken over 25 binary orbits at a sample rate of 20 per orbit.

The top row of Fig. 5 shows the binary and viscous torque densities for $q = 1$ over the inner $5r/a$ of the disc, during both the initial transient state (left-hand panel) and the subsequent quasi-steady state (right-hand panel). There is indeed a well-defined central region, where the binary torques exceed the viscous torques and can be expected to clear a cavity. The transition (computed via equation 10) is located at $r_c \approx 1.85$ and $r = 2.02a$ in the transient and quasi-steady state, respectively, and is marked in both panels by a vertical dotted line. These vertical lines are also shown in Fig. 4 and indeed lie very close to radii where the disc surface densities remain truncated.

The small outward drift of the average position of the cavity wall from the transient to the quasi-steady state is also visible in Fig. 4. As the disc transitions to the quasi-steady state, the binary torque-density wavelength increases in $r/a$, while keeping approximately the same amplitude. These effects can be attributed to the increasingly elongated and lopsided shape of the inner cavity; when azimuthally averaged, this results in a larger cavity size.

### 3.1.4 Accretion rates

The most interesting consequence of the lopsided cavity shape is on the accretion rate. In the top-left pair of panels in Fig. 6, we show the accretion rate, measured across the inner boundary of the simulation ($r_{\text{min}} = a$) during the quasi-steady state, after 4000 binary orbits. The upper panel shows the accretion rate as a function of time for $\sim 16$ binary orbits; the solid horizontal (blue) line shows the time-averaged accretion rate during this time, and, for reference, the horizontal dashed (green) line shows the accretion rate over the same orbits in the $q = 0$ reference simulation. The average accretion rate on to the binary is approximately 2/3 of the accretion measured for the $q = 0$ run at the same resolution. Note that this ratio stays constant over the course of the quasi-steady state. The lower panel shows the corresponding Lomb–Scargle periodogram, measured over 100 binary orbits at a sampling rate of once per simulation time step ($\sim 500$ per orbit).

As mentioned above, once the quasi-steady state is reached, the accretion rate increases from that in the transient state by an order of magnitude. As Fig. 6 shows, it also begins to exhibit strong (factor of $\sim 3$ above average) variability. Although this figure samples the accretion only between 4000 and 4016 orbits, the pattern is remarkably steady, and repeats itself until the end of the simulation. (However, see point (vi) in Section 4.2.)

The accretion is clearly periodic, and displays two prominent periods $(1/2)\dot{h}_{\text{bin}}$ and $\sim 5.7\dot{h}_{\text{bin}}$. The stronger variability at one half the orbital time is due to the passage of each BH by the near-side of the lopsided disc and the corresponding stripping of gas streams from the cavity wall. These streams are then driven into the opposite side of the cavity (as seen in the top-right panel in Fig. 3; approximately 135° from the generation point). The second, longer time-scale corresponds to the orbital period at the cavity wall. As mentioned above, when the non-accreted material from the streams hits the far-side of the cavity, it creates an overdensity which orbits at the disc’s orbital period there which ranges from $\sim 2\tau(2.0a)^{1/2}(GM)^{-1/2} \sim 2.9\dot{h}_{\text{bin}}$ out to $\sim 2\tau(3.3a)^{1/2}(GM)^{-1/2} \sim 6.0\dot{h}_{\text{bin}}$. The larger streams pulled from the lump in turn create a new lump and the cycle repeats once every $\sim 5.7\dot{h}_{\text{bin}}$. Similar overdense lumps have also been found and described in the 3D MHD simulations of Shi et al. (2012); more recently, Roedig et al. (2011, 2012) have also mentioned the contributions of such lumps to fluctuations in the accretion rate.

### 3.2 Unequal-mass binaries

We next turn to the main new results of this paper, and examine the disc behaviour as a function of the mass ratio. We start with a qualitative description of how the accretion pattern changes as we decrease $q$.

#### 3.2.1 Three-time-scale regime: $0.25 < q < 1$

The behaviour of systems with mass ratio in the range $0.25 \lesssim q < 1$ is illustrated in Figs 3, 4 and 6. These show snapshots of the 2D surface density in the quasi-steady state, the evolution of the azimuthally averaged density profile, and the time-dependent accretion rates, respectively. Additionally, in Table 2, and in the corresponding Fig. 7, we show the time-averaged accretion rate as a function of $q$. In each case, the accretion rate is averaged over 1000 binary orbits in the quasi-steady state (unless noted, from 3500 to 4500 binary orbits), and is quoted in units of the corresponding rate for a single-BH ($q = 0$) disc. This ratio changes very little over the course of the quasi-steady-state regime.

These figures and table illustrate several trends as $q$ is lowered from $q = 1 \rightarrow 0.75 \rightarrow 0.5 \rightarrow 0.25$.

(i) The cavity becomes more compact, and less lopsided, as one naively expects when the binary torques are reduced. These effects are clearly visible in the middle row of Fig. 3, and also in the corresponding azimuthally averaged density profiles in Fig. 4: the profiles look remarkably similar for $q = 0.5$ and 0.25, except the cavity for $q = 0.25$ is smaller, and its wall is visibly sharper (as a result of azimuthally averaging over a less lopsided 2D distribution).

(ii) The secondary (primary) moves closer to (farther from) the cavity wall as $q$ is reduced. Thus, occurs for two reasons. First, the position of the secondary (primary) moves away from (towards) the binary’s centre of mass,

$$r_s(q) = a(1 + q)^{-1} \quad r_p(q) = a(1 + 1/q)^{-1}. \quad (13)$$

Secondly, as mentioned in (i), the size of the central cavity decreases. Fig. 4 shows the locations of the cavity edge $r_c$, expected from balancing the azimuthally averaged gravitational and viscous torques (Fig. 5), which agree well with the observed cavity sizes. In Fig. 8, we explicitly show $r_c$ as a function of $q$. The points in the figure have been obtained by balancing the azimuthally averaged viscous and gravitational torques measured in the simulation (equation 10). The black line is an empirical fit to the data points at fiducial resolution, given by

$$r_c(q) \simeq A + B \ln (q^{1/2} + 1) + C \ln (q + 1)$$

$$A = 1.191 \quad B = 2.541 \quad C = -1.350. \quad (14)$$

(iii) The dense lump, created by the shocks due to the ‘regurgitated’ stream material thrown back out by the binary, is still present for $q = 0.5$ (see the corresponding panel in Fig. 3, around 9 o’clock at $r \sim 2a$), but is much less discernible for $q = 0.25$. Again, this trend is unsurprising – as the torques diminish, one expects weaker shocks and smaller overdensities in any resulting lump.

(iv) The accretion streams in the $q = 0.25$ panel of Fig. 3, and the corresponding ripples in the azimuthally averaged density profiles of Fig. 4, become noticeably weaker. However, the average accretion rate, shown in Figs 6 and 7 and in Table 2, stays at approximately $\lesssim 0.7M_{\odot}$ over the range $0.5 \lesssim q \lesssim 1$. For $q \lesssim 0.5$, the average accretion rate drops more rapidly, falling by nearly a factor of 2 to $\lesssim 0.36M_{\odot}$ by $q = 0.1$. 

Downloaded from https://academic.oup.com/mnras/article-abstract/436/4/2997/984794 by guest on 26 July 2018
Figure 5. Azimuthally and time-averaged torque-density profiles in the inner disc for the equal-mass binary (top two panels) and for unequal-mass binaries (other panels, with different mass ratios $q$ as labelled). The top-left panel corresponds to the point-symmetric transient stage (after $\sim 2000$ orbits) and the top-right panel to the asymmetric quasi-steady state (after $\sim 4000$ orbits). Only the quasi-steady state is shown for the $q < 1$ cases. In each panel, the dashed (black) curves show the gravitational torques from the binary and the red (dot–dashed) curves show the negative viscous torques. The vertical dotted line marks the radius where the viscous and gravitational torques balance (equation 10); these are close to where the azimuthally averaged surface density profiles are found to be truncated (see Fig. 4). See Fig. 8 for a plot of this cavity edge radius versus $q$. Time averages are taken over 25 orbits at a sample rate of 20 per orbit.
Figure 6. The time variable accretion rate across the inner boundary of the simulation measured at $r = r_{\text{min}} = a$ (top of each pair of panels) and the corresponding Lomb–Scargle periodogram (bottom of each pair) computed over 100 orbits. Panels are displayed in order of decreasing binary mass ratio, starting from $q = 1.0$ at top left to $q = 0.003$ at the bottom right. The average accretion rate in each panel is denoted by the solid (blue) horizontal line. The dashed (green) horizontal line in each plot shows the average accretion rate for the point mass ($q = 0$) case for reference. For $q > 0.05$, the accretion rate is strongly modulated by the binary, with either one, two or three distinct periods present simultaneously, depending on the value of $q$ (see the text for detailed explanations). For $q \lesssim 0.05$, the binary still reduces the mean accretion rate noticeably, but does not imprint strong periodicity; the $q = 0.003$ binary is nearly indistinguishable from a single BH.
Table 2. The mean accretion rate $\dot{M}_{\text{bin}}$, averaged over 1000 orbits in the quasi-steady state, for binaries with different mass ratios. The rates are shown in units of the corresponding rate $\dot{M}_q$ found in a single-BH ($q = 0$) simulation. This ratio is computed as an average from 3500 to 4500 orbits unless the quasi-steady state is not reached until after (or for large $\alpha$ much before) 3500 orbits; in this case the value in the table is denoted by $\ast$. The first four rows show results for different combinations of radial and azimuthal resolutions. The first row is our fiducial resolution. The last three rows are for runs at the fiducial resolution but different magnitudes of the viscosity parameter $\alpha$.

| $q$ | 1.0  | 0.75 | 0.5  | 0.25 | 0.075 | 0.05 | 0.025 | 0.01 | 0.003 |
|-----|------|------|------|------|-------|------|-------|------|------|
| $M_{\text{bin}}/\dot{M}_0$ [Mid$\Delta r$, Lo$\Delta \phi$] | 0.671 | 0.700 | 0.655 | 0.382 | 0.355$\ast$ | 0.228$\ast$ | 0.025 | 0.100 | 0.341 | 0.814 |
| $M_{\text{bin}}/\dot{M}_0$ [Lo$\Delta r$, Lo$\Delta \phi$] | 0.544 | 0.518 | 0.499 | 0.406 | 0.021 | 0.027 | 0.055 | 0.147 | 0.381 |
| $M_{\text{bin}}/\dot{M}_0$ [Mid$\Delta r$, Mid$\Delta \phi$] | 0.821 | 0.426 | 0.021 |
| $M_{\text{bin}}/\dot{M}_0$ [Hi$\Delta r$, Hi$\Delta \phi$] | 0.930 | 0.724 | 0.172$\ast$ |
| ($\alpha = 0.02$) | 0.899 | |
| ($\alpha = 0.04$) | 0.921 | |
| ($\alpha = 0.1$) | 1.015$\ast$ | |

(v) As the average accretion rate decreases, so does the maximum accretion rate (i.e. the amplitude of the spikes in Fig. 6) keeping an approximately constant enhancement of a factor of $\sim 3$ as the mass ratio is decreased to $q \sim 0.1$.

(vi) The percentage of a stream which leaves the domain at $r_{\text{min}} = a$ as opposed to being flung back out also decreases by a factor of $\sim 2$ as $q$ decreases from 1.0 to 0.1. We measure this percentage from the simulations by computing the ratio $M(r_{\text{min}})/M(r_{\text{ce}})$ averaged over 25 orbits in the quasi-steady state. Here, $r_{\text{ce}} = 0.95r_{\text{cc}}$ (with $r_{\text{cc}}$ given by equation 14) is chosen to be just inside the cavity wall, where the accretion rate is dominated by the streams. The percentage drops by approximately a factor of 2, from $\sim 3.3$ per cent at $q = 1$ to $\sim 1.8$ per cent at $q = 0.1$, suggesting that the drop in average accretion rate in the three-time-scale regime is largely due to the amount of stream material which can penetrate beyond the binary torque barrier at small $r$.

(vii) Perhaps the most interesting result is shown by the Lomb–Scargle periodograms in Fig. 6. As $q$ decreases, power is traded from both the $(1/2)t_{\text{bin}}$ and the $5.7t_{\text{bin}}$ variability time-scales into the $t_{\text{bin}}$ time-scale. This is because of the increased proximity between the secondary and the cavity wall, and a corresponding larger distance between the primary and the cavity wall noted above. As a result, as $q$ is decreased, the secondary begins to dominate the variability, pulling accretion streams off of the cavity wall once per binary orbit.

Focusing on the last finding: in the $0.25 \lesssim q < 1$ case, we find that the time-dependent accretion rate displays three distinct and sharply defined periods, with well-defined ratios at $0.5$, $1$ and $5.7t_{\text{bin}}$. While the last of these reflects the orbit of the dense lump at the elongated cavity wall and could depend on details of the disc properties, the first two periods are fixed by the binary alone and are independent
The position of the cavity wall as a function of \( q \) is not always as expected. For instance, at \( q \approx 0.075 \), the fluctuations are still large (factor of \( \leq 0.25 \)), the lump may not survive. However, the reason this lump disappears is less obvious, and warrants some discussion.

(i) As \( q \) decreases, the cavity becomes less lopsided, and the accretion rate spikes become weaker. This suggests that when these weaker accretion streams are flung back to the cavity wall, they create less overdense lumps. For \( q \leq 0.25 \), the lump may not survive shear stresses and pressure forces, and may dissolve in less than an orbital time.

(ii) As can be eyeballed from Fig. 3, the stream impact zone (dense region outside the cavity) extends over an azimuth of \( \Delta \phi \approx 100^\circ - 120^\circ \). The orbital period at the cavity edge is \( \leq 6t_{\text{bin}} \) (the orbital period at the furthest edge of the cavity), implying that multiple streams can hit parts of the same lump if streams are generated more than once per binary orbit.

To test whether the strength or the frequency of the streams is more important for lump generation, we repeat our simulation for an equal-mass binary, but we placed one of the BHs artificially at what would be the real binary’s centre of mass. The second hole still orbits at \( r = a/2 \) as usual. In this setup, the cavity wall is perturbed by streams with a similar strength as in the real \( q = 1 \) simulation (however, much less of the streams reach the inner edge of the simulation domain in this one-armed perturber case), but now only once, rather than twice per orbit. We found that, while this ‘one-armed’ binary does generate significant stream impacts at the cavity edge, it does not create an orbiting overdensity at the cavity edge, nor does it excite a significant elongation of the cavity. We therefore conclude that multiple, overlapping streams are required to generate a strong lump that survives for an orbital time. However, as the \( q = 0.25 \) accretion rate and periodogram show in Fig. 6, simply generating two streams is not a sufficient condition for generating a cavity wall lump. Both streams must also be sufficiently large. This explains the disappearance of the cavity wall frequency for \( q \leq 0.25 \); as the mass ratio decreases, the primary generates less significant streams and the overlap of two large streams crashing into the cavity wall can no longer occur to generate an overdense lump there. This result is also consistent with the cavity becoming less lopsided as \( q \) decreases.

3.2.3 Steady-accretion regime: \( q \leq 0.05 \)

As we continue to decrease \( q \) from 0.075 through 0.05 to 0.01, we find yet another distinct regime. The overall morphology of the snapshot of a \( q = 0.01 \) and 0.05 (not shown) disc in Fig. 3 looks similar to the \( q = 0.075 \) case, except the nearly concentric perturbations are even weaker, and the cavity still smaller. However, the similarity is quite deceptive, with the movie versions of these figures showing a striking difference.\(^5\) In the \( q = 0.075 \) case, accretion streams form and disappear periodically, but in the \( q = 0.05 \) case, the disc pattern becomes constant and unchanging (in the frame corotating with the central disc).

\(^5\) Movie versions of the snapshots in Fig. 3 are available at http://www.astro.columbia.edu/~dorazio/moviespage
binary). There is still a visible accretion stream, hitting the inner boundary of the simulation just ahead of the secondary’s orbit, but the stream steadily coagulates with the binary.

As Fig. 6 shows (see also Fig. 7 and Table 2), the average accretion rate has reached its minimum at \( q = 0.05 \). For \( q \leq 0.05 \), the accretion rate becomes steady, with no fluctuations, and its value increases back towards the \( q = 0 \) rate [dashed horizontal (green) line in Fig. 6]. For such a light secondary, the system begins to resemble a disc with a single BH. Although a cavity is still clearly present (moving from \( r \simeq 1.6a \) at \( q = 0.05 \) to \( r \simeq 1.4a \) at \( q = 0.01 \)), it is being refilled, as for a single BH. Indeed, after 4000 orbits, the \( q = 0.01 \) azimuthally averaged density profile is approaching the \( q = 0 \) profile (see Fig. 4).

Although the secondary still excites small but visible ripples in the disc (Fig. 3), by \( q \leq 0.05 \) it can no longer exert a large enough torque to pull in large streams and drive them back out to produce a lopsidedness in the circumbinary disc. The ripples are in the linear regime, and they resemble the tightly wound spiral density waves launched in protoplanetary discs (e.g. Goldreich & Tremaine 1980; Dong et al. 2011a; Duffell & MacFadyen 2012), except here our background discs have a pre-imposed central cavity. In the linear regime, the waves are excited by resonant interactions with the disc, and non-linear coupling to an \( m = 1 \) mode is no longer possible. Note the disappearance of any ripples in the azimuthally averaged surface density for \( q = 0.01 \) (Fig. 4).

4 SUMMARY AND DISCUSSION

In summary, we find that the behaviour of the accretion rate across the circumbinary cavity as a function of \( q \) can be categorized into four distinct regimes.

(i) Two-time-scale regime; \( q = 1 \). Confirming previous results, an equal-mass binary maintains a central low-density cavity of size \( r \sim 2a \) and the time-averaged accretion rate is \( \sim 2/3 \) of that for a point-mass case. There are up to factor of \( \sim 3 \) fluctuations around the average on two prominent time-scales \( (1/2)\bmin \) and \( \sim 5.7\bmin \).

(ii) Three-time-scale regime; \( 0.25 < q < 1 \). The time-averaged accretion rate drops by a factor of \( q = 0.25 \); however, the maximum fluctuations continue to occur with amplitude \( \sim 3 \) times the average rate. There are three time-scales present \( (1/2)\bmin, \bmin \) and \( \sim 5.7\bmin \).

(iii) Single- and orbital-time-scale regime; \( 0.05 \lesssim q < 0.25 \). In this regime, the average accretion rate and its fluctuations continue to drop with decreasing \( q \). Variability is dominated by the secondary, and is nearly sinusoidal on the binary period \( \bmin \) though accompanied by small accretion spikes due to the primary with maxima below the average rate.

(iv) Steady-accretion regime; \( q \lesssim 0.05 \). The accretion becomes steady while reaching its lowest rate at \( q \sim 0.05 \). By \( q = 0.003 \), the accretion rate rises again to \( \sim 0.8 \) of the \( q = 0 \) case. The system overall resembles a cavity-filling single-BH disc, with small perturbations due to the secondary in the linear regime.

4.1 Comparison with MM08

Our main qualitative conclusions for the equal-mass binary case, including the morphology of the disc, and the accretion rate, are in good agreement with MM08. Nevertheless we do find a small discrepancy in the time-averaged accretion rates.

Comparing MM08’s figs 7 and 8 with the top-left panel of our Fig. 9, which is run at the same resolution as the highest resolution

![Figure 9. Accretion rates at the inner boundary \( r_{\text{bin}} = a \) for the equal-mass binary as in the top-left panels of Fig. 6, except for the lowest resolution runs used in this study which matches the highest resolution used in the disc simulated by MM08. The solid horizontal (blue) line in the top panel is the average accretion rate and the bottom panel shows the Lomb–Scargle periodogram computed over 100 binary orbits.](https://academic.oup.com/mnras/article-abstract/436/4/2997/984794)
it is that the time-scale associated with the orbital period of the overdense lump at the cavity wall is more prominent for larger $\alpha$. However, it is possible that, for even larger $\alpha$, the overdense lump could break up before it can seed another lump via stream generation.

(vi) For the larger $\alpha$ runs, the top panels of Fig. 11 show the appearance of a longer variability time-scale with the same period as the lopsided cavity precession – once per $\sim 400$ orbits. This variability manifests itself in a modulation of the maximum accretion rate achieved by the largest streams pulled from the cavity edge lump; every $\sim 400$ orbits the largest accretion rate spikes reach 30 per cent higher above the average than they do $\sim 200$ orbits later. Note that in the fiducial $\alpha$ case, there is a similar long-term variation in the strength of the $5.7r_{bin}$ modulation, but it occurs more erratically and with approximately half of the total variation.

(vii) Finally, the top panels of Fig. 11 also show that the quasi-steady, lopsided mode occurs much earlier for larger $\alpha$. For the fiducial case the transition takes place after $\sim 2500r_{bin}$, for $\alpha = 0.02$ after $\sim 1000r_{bin}$ and for larger $\alpha = 0.04, 0.1$ after less than a few 100 orbits which is set largely by the time for fluid to diffuse to the inner regions of the disc.

A more detailed investigation of the effects of viscosity should be carried out in future studies and understood self-consistently from simulations which generate turbulent viscosity via the magnetorotational instability (see Noble et al. 2012; Shi et al. 2012).

4.3 Resolution study

Up to now we have discussed the results from our fiducial set of medium-radial, low azimuthal resolution runs ([Mid$\Delta r$ and Lo$\Delta \phi$] in Table 1). Ideally, we would repeat these runs at increasingly high radial and azimuthal resolutions, until the results converge. Unfortunately, this is computationally prohibitive, and we instead choose the following approach.

(i) For the $q = 1.0$ case, we perform two higher resolution runs ([Mid$\Delta r$ and Mid$\Delta \phi$] and [Hi$\Delta r$, Hi$\Delta \phi$] in Table 1) and one lower resolution run ([Lo$\Delta r$ and Lo$\Delta \phi$] in Table 1) to look for signs of convergence.

(ii) We then explore the resolution sensitivity of the boundaries between each accretion regime. The $q = 0.1$ case is at the cusp of the three-time-scale and single-time-scale regimes, where we expect the accretion behaviour to be particularly sensitive to $q$. The $q = 0.05$ case is at the cusp of the single-time-scale and steady-accretion regimes, where the minimum accretion rates are achieved. Thus, we also run the $q = 0.1$ and 0.05 cases for the same set of resolutions as the $q = 1.0$ case.

(iii) We repeated each of our runs at the lowest resolution. Re-doing the entire set of runs allows us to assess whether different mass-ratios are affected by resolution differently.

Fig. 12 gives a visual impression of the surface density distribution at the four different combinations of resolutions for $q = 0.1$. They show a clear trend: as the resolution is increased, the lumps near the cavity wall become sharper and more overdense, and the cavity becomes larger and more lopsided.\footnote{Fig. 8 shows that the average position of the cavity wall, found from the azimuthally averaged torques in equation (10), is much less affected over the range of resolutions studied here.}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figures/figure10.png}
\caption{2D surface density distributions during the quasi-steady-state phase, as in Fig. 3, except for the single mass ratio $q = 1.0$, and for four different values of the viscosity parameter $\alpha$, as labelled. Increasing $\alpha$ causes ripples created by streams impacting the cavity wall to smear out more quickly causing the surface density snapshots to appear smoother. For all values of $\alpha$ shown here, the over-dense lumps still survive long enough to create the $\sim (5-6)r_{bin}$ modulation of the accretion rate. Also for larger $\alpha$, the near-side of the disc extends in closer to the binary causing a larger fraction of streams to exit the integration domain at $r = a$. This results in higher measured accretion rates relative to the point mass values for the same $\alpha$s.}
\end{figure}
the resolution implies weaker numerical diffusion, stronger accretion streams, more momentum carried by these streams into the disc and an overall more efficient driving of the $m = 1$ mode. This is further evidenced by an earlier onset of the elongated mode for the higher resolution $q = 0.1$ simulations. The $q = 0.1$ simulations develop an elongated cavity after $\sim 1500$ (highest resolution), $\sim 2500$ (medium resolution) and $\sim 3500$ (fiducial resolution) binary orbits. The lowest resolution run never develops an elongated cavity even after $\sim 7000$ binary orbits.

The 2D surface density profiles for the $q = 1.0$ runs at different resolutions remain qualitatively the same as the fiducial resolution counterpart. We do find that as the azimuthal resolution is increased the cavity becomes slightly more elongated likely again due to more efficient stream impacts. The 2D surface density profiles for the $q = 0.05$ runs at different resolutions remain qualitatively the same for all resolutions except the highest resolution. For the highest resolution $q = 0.05$ run, the disc transitions into the single-time-scale regime only after...
Accretion into the central cavity

The accretion rates for the \( q = 0.05 \) runs at different resolutions again remain nearly identical for all resolutions except the highest resolution. For the highest resolution \( q = 0.05 \) run, the disc transitions into the single-, orbital-time-scale regime after \( \sim 4000 \) binary orbits and exhibits modulation of the accretion rate at the orbital frequency, mimicking the \( q = 0.075 \) accretion rates at the fiducial resolution.

Table 2 and the corresponding Fig. 7 show the time-averaged accretion rates as a function of \( q \) at different resolutions. In all cases, at the same fixed \( q \), we find that increasing resolution produces a higher accretion rate. This is consistent with the interpretation above that higher resolution allows stronger accretion streams. Interestingly, we find a strong correlation between the values of \( \frac{M_{\text{bin}}}{M_{q0}} \) listed in Table 2 and the accretion patterns seen in Fig. 13: runs at different resolutions but with similar values of \( \frac{M_{\text{bin}}}{M_{q0}} \) have very similar accretion patterns (including the variability and the values of the maxima). The result of increasing (decreasing) the resolution can therefore be interpreted as a shift of the accretion behaviour to lower (higher) mass ratios.

Comparing the full set of mass ratio runs for the lowest resolution to the fiducial resolution runs, we observe the same progression of the accretion rate through each of the accretion variability regimes discussed above; a difference being that, as discussed in the previous paragraph, the boundaries between each regime are delineated at larger mass ratios in the lowest resolution runs. We also notice that in the three-time-scale regime, there is more power in the periodogram peak associated with the cavity wall orbital period (e.g. compare the top left of Fig. 6 with Fig. 9). The cavity wall peak still disappears for \( q = 0.25 \) when the stream due to the primary becomes much smaller than the stream due to the secondary and the overlapping of large streams at the cavity wall no longer generates an overdense lump.

Encouragingly, the two higher resolution runs at \( q = 1.0 \), also plotted in Fig. 7, lie closer to each other than the two lower resolution runs. Since the resolution steps are evenly spaced, we consider this evidence that the simulations are converging monotonically with resolution. Since the \( q = 0.1 \) and 0.05 discs are positioned at the boundary of different accretion regimes, we find a large dependence of disc response and accretion rate on resolution, but with a clear trend of increasing resolution moving the boundaries between the accretion regimes described in this study to slightly lower values of the binary mass ratio.

4.4 Physical regime: black hole binary parameters

The simulations presented above can be scaled, in principle, to any BH mass and orbital separation. In this section, we discuss the physical scales for which our simulations could be relevant (i.e. physically viable and observationally interesting). The shaded region in Fig. 14 plots this relevant portion of parameter space by imposing the following restrictions.

(i) \( 10^{-6} M_\odot \lesssim M_p \lesssim 10^4 M_\odot \). It is not clear whether smaller BHs exist in galactic nuclei, and, in any case, the radiation from such a low-mass BHs would likely be too faint to detect. Likewise, much more massive BHs are known to be rare.

(ii) \( 10^{-2} \lesssim q \lesssim 1 \). As we have shown, for the set-up we study (i.e. with a cavity inserted by hand into the disc), the accretion pattern converges as we decrease the mass ratio to \( q = 0.01 \) and below. In practice, a physical lower limit of \( q \approx 0.01 \) may arise from the fact that bound binary BHs can be created only in relatively major mergers. In a minor merger, the smaller satellite galaxy may be
Figure 13. The time-dependent accretion rates as in Fig. 6, except for the single mass ratio \( q = 0.1 \) and the four different resolutions also shown in Fig. 12. The \( q = 0.1 \) binary is at the cusp of the transition from the three-time-scale to the single-time-scale regime, and is particularly sensitive to resolution, as seen especially in the maxima of the accretion spikes.
Coupled with the well-established correlations between the mass of an MBH and its host galaxy, this suggests that the $q$-distribution may not extend to values significantly below $q \sim 0.01$. Fig. 14 plots the restriction of $a$, $M$ parameter space for these limiting mass ratios $q = 1$ (left) and $q = 0.01$ (right).

(iii) The binary is embedded in a thin gaseous disc. Following a merger of two MBH-harbouring galaxies, the MBHs sink to the bottom of the new galactic potential via dynamical friction in approximately a galactic dynamical time-scale (Begelman, Blandford & Rees 1980). In addition to stellar interactions (e.g. Preto et al. 2011), many studies have shown that gas in the vicinity of the binary could aid in hardening the binary down to $\sim$parsec separations (e.g. Escala et al. 2005; Dotti et al. 2007; Mayer et al. 2007; Cuadra et al. 2009; Lodato et al. 2009; Nixon et al. 2011a; Chapon, Mayer & Teyssier 2013). We have assumed that such gas in the vicinity of the binary cools efficiently and forms a rotationally supported thin disc. For a given $M_1$ and $M_2$, we are still free to choose a physical distance for the orbital radius $a$, which could correspond to a snapshot of the binary anywhere along its orbital decay. The assumption that the binary is embedded in a thin disc allows us to make the following additional constraints on $a$ given $M_2$.

(a) The accretion disc is gravitationally stable. Accretion discs become self-gravitating, and unstable to fragmentation, beyond a radius of $r_{\text{in}} \sim M_{\text{bin}}/10^3 M_\odot \cdot a_{\text{bin}}$ where $a_{\text{bin}}$ is the Schwarzschild radius; see, e.g. Goodman 2003; Haiman et al. 2009b for the formula used to generate the $Q \leq 1$ criteria in Fig. 14). Since the binary has to fit inside a gravitationally stable disc, this puts an upper limit on the orbital separation denoted by the red lines in Fig. 14.

(b) Variability occurs on an observable time-scale. The binary orbital time is given by

$$t_{\text{bin}} = \frac{2\pi}{\Omega} = 0.88 \left(\frac{M}{10^3 M_\odot}\right) \left(\frac{a}{10^3 r_s}\right)^{3/2} \text{yr.} \quad (15)$$

As we have shown, the accretion rate shows periodicity on a time-scale of $\sim t_{\text{bin}}$. In a realistic survey, it will be feasible to look for periodic variations between $0.1 h \lesssim t_{\text{bin}} \lesssim 1$ yr denoted by the solid blue lines in Fig. 14. Here, the lower limit comes from the integration time required to measure the flux variations for MBHBs in the above mass range [for a survey instrument with a sensitivity similar to Large Synoptic Survey Telescope (LSST); Haiman et al. 2009b], and the upper limit comes from the duration of proposed time domain surveys. As a guide, the dashed blue lines in Fig. 14 are contours of constant orbital times drawn at 10 d and 1 yr.

---

Footnote:

8 In principle, less massive BHs may grow from the accretion discs around the primary (McKernan et al. 2012); the long-term evolution of such systems would be worthy of further study.
Throughout our simulations, $\alpha$ is assumed to be of the order of $a^{-1}$, as a function of the binary separation $a$, due to migration of the secondary through the disc and due to GW decay at small enough binary separations. In Fig. 14, the green lines denote the requirement that $10^5 \leq t_{\text{res}} \leq 10^{10}$ yr. A residence time of greater than $10^{10}$ yr does not on its own exclude a binary system from observation. Nevertheless, we include this limit in order to show which binaries will not merge (due to migration through a gaseous disc) in a Hubble time. Note also that there is a trade off: a longer residence time is desirable since it increases the probability of finding such a system; however, longer residence times occur at larger separations and longer orbital times, which will make it more difficult to verify any periodic behaviour.

(d) A cavity is maintained. For consistency with the initial conditions adopted here, we require that the binary + disc systems will indeed form a cavity during earlier stages of their evolution. The region of parameter space for which a cavity may be maintained is denoted by the orange lines in Fig. 14 and is calculated using the steady-state disc solutions detailed in Kocsis et al. (2012b).

(e) The orbital separation is fixed. Throughout our simulations, we fix the binary separation; we therefore require that the orbital decay should be slow enough for the binary’s orbit not to change significantly over a few thousand orbits. This is denoted by the dashed black lines in Fig. 14 which are drawn where $N_{\text{orb}} = t_{\text{res}}/\tau_{\text{bin}} = 10^5$.

(iv) Though not expressed in Fig. 14, for our simulations to be self-consistent, we also require $a \gtrsim 100 r_5$, since our Newtonian treatment ignores general relativity. Furthermore, at approximately the same binary separation, the orbital decay of the binary due to GW emission becomes more rapid than the viscous time at the edge of the cavity. As a result, the disc decouples from the binary and is ‘left behind’, rendering our initial conditions inconsistent in this regime (e.g. Milosavljević & Phinney 2005; although see Farris et al. 2012b; Noble et al. 2012 whose MHD simulations suggest that the gas can follow the binary down to smaller separations).

4.5 Caveats

However instructive, the simulations presented here are still of course simplified models of a real binary disc system, and it is worth listing some major caveats.

(i) Our simulations are 2D – we expect that the 3D vertical structure could modify the structure of the accretion streams, including their $q$-dependence. However, note that Roedig et al. (2012) find similar features in the accretion rate periodograms measured from 3D simulations.

(ii) Our discs are assumed to have angular momentum co-aligned with that of the binary (prograde discs). In principle, a random accretion event on to the binary could result in a misaligned disc which will eventually be torqued into co- or counter-alignment with the binary angular momentum (Nixon, King & Pringle 2011b).

(iii) Our simple $\alpha$-viscosity prescription may be inaccurate, especially in the nearly radial accretion streams. Recent MHD simulations in the $q = 1$ case find a larger effective $\alpha$ than the fiducial value adopted here (Noble et al. 2012; Shi et al. 2012). It is interesting, however, that our highest $\alpha = 0.1$ simulation for an equal-mass binary exhibits the same variability as our fiducial case but a larger accretion rate, in agreement with the above mentioned 3D MHD simulations.

(iv) We assumed a locally isothermal equation of state; more realistic equations of state could have a especially large impact on the strength and dissipation of shocks at the cavity wall.

(v) Our initial conditions correspond to an unperturbed, near-Keplerian, circular disc and circular binary orbit with a significant pile-up of gas. In reality, accretion on to the binary could produce significant binary (as well as disc) eccentricity (e.g. Cuadra et al. 2009; Lodato et al. 2009; Roedig et al. 2011). In this case, the variability we find would most likely have a more complex structure (e.g. Hayasaki et al. 2007) due to the plethora of resonances available in an eccentric binary potential (e.g. Artymowicz & Lubow 1994). In a study of circumbinary discs around eccentric binaries (with mass ratio 1/3), Roedig et al. (2011) find similar accretion rate periodograms as in this study; periodogram peaks exist at the orbital frequency, twice the orbital frequency and the cavity wall orbital frequency. They find that an increasingly eccentric binary: (a) increases the size of the cavity and thus decreases the overall magnitude of the accretion rate, (b) enhances power at harmonics of the orbital frequency and (c) increases power in a peak located at the beat frequency of the orbital frequency and the cavity wall frequency.

(vi) As previously stressed, these simulations do not allow the binary orbit to evolve in response to forces exerted by the gas disc. When this assumption of a massless disc is lifted, in addition to changes in binary eccentricity and semimajor axis, the binary centre of mass could oscillate around the disc centre of mass due to the orbiting eccentric disc. For massive discs, the above effects could alter the description of disc evolution, and hence accretion, presented in this zero disc-mass study.

(vii) Although we begin with an ‘empty’ cavity, this cavity may overflow already at a large radius (Kocsis et al. 2012b). Future studies should construct a self-consistent initial density profile, by evolving the binary’s orbit from large radius through gap clearing, and thus determining whether a true pile-up occurs.

(viii) We have ignored the radiation from the gas accreted on to the BHs. Given that we find high accretion rates – comparable to those for a single BH – the secondary BH can be fed at super-Eddington rates, and the flow dynamics can then be strongly affected by the radiation.

(ix) We have not allowed accretion on to the BHs, and have excised the inner region $r < a$ from the simulation domain. This could have an impact on the dynamics of the streams that are flung back towards the cavity wall, and therefore on the formation of the dense lump, the lopsidedness of the cavity, and the variable accretion patterns.

These caveats should all be pursued in future work, to assess the robustness of our results. We expect that our main conclusions, namely that the accretion rate is strongly modulated by the binary, and that the power spectrum of the accretion shows distinct periods, corresponding to the orbital periods of the binary and the gas near the location of the cavity wall, will be robust to all of the above caveats. However, the numerical values, such as the mean accretion rate, and the critical value of $q$ for the transition between variable and steady accretion, will likely be affected.

5 CONCLUSIONS

We have investigated the response of an accretion disc to an enclosed binary via 2D, Newtonian, hydrodynamical simulations. As previous work has shown (Artymowicz & Lubow 1994; Hayasaki et al. 2007; MM08; Cuadra et al. 2009; Roedig et al. 2012; Shi et al.
2012), for non-extreme mass ratios, the binary carves out a cavity in the disc, but gas still penetrates the cavity in streams which possibly accrete on to the binary components. Here, we have followed up on the work of MM08 by investigating the nature of this inflow across the circumbinary cavity, as a function of binary mass ratio $q$. We have simulated 10 different mass ratios in the range $0.003 \leq q \leq 1$. This corresponds to the expected range of $q$-values for massive BH binaries produced in galaxy–galaxy mergers. We find that while the binary ‘propellers’ are effective at maintaining a low-density cavity at the centre of the disc, they cannot efficiently suppress accretion across the cavity. For $q = 1$, the average accretion rate is of the order of 2/3 that of a single BH with accretion spikes of $\sim 3$ times larger. As long as the circumbinary disc is fuelled at a near-Eddington rate from large radius, these binaries could therefore have quasar-like luminosities. This should facilitate finding counterparts to GW events (Kocsis et al. 2006), and should also allow their detection in EM surveys (Haiman et al. 2009b).

We have found that the accretion is not only strong, but can be strongly variable (by a factor of $\sim 3$), with a characteristic $q$-dependent frequency pattern. While the accretion for $q < 0.05$ is steady, for $q \gtrsim 0.05$ there is a strong modulation by the binary, and a clear dependence on $q$ of both the variability pattern and the magnitude of the time-averaged accretion rate. For an equal-mass binary, the accretion rate is modulated at twice the orbital frequency and $\sim 1/6$ the orbital frequency. As the mass ratio is lowered, the power in the $1/2h_{\text{bin}}$ and $(5-6)h_{\text{bin}}$ variability time-scales is reduced, and traded for a third variability time-scale at $h_{\text{bin}}$. In the range $0.05 \lesssim q \lesssim 0.25$, the single $h_{\text{bin}}$ time-scale is dominant.

Increasing the magnitude of viscous forces has little effect on the above findings except to increase the magnitude of the accretion rate (both absolute and relative to $q = 0$) and to bring out a long-term accretion variability time-scale with a periodicity of $400 h_{\text{bin}}$. However, accretion discs with even larger viscous forces could quench the $(5-6)h_{\text{bin}}$ variability time-scale if the overdense lump responsible for its generation can be broken up before it repeats $\sim$ an orbit at the cavity wall. Hence, further investigation into the effects of viscosity is warranted.

Strong and highly variable accretion, with characteristic frequencies, should aid in identifying massive BH binaries in galactic nuclei. The presence of two frequencies, in the ratio 1:2 for unequal-mass binaries ($0.1 \leq q < 1$), is an especially robust prediction that is independent of disc properties, and could serve as a ‘smoking gun’ evidence for the presence of a binary. Our results suggest that the ratio of the power at these two frequencies could probe the mass ratio $q$, while other features of the periodogram could probe properties of the disc, such as its viscosity.

The variability time-scales are of the order of the orbital period, and we have argued that the most promising candidates in a blind EM search would be those with total mass and separation contained in the shaded regions of Fig. 14; $10^{6-7} M_{\odot}$ binaries, preferably with orbital periods of days to weeks. The time-varying accretion to the central regions could produce corresponding variability in broadband luminosities, allowing a search in a large time-domain survey, such as LSST, without spectroscopy. Additionally, the emission lines could exhibit periodic shifts in both amplitude and frequency; kinematic effects from the binary’s orbit could be distinguished from those due to the fluctuating accretion rate, whenever the latter contains multiples of the binary period.

A few per cent of the accretion streams generated periodically fuel the BHs, but the majority of the stream material is flung back and hits the accretion disc farther out. The shocks produced at these impact sites are prominent for $q \gtrsim 0.1$, and can provide additional observable signatures. In particular, radiation from these shocks should be temporally correlated with the luminosity modulations arising near the secondary and/or primary BH, with a delay time of the order of a binary orbit.

GW observatories, such as eLISA and PTAs will be able to constrain the mass ratios of inspiralling MBHB’s at the centres of galactic nuclei ab initio, providing a template for the expected variability pattern. This should be helpful in identifying the unique EM counterpart among the many candidates (with luminosity variations) in the eLISA/PTA error box, as the source with a matching period.

In summary, our results imply that massive BH binaries can be both bright and exhibit strongly luminosity variations, at the factor of several level. This raises the hopes that they can be identified in a future, suitably designed EM survey, based on their periodic variability. Although encouraging, these conclusions are drawn from simplified 2D hydrodynamical models of a real binary disc system, and should be confirmed in future work.

ACKNOWLEDGEMENTS

We thank Brian Farris and Bence Kocsis for useful discussions. Also we thank Paul Duffell for useful discussions, verifying some of our results in independent runs with the code discos and also for suggesting the Cartesian shear flow test of our viscosity implementation. We also thank the anonymous referee for constructive comments which have improved this manuscript. We acknowledge support from NASA grant NNX11AE05G (to ZH and AM). DJD acknowledges support by a National Science Foundation Graduate Research Fellowship under grant no. DGE1144155.

REFERENCES

Amaro-Seoane P. et al., 2013, GW Notes, 6, 4
Artymowicz P., Lubow S. H., 1994, ApJ, 421, 651
Artymowicz P., Lubow S. H., 1996, ApJ, 467, L77
Barnes J. E., Hernquist L., 1992, ARA&A, 30, 705
Begel M. C., Blandford R. D., Rees M. J., 1980, Nat, 287, 307
Escala A., Larson R. B., Coppi P. S., Mardones D., 2005, ApJ, 630, 152
Chapon D., Mayer L., Teysier R., 2013, MNRAS, 429, 3114
Cuadra J., Armitage P. J., Alexander R. D., Begelman M. C., 2009, MNRAS, 393, 1423
Dong R., Rafikov R. R., Stone J. M., Petrovich C., 2011a, ApJ, 741, 56
Dong R., Rafikov R. R., Stone J. M., 2011b, ApJ, 741, 57
Dotti M., Colpi M., Haardt F., Mayer L., 2007, MNRAS, 379, 956
Duffell P. C., MacFadyen A. I., 2012, ApJ, 755, 7
Edgar R. G., 2006, preprint (astro-ph/0609756)
Eracleous M., Boroson T. A., Halpern J. P., Liu J., 2012, ApJS, 201, 23
Escala A., Larson R. B., Coppi P. S., Mardones D., 2005, ApJ, 630, 152
Farris B. D., Gold R., Paschalidis V., Etienne Z. B., Shapiro S. L., 2012a, Phys. Rev. Lett., 109, 221102
Farris B. D., Gold R., Paschalidis V., Etienne Z. B., Shapiro S. L., 2012b, abstract #G8.044, APS April Meeting 2012. Am. Phys. Soc., New York
Fryxell B. et al., 2000, ApJS, 131, 273
Gergely L. Á., Biermann P. L., 2012, preprint (arXiv:1208.5251)
Goldreich P., Tremaine S., 1980, ApJ, 241, 425
Goodman J., 2003, MNRAS, 339, 937
Goodman J., 2003, MNRAS, 339, 937
Haiman Z., Kocsis B., Menou K., Lippai Z., Frei Z., 2009a, Class. Quantum Grav., 26, 094032
stress tensor $\sigma$. Writing out (A1) with (A2),
\[ \partial_t (\rho v_i) = \nabla_i (P g_j^j) + \nabla_j (\rho v_i v^j - \sigma_j) \]
or in vector notation
\[ \partial_t (\rho v) = - \nabla P - \nabla \cdot (\rho v v - \sigma). \] (A3)

Thus, we see that the effects of viscosity can be incorporated by computing viscous momentum fluxes from $\sigma$ and subtracting them from the mechanical transport term $\rho v v$. This is what FLASH does currently to effect the use of viscosity in Cartesian coordinates. However, in non-Cartesian coordinates, there will also be geometric source terms from taking the divergence of the rank-two-tensors $\varepsilon$ and $\sigma$. Thus, we must compute not only the components of these terms, but also the divergence in order to identify geometric source terms.

A3 Components of $\sigma$ in polar coordinates

To compute the viscous stress tensor components, we work in a coordinate basis to evaluate the covariant derivatives in terms of Christoffel symbols,
\[ \nabla_i T_{ij} = \partial_i T_{ij} - \Gamma^k_{ij} T_{ki}. \]

Working in 2D polar coordinates, $(r, \phi)$, the non-zero Christoffel symbols are
\[ \Gamma^r_{\phi \phi} = - r, \quad \Gamma^\phi_{\phi \phi} = \Gamma^\phi_{\phi r} = 1/r. \]

In the polar coordinate basis equations (A4) becomes
\[ \sigma_{rr} = \Sigma \left[ 2 v \partial_r v_r + (\zeta - 1) \nabla \cdot v \right] \]
\[ \sigma_{\phi \phi} = \Sigma \left[ 2 v \left( \partial_\phi v_\phi + \partial_\rho v_\rho - \frac{2}{r^2} v_{\rho \phi} \right) \right], \]
\[ \sigma_{\phi r} = \Sigma \left[ 2 v \left( \partial_\phi v_r + (\zeta - 1) \nabla \cdot v \right) \right], \]
\[ r^2 \sigma_{\phi \phi} = \Sigma \left[ 2 v \left( \partial_\phi v_\phi + (\zeta - 1) r^2 \nabla \cdot v \right) \right], \]
where $\Sigma$ is the height integrated 2D surface density, $v^\phi = \Omega$ and $v_\phi = r^2 \Omega$, $\Omega$ being the angular frequency. Transforming to an orthonormal basis (used in FLASH) these components become, \[ \sigma_{rr} = \Sigma \left[ 2 v \partial_r v_r + (\zeta - 1) \nabla \cdot v \right] \]
\[ r^2 \sigma_{\phi \phi} = \Sigma \left[ 2 v \left( \partial_\phi v_\phi + (\zeta - 1) r^2 \nabla \cdot v \right) \right], \]
\[ r^2 \sigma_{\phi r} = \Sigma \left[ 2 v \left( \partial_\phi v_r + (\zeta - 1) \nabla \cdot v \right) \right]. \]

For this conversion, one needs to contract the coordinate tensor components with the orthonormal components of the coordinate basis vectors. This simply amounts to multiplying each $\phi$-up component by $r$ and each $\phi$-down component by $1/r$ (e.g. $v^\phi = rv^\phi = v_\phi = v_\phi / r = r \Omega$).
where \( v^\phi = v_\phi = r \Omega \). Since the value of the bulk viscosity coefficient \( \zeta \) is somewhat arbitrary, we set it to 0. Then simplifying the above using \( \nabla \cdot \mathbf{v} = \partial_r v_r + \frac{1}{r} \partial_\phi v_\phi + \frac{\partial_z v_z}{r} \),

\[
\begin{align*}
\sigma_{\phi} &= \Sigma \nu \left[ \partial_r v_r - \frac{1}{r} \partial_\phi v_\phi - \frac{v_\phi}{r} \right], \\
\sigma_{\phi \phi} &= \Sigma \nu \left[ \partial_\phi v_\phi + \frac{1}{r} \partial_r v_r - \frac{v_\phi}{r} \right], \\
\sigma_{\phi \phi} &= \Sigma \nu \left[ \frac{1}{r} \partial_\phi v_\phi - \partial_r v_r + \frac{v_\phi}{r} \right].
\end{align*}
\]

(A5)

**A4 Divergence of \( \sigma \) in polar coordinates**

To compute the geometric source terms which will modify the 2D polar momentum equation we compute the divergence of the second rank tensor \( \sigma \). Starting again in coordinate bases we may write

\[
(\nabla \cdot \sigma)^i = \partial_i \sigma^{ii} + \partial_\phi \sigma^{i \phi} + \frac{1}{r} \sigma^{ir} - r \sigma^{i \phi}
\]

(A6)

giving us the components of the viscous force

\[
(\nabla \cdot \sigma)^i = \partial_i \sigma^{ii} + \partial_\phi \sigma^{i \phi} + \frac{1}{r} \sigma^{ir} - \sigma^{i \phi}
\]

\[
(\nabla \cdot \sigma)^r = \partial_r \sigma^{rr} + \frac{1}{r} \partial_\phi (r \sigma^{i \phi}) - \frac{3}{r} \sigma^{r \phi}.
\]

Transforming again to an orthonormal basis for implementation in FLASH,

\[
(\nabla \cdot \sigma)^i = \partial_i \sigma^{ii} + \partial_\phi \sigma^{i \phi} + \frac{1}{r} \sigma^{ir} - r \sigma^{i \phi}
\]

\[
\frac{1}{r} (\nabla \cdot \sigma)^\phi = \partial_\phi \left( \frac{1}{r} \sigma^{r \phi} \right) + \partial_r \left( \frac{1}{r} \sigma^{r \phi} \right) + \frac{3}{r} \sigma^{r \phi}.
\]

(A7)

Simplifying,

\[
(\nabla \cdot \sigma)^i = \frac{1}{r} \partial_i \left( r \sigma^{ii} \right) + \frac{1}{r} \partial_\phi (r \sigma^{i \phi}) - \frac{3}{r} \sigma^{r \phi}
\]

\[
(\nabla \cdot \sigma)^r = \frac{1}{r} \partial_r (r \sigma^{rr}) + \frac{1}{r} \partial_\phi (r \sigma^{r \phi}) + \frac{3}{r} \sigma^{r \phi}.
\]

Plugging in the values of the components from (A5) we have

\[
(\nabla \cdot \sigma)^i = \frac{1}{r} \partial_i \left[ r \Sigma \nu \left( \partial_r v_r - \frac{1}{r} \partial_\phi v_\phi - \frac{v_\phi}{r} \right) \right] \\
+ \frac{1}{r} \partial_\phi \left[ \Sigma \nu \left( \partial_r v_r + \frac{1}{r} \partial_\phi v_\phi - \frac{v_\phi}{r} \right) \right] \\
- \frac{1}{r} \partial_\phi \left( \partial_r v_r - \partial_\phi v_\phi + \frac{v_\phi}{r} \right) \\
\]

\[
(\nabla \cdot \sigma)^r = \frac{1}{r} \partial_r \left[ r \Sigma \nu \left( \partial_r v_r + \frac{1}{r} \partial_\phi v_\phi - \frac{v_\phi}{r} \right) \right] \\
+ \frac{1}{r} \partial_\phi \left[ \Sigma \nu \left( \partial_r v_r - \partial_\phi v_\phi + \frac{v_\phi}{r} \right) \right] \\
+ \frac{1}{r} \partial_\phi \left( \partial_r v_r + \frac{1}{r} \partial_\phi v_\phi - \frac{v_\phi}{r} \right) \\
\]

(A8)

The first two terms in each of the above look like a normal divergence of a rank-one-tensor and the third terms are the geometric source terms that we must add to the momentum equation. They are akin to the centrifugal and Coriolis terms which arise from a similar exercise performed on the \( \rho \mathbf{v} \) term of (A3). Note that each term has the units of force per volume while the components of the stress tensor have units of density times velocity squared which matches the \( \rho \mathbf{v} \) term in the hydroequations.

**A5 Implementation in FLASH**

In FLASH 3.2, the components (with respect to the orthonormal basis) of the viscous stress tensor are computed in the routine Diffuse_visc.F90. This routine then subtracts the flux indicated by the viscous stress tensor from the \( \Sigma \nu \) flux. Note that since FLASH computes fluxes on cell boundaries, the stress tensor components in Diffuse_visc.F90 must be computed on the lower face of the current sweep direction (see e.g. Edgar 2006). We compute the viscous source terms along with the centrifugal and Coriolis source terms in the routine hy_ppm_force.F90.

We test the above implementation in FLASH with two tests, the viscously spreading ring of Pringle (1981) and a Cartesian shear flow.

**A6 Viscously spreading ring**

The viscously spreading ring test begins with a delta function initial density distribution which spreads solely due to viscous forces. This test assumes axis-symmetry \( \partial_\phi = 0 \) and exercises only the terms with \( \sigma^{i \phi} \) in (A8). The analytic solution assumes a constant coefficient of kinematic viscosity is well known and given by Pringle (1981).

To implement the viscously spreading ring test we choose \( v = \text{cst.} = \alpha / \mathcal{M}^2 \) with dimensionless parameter \( \alpha = 0.1 \), and Mach number \( \mathcal{M} = 100 \). This choice of Mach number mitigates pressure effects and also allows time for a small initial transient to pass through the simulation domain without greatly affecting the evolution of the solution (wiggles in the earliest blue dashed radial velocity curve of Fig. A1). As required by the analytic solution, we also turn off all terms in (A8) which do not include \( \sigma^{i \phi} \). The outer and inner boundaries are at \( r_{\text{min}} = 0.2 \) and \( r_{\text{max}} = 2.2 \) where boundary values are set by the time-dependent analytic values. We use a spatial resolution of 128 radial cells by 64 azimuthal cells and we start with initial conditions corresponding to the dimensionless initial time parameter \( \tau = 12 \nu t / r_0^2 = 0.032 \), where \( r_0 \) is the initial position of the delta function ring.

Fig. A1 shows the result of the test. Besides some expected deviation at the inner boundary, the numerical solution (dashed lines) agrees well with the analytic solution (solid lines).

**A7 Cartesian shear flow**

The viscously spreading ring test confirms that the most important terms for thin disc accretion (those containing \( \sigma^{i \phi} \)) are implemented properly. The Cartesian shear flow tests all of the terms in (A8). The idea is to choose a problem which is analytic in Cartesian coordinates and use the computer to solve it in polar coordinates thereby exercising all of the polar derivatives to make up the simple Cartesian derivatives.

To set-up the Cartesian shear flow problem, we start with the Navier–Stokes equation for an incompressible fluid with constant
coefficient of viscosity $\nu$,
\[ \partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v} = -\frac{1}{\Sigma} \nabla P + \nu \nabla^2 \mathbf{v}. \] (A9)

Choosing constant pressure and $\mathbf{v} = \mathbf{v}(y, t) \mathbf{e}_y$ reduces (A9) to a simple 1D diffusion equation in $v_y(y, t)$:
\[ \partial_t v_y(y, t) = \nu \partial^2_y v_y(y, t). \] (A10)

A solution is
\[ v_y(y, t) = \frac{v_0}{\sqrt{2\pi \nu t}} \exp \left[ -\frac{(y - y_0)^2}{4\nu t} \right], \] (A11)

while $v_x$ is 0 for all time.

In practice, we implement the Cartesian shear flow test with initial conditions,
\[ \Sigma(r, \phi, t_0) = 1.0 \]
\[ v_i(r, \phi, t_0) = v_i(y, t_0) \cos \phi \]
\[ v_\phi(r, \phi, t_0) = -v_x(y, t_0) \sin \phi, \] (A12)

with $y_0 = 1.0$, $v_0 = 1.0$, $v = 0.1$ and $t_0 = 0.5$. Choice of outer and inner boundaries of $r_{\text{min}} = 0.5$, $r_{\text{min}} = 5.0$ allow the solution to not be greatly affected by the boundaries while supplying reasonable resolution requirements. Fig. A2 plots the results of this test set up in FLASH for a number of different resolutions and cell aspect ratios.

This paper has been typeset from a \TeX/\LaTeX\ file prepared by the author.