

The dynamics and electromagnetic signatures of accretion in unequal-mass binary black hole inspirals

Madeline Clyburn   and Jonathan Zrake  

Department of Physics and Astronomy, Clemson University, 105 Sikes Hall, Clemson, SC 29634, USA

Accepted 2025 April 7. Received 2025 April 7; in original form 2024 May 15

ABSTRACT

We present a theoretical study of the gravitational wave (GW) driven inspirals of accreting black hole binaries with mass $M = 10^7 M_\odot$ and mass ratios between 10^{-3} and 10^{-1} . Our results are based on analytic estimates, and grid-based hydrodynamics simulations run for many thousands of binary orbits before the merger. We show that the GW inspiral is evident in the light curves and colour evolution of a binary-hosting quasar over years to decades before a merger. The long-term electromagnetic (EM) signature is characterized by a gradual UV brightening and X-ray dimming, followed by an X-ray disappearance hours to days before the GW burst, and finally, a years-like re-brightening as the disc relaxes and refuels the remnant black hole. These time-scales are surprisingly insensitive to the normalization of the kinematic viscosity in the disc. The spectrum of quasi-thermal disc emission shows two peaks: one in the UV and another in the X-ray, associated with the outer and circum-secondary discs, respectively; emission from the inner disc is suppressed because the secondary consumes most of the inflowing gas. We discuss implications for real-time and archival EM follow-up of GW bursts detected by *LISA*.

Key words: accretion discs – black hole physics – hydrodynamics – black hole mergers.

1 INTRODUCTION

The gravitational wave (GW) driven inspirals of massive black hole binaries (MBHBs) are high-priority targets for the future space-based GW detectors such as *LISA*, *Taiji*, and *TianQin* (Ruan et al. 2018; Gong, Luo & Wang 2021; Amaro-Seoane et al. 2023). Within the next 15 yr, *LISA* will observe the GWs emitted by MBHBs in the final hours to months of their lives (e.g. Sesana 2021), including the GW burst produced as they merge, and the ring-down of the final black hole remnant (Baibhav, Berti & Cardoso 2020; Pitte et al. 2023). These observations will directly constrain the binary component masses and spins (e.g. Trias & Sintes 2008), and could reveal the presence of a circumbinary gas disc (Yunes et al. 2011; Derdzinski et al. 2019, 2021; Garg et al. 2024).

The science potential of a *LISA* detection is greatly increased if the host galaxy can be identified. The host morphology could reveal evidence of recent galaxy mergers (e.g. Hughes 2002; Izquierdo-Villalba et al. 2022), potentially constraining the time-scale for MBHB formation (Begelman, Blandford & Rees 1980). Also, the host redshift, combined with the GW luminosity distance, yields a measurement of the Hubble constant (Schutz 1986). However, *LISA* observations of MBHB inspirals could be localized to a region of $1\text{--}10 \text{ deg}^2$, generally too large to identify the host galaxies of individual events prior to merger (Mangiagli et al. 2020). It should be noted that post-merger, the *LISA* localization region can be potentially as small as 0.01 deg^2 . Nevertheless, host galaxy identification in the pre-merger phase will almost certainly require an electromagnetic (EM)

counterpart, and pre-merger galaxy selection can uniquely allow for multiwavelength monitoring of the source during and after the merger.

The MBHBs that will merge as *LISA* sources likely exist in the gas-rich nuclear regions of post-merger galaxies (Barnes & Hernquist 1996) and could thus be surrounded by an optically thick accretion flow. The flow would consist of an outer (circumbinary) disc, an inner disc around the primary (more massive) black hole, and a smaller ‘minidisc’ around the secondary (less massive) black hole. The quasi-thermal radiation released from the accretion flow could exhibit distinctive temporal or spectral characteristics, which may aid in the identification of a host galaxy, either in archival data from time-domain EM surveys or with contemporaneous tiling of the *LISA* error box.

Predicting such EM signatures requires a detailed understanding of how the accretion flow evolves in time throughout the binary inspiral, and many studies have now been carried out based on magnetohydrodynamics simulations in both 2D (e.g. Noble et al. 2012; Farris et al. 2015; Tang, Haiman & MacFadyen 2018; Derdzinski et al. 2019, 2021; Dittmann, Ryan & Miller 2023; Krauth et al. 2023b) and 3D (e.g. Farris, Liu & Shapiro 2010, 2011; Farris et al. 2012; Bowen et al. 2018, 2019; Pereira et al. 2019; Combi et al. 2021; Gutiérrez et al. 2022; Ruiz, Tsokaros & Shapiro 2023; Avara et al. 2024; Cocchiararo et al. 2024; Franchini et al. 2024). Each of these is based on equal-mass binary inspirals, except for Ruiz et al. (2023), which includes a model with a mass ratio as low as $q \equiv M_2/M_1 = 0.25$, and Pereira et al. (2019) and Derdzinski et al. (2019, 2021), which include intermediate mass ratio inspirals (IMRIs; Arca Sedda, Amaro Seoane & Chen 2021), in the range of $q = 10^{-2}$ to 10^{-4} . The study led by Pereira reports the time-evolving

* E-mail: clyburn@clemson.edu (MC); jzrake@clemson.edu (JZ)

mass accretion rates, but not EM signatures, and the studies led by Derdzinski analyse the magnitude of gas-induced deviations from the vacuum post-Newtonian inspiral, but also do not characterize the EM signatures. Very recently, Cocchiararo et al. (2024) reported EM signatures for mass ratios as small as $q = 0.1$.

In this paper, we use multidimensional hydrodynamics simulations to characterize the EM signatures of unequal-mass inspirals, with q in the range of 10^{-3} to 10^{-1} . Our interest in unequal-mass inspirals is motivated by the possibility that a large fraction of the MBHB population could have significantly unequal masses (Sesana et al. 2012; Bellovary et al. 2019), and that accretion in low mass ratio binaries seems to operate with reduced stochastic variability, especially below $q \sim 0.01$ – 0.04 (e.g. Farris et al. 2014; Shi & Krolik 2015; D’Orazio et al. 2016; Duffell et al. 2020; Dittmann & Ryan 2024), which could allow for periodic or secular variability signatures to stand out more clearly. IMRIs are also of particular interest because the lower mass component black holes will be in the range of $10^{3.5} M_{\odot}$, meaning that *LISA* observations of them would directly probe a population of intermediate-mass black holes, which has so far been elusive (e.g. Miller 2009).

Accreting IMRIs might also present distinctive spectral signatures to distinguish them from MBHBs with nearly equal component masses. This is because the less massive black hole is generally expected to consume a majority of the inflowing gas as seen in numerical simulations with thin discs (e.g. Lubow, Seibert & Artymowicz 1999; Duffell et al. 2020; Muñoz et al. 2020; Siwek et al. 2023a), an effect referred to in literature as preferential accretion (Muñoz & Lai 2016). The effect implies that emission from the secondary minidisks could dominate the system’s overall EM luminosity or that the spectral energy distribution of quasi-thermal disc emission could exhibit multiple peaks. Also, if the degree of preferential accretion changes throughout the late inspiral, that might imprint a distinctive colour evolution that could be used to photometrically select host galaxy candidates from time-domain surveys. Thus, the main goal of this study is to measure the evolution of the component accretion rates \dot{M}_1 and \dot{M}_2 throughout a GW-driven inspiral, and use those measurements to approximate the colour evolution of accreting IMRIs.

To the best of our knowledge, this study is the first to use multidimensional hydrodynamics simulations to model the EM emission of accreting IMRIs. However, the earliest papers on EM emission from MBHB inspirals, which were based on one-dimensional disc models, focused specifically on systems with mass ratios $q \lesssim 0.1$. Armitage & Natarajan (2002) used the ‘standard disc’ equations (e.g. Shakura & Sunyaev 1973; Pringle 1981) to model how the inner and outer discs evolve throughout the inspiral. They used an empirical term derived from the theory of disc–satellite interactions (Goldreich & Tremaine 1980; Lin & Papaloizou 1986) to model the gravitational torque applied to the disc by the secondary. Such treatments can lead to the formation of a vacuum gap surrounding the secondary’s orbit, prohibiting mass from moving between the inner and outer discs or on to the secondary black hole. As the GW inspiral speeds up, the vacuum gap implies ‘tidal squeezing’ of the inner disc towards the primary black hole. Armitage & Natarajan (2002) speculated that this process (now sometimes referred to as a ‘snowplow’ mechanism) force-feeds the larger black hole and could drive a massive outflow and EM flare in the very late inspiral. The studies of Lodato et al. (2009) and Chang et al. (2010) are similarly based on time-dependent one-dimensional disc models and also support the possibility of a snowplow mechanism.

However, the vacuum gap around the secondary’s orbit is now understood to be an artefact of one-dimensional disc treatments.

Two-dimensional calculations show that gas crosses the secondary orbit on so-called ‘horseshoe’ orbits (Lubow et al. 1999; Duffell et al. 2014; Fung, Shi & Chiang 2014). It means that the inner disc gas could be expelled to the outer disc or consumed by the secondary and that a late-inspiral enhancement of the accretion rate may not take place (Baruteau, Ramirez-Ruiz & Masset 2012). In fact, the opposite effect has been seen in simulations of equal-mass binaries (e.g. Farris et al. 2015; Dittmann et al. 2023; Krauth et al. 2023b; Franchini et al. 2024); the binary accretion rate drops throughout the late inspiral. We refer to such mass starvation as ‘viscous decoupling’, an effect first described in Liu, Wu & Cao (2003). It happens because the binary eventually contracts faster than the inward viscous spreading of the outer disc, and this leads to a diminishing rate of mass supply to the binary components. It was also pointed out in Milosavljević & Phinney (2005) that viscous decoupling of the binary should lead to a post-merger re-brightening as the disc spreads inwards following the merger and eventually begins fuelling the black hole remnant. The refuelling has been simulated for the case of post-merger equal-mass binaries in Farris et al. (2015) and Krauth et al. (2023b), and we examine it in this study for the case of intermediate mass ratios $q < 0.1$.

A cautionary note for readers of this paper and similar works: previous works have not always made the distinction between two different time-scales in the binary–disc interaction. One time-scale, the viscous decoupling time-scale, as mentioned in the preceding paragraph, describes the time at which the binary merger rate begins to contract faster than the viscous spreading at the inner edge of the outer disc. Previous works (e.g. Sesana et al. 2012) have referred to this time-scale as the ‘disc freezing time-scale’. The second time-scale is the time at which the binary switches from contraction/expansion due to the exchange of angular momentum with the gas disc to contraction due to GW emission. In this paper, we will refer to this time-scale as the ‘transition time-scale’, but others have referred to this time-scale as a ‘decoupling time-scale’ (e.g. Armitage & Natarajan 2002; Sesana et al. 2012). These time-scales are distinct and should be calculated separately; see Sections 2.3 and 2.5 for derivations of the viscous decoupling time-scale and transition time-scale, respectively.

Our study is based on time-dependent numerical solutions of the two-dimensional (vertically averaged) Navier–Stokes equations, where gas is subject to the time-dependent gravitational potential of an unequal-mass black hole binary, undergoing a GW-driven post-Newtonian inspiral (Peters 1964). We adopt a simplified thermodynamic treatment, in which the speed of sound is set to a small fraction, $1/\mathcal{M}$, of the nominal orbital speed (\mathcal{M} is a global Mach number, selected to be in the range of 10–20). The bolometric EM luminosity is reasonably well predicted by measurements of the mass accretion rate, provided the flow is radiatively efficient. Owing to the use of simplified thermodynamic treatment, our simulations do not yield a self-consistent disc photosphere temperature. To model the colour evolution, we have thus developed a ‘three-disc’ toy model for the disc surface temperature, based on approximating the inner, outer, and secondary discs as stationary α -discs (Shakura & Sunyaev 1973), with simulation-calibrated mass inflow rates. We believe this approach yields a reasonable first approximation of the colour evolution of accreting IMRIs. Models with self-consistent thermodynamics (e.g. Westernacher-Schneider et al. 2022, 2024) for unequal-mass binaries will be reported in a follow-up study.

Our paper is organized as follows. In Section 2, we describe the physical picture of the MBHB inspiral and provide analytical estimates for the EM emission from these systems. In Section 4.7, we describe the equations of motion, initial conditions, and the

numerical set-up. Section 6.1 contains the key dynamical measurements obtained from the numerical simulations. Section 5 includes the predictions of spectral evolution and light curves based on the three-disc spectral toy model and discusses prospects for EM detections of *LISA* sources. Finally, in Section 6, we summarize our results, discuss some caveats, and describe plans for future work.

2 ACCRETING UNEQUAL-MASS BLACK HOLE BINARIES

2.1 Fiducial system

We adopt a fiducial black hole binary that will be detectable by *LISA* out to a high redshift. The total mass of the fiducial system is $M = 10^7 M_\odot$, the mass ratio is $q = 0.01$, and the orbit is circular. The assumption of a circular orbit is justified on the basis that in the late-inspiral stage, the GW radiative back-reaction exceeds the gravitational forces of the gas, so the orbital eccentricity, driven up at wider binary separations (D’Orazio & Duffell 2021; Zrake et al. 2021; Siwek, Weinberger & Hernquist 2023b), has already been dissipated. At the leading post-Newtonian order, the rate of orbital contraction due to GW’s is given in Peters (1964) as

$$\dot{a}_{\text{gw}} = -\frac{64G^3M^3\eta}{5c^5a^3}, \quad (1)$$

where $\eta \equiv q/(1+q)^2 = M_1M_2/M^2$ and a is the size of the orbit. If the fiducial binary were on track to merge in the next 50 yr, then it would presently have a semimajor axis $a_0 \simeq 6.3$ au and an orbital period $P_0 \simeq 1.8$ d. The time-to-coalescence is denoted as $t_{\text{minus}} \equiv \tau/4$, where $\tau \equiv -a/\dot{a}_{\text{gw}}$ is the nominal orbital contraction timescale. The fiducial system is very mildly relativistic, with the orbital velocity roughly equivalent to $0.17c$ and gravitational radius $r_g \equiv GM/c^2$ about 64 times smaller than the initial semimajor axis. For later reference, the three (small) relativistic parameters v/c , r_g/a , and P/τ are mutually related via equation (1),

$$\left(\frac{v}{c}\right)^2 = \frac{r_g}{a} = \left(\frac{5P}{128\pi\eta\tau}\right)^{2/5}. \quad (2)$$

The Eddington luminosity of a $10^7 M_\odot$ black hole is $L_{\text{edd}} \simeq 1.15 \times 10^{45} \text{ erg s}^{-1}$, corresponding to a bolometric absolute magnitude of $M_{\text{bol}} \simeq -24$. For reference, such modest-luminosity active galactic nuclei (AGNs) will be detectable by *LSST* out to redshifts of $z \simeq 5.5$. (Brandt et al. 2009). The *LISA* detection horizon for these events will be at redshifts of $z \simeq 20$ (Sesana 2021).

2.2 Physical picture and spectral appearance

The binary is surrounded by a thin, centrifugally supported, optically thick gas flow. In binaries with a small mass ratio, the overall flow is similar to that of planet-forming discs. It comprises an outer circumbinary disc, a low-density annular gap surrounding the secondary’s orbit, and an inner disc around the primary.

The accretion flow appears as a multitemperature blackbody; the spectrum is determined by the distribution of the local disc photosphere temperatures. To get a sense for the composite spectrum of the accretion flow, including the outer disc, the inner disc, and the secondary minidisc, it is instructive to consider each of these discs as a steady-state α -disc (Shakura & Sunyaev 1973), and examine the composite quasi-thermal emission spectrum. The characteristic

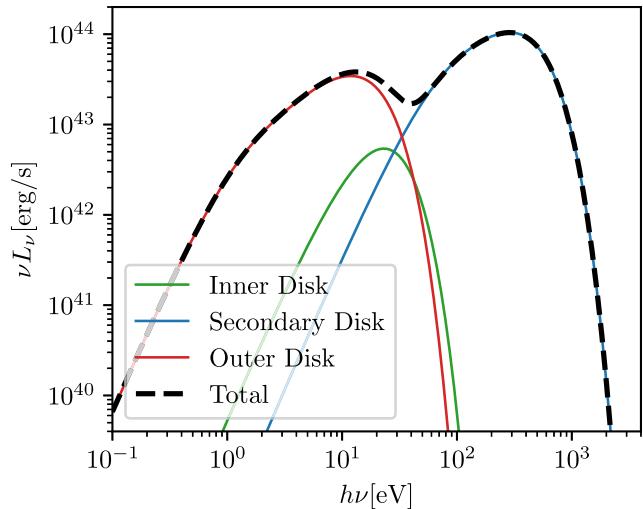


Figure 1. Toy model of the spectral energy distribution of thermal emission produced by an accretion flow around an $M = 10^7 M_\odot$ binary with mass ratio $q = 0.01$. The model assumes that the three components of the accretion flow: inner disc (green), secondary disc (blue), and outer disc (red) can each be characterized as a steady-state α -disc, with radial temperature profile consistent with equation (3).

temperatures are given by

$$T_{\text{outer}} = (3GM\dot{M}/8\pi a^3\sigma)^{1/4} \quad (3)$$

$$T_i = (3GM_i\dot{M}_i/8\pi R_i^3\sigma)^{1/4} \propto \dot{M}_i^{1/4} M_i^{-1/2},$$

where $i = 1$ for the primary and $i = 2$ for the secondary, and $R_i = 6GM_i/c^2$ are the inner cut-off radii of the component discs. For the fiducial system discussed in Section 2.1, accreting at the Eddington limit of the primary,¹ with $M_2/M_1 = 10$, the characteristic temperatures of the three discs are

$$kT_{\text{outer}} \simeq 28 \text{ eV}$$

$$kT_1 \simeq 41 \text{ eV}$$

$$kT_2 \simeq 1.3 \text{ keV}.$$

The secondary black hole has a more compact disc and typically accretes faster, so the secondary disc emission is bluer and brighter than the primary and outer discs.

Fig. 1 shows a toy model spectrum based on this picture and the temperature profiles in equation (3). The inner disc extends to $r = a$, and the secondary disc extends to the Hill radius $r_{\text{hill}} \simeq a(q/3)^{1/3}$. The outer disc starts at $r = a$ and extends to infinity. The preferential accretion parameter is set to $\dot{M}_2/\dot{M}_1 = 10$. The spectrum is double-peaked, with the UV and X-ray bumps formed by the outer and secondary discs, respectively. The inner disc is outshined by the other components, a consequence of the secondary black hole dominating the accretion. Note that a double-peaked spectrum similar to that

¹If the circumbinary disc carries mass inwards at a large fraction of the primary’s Eddington limit, the secondary would need to accrete in a super-Eddington regime to consume a significant fraction of the overall mass flow. Thus, our assumption that the secondary disc remains thin is not technically compatible with assumptions elsewhere that the binary accretion rate is comparable to the Eddington rate. This very interesting aspect of accretion in unequal-mass binary systems needs to be addressed but is out of the scope of this study. Most AGNs accrete below Eddington, so we may be overestimating typical luminosities here while not violating the physical assumption of a thin secondary disc.

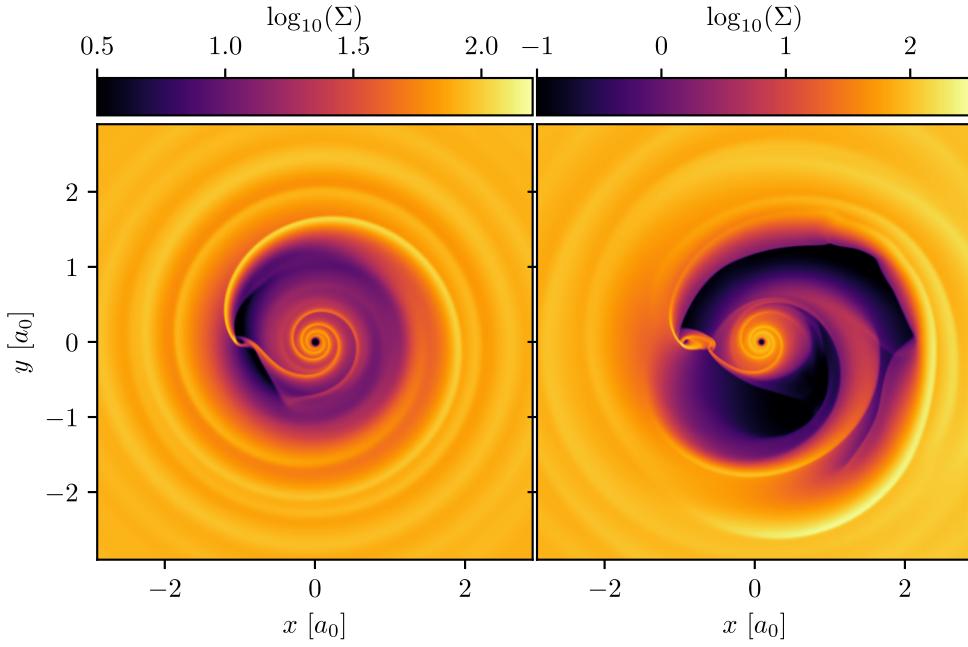


Figure 2. Image of the surface density of the disc around an MBHB with mass ratio $q = 0.01$ (left) and $q = 0.1$ (right), accreting from a circumbinary gas disc. The image is from a simulation snapshot taken well after the disc is viscously relaxed but before the binary separation has been reduced much from a_0 . The axes show the $x - y$ plane in units of a_0 .

shown in Fig. 1 is also presented in Chang et al. (2010), but the reason is different. In that study, the secondary does not accrete, and the two spectral bumps are produced by the inner and outer discs.

2.3 Viscous decoupling time-scale

At some point during the GW inspiral, the GW contraction rate τ^{-1} exceeds the viscous relaxation rate (e.g. Liu et al. 2003; Milosavljević & Phinney 2005). After that time, the disc is said to have ‘viscously decoupled’ from the binary, as it can no longer spread viscously inwards as fast as the speed \dot{a}_{gw} of the binary contraction. The secular evolution of the brightness and colour of the accretion flow around the binary could be significantly altered by viscous decoupling. Note that previous works (e.g. Sesana et al. 2012) have referred to this time-scale as the ‘disc freezing time-scale’. Also, the term ‘decoupling time-scale’ has been used in other works to refer to the time after which the GW power exceeds the rate of energy loss due to tidal interactions with the disc. In this paper, we refer to the transition from gas to GW driving as the ‘transition time-scale’. The transition time-scale is estimated later in Section 2.5.

The nominal viscous relaxation time-scale of the accretion disc is given by

$$t_{\text{visc}}(r) = \frac{2}{3} \frac{r^2}{\nu(r)}, \quad (4)$$

where $\nu(r)$ is the kinematic viscosity at radius r . In a standard α -disc with constant aspect ratio h/r , the viscosity coefficient can be written as $\nu(r) = \bar{\nu} \sqrt{GMr}$. Here, $\bar{\nu} \equiv \alpha/\mathcal{M}^2$, where $\alpha \lesssim 0.1$ represents the Reynolds stress associated with small-scale (numerically unresolved) turbulence in the disc (Shakura & Sunyaev 1973), and \mathcal{M} is the orbital Mach number. In addition to the fiducial binary parameters introduced in Section 2.1, we also define a set of fiducial disc parameters $\alpha = 0.1$ and $\mathcal{M} = 10$.

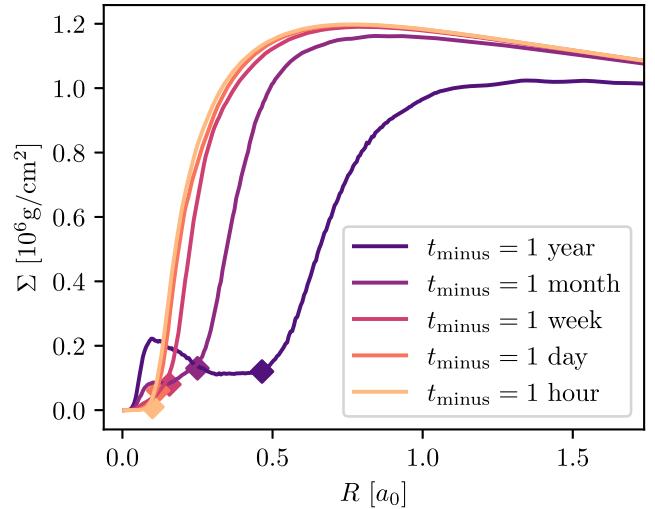


Figure 3. The radial profile of the surface density $\Sigma(r)$ in a binary with mass ratio $q = 0.01$ and orbital Mach number $\mathcal{M} = 10$, shown at representative times before the merger. The radial profiles are obtained from the azimuthal average of the two-dimensional disc surface density. The diamond markers show the location of the secondary at the representative times.

The ratio of the viscous time near the binary to the GW inspiral time is

$$\frac{t_{\text{visc}}(a)}{\tau} = \frac{128G^3M^3\eta}{15\nu(a)c^5a^2} \propto a^{-5/2}. \quad (5)$$

Note that we have chosen to equate the inspiral time-scale to the viscous time-scale at $r = a$ for the viscous decoupling condition. This choice seems more appropriate when the mass ratio is small (see Figs 2 and 3) rather than at $r = 2.5a$, which is typically used as the radius for the outer disc cavity edge in equal-mass binary studies. This ratio, as shown in equation (5), starts off small and increases

as the inspiral proceeds. It exceeds unity at the viscous decoupling radius a_{dec} ,

$$a_{\text{dec}} = r_g \left(\frac{128\eta}{15\bar{v}} \right)^{2/5} \simeq 0.6 \text{ au} \times \left(\frac{M}{10^7 M_\odot} \right) \left(\frac{\eta}{10^{-2}} \right)^{2/5} \left(\frac{\bar{v}}{10^{-3}} \right)^{-2/5}, \quad (6)$$

which we have written here, first in terms of r_g , and then for a typical *LISA* binary with the fiducial parameters described in Section 2.1. The method we have used to determine the decoupling separation in equation (6) has been questioned in works such as Dittmann et al. (2023), where the authors show that a more accurate method to derive the decoupling separation is to equate the binary contraction speed to the viscous drift speed rather than equating time-scales. Nevertheless, our result in equation (6) is the same regardless of equating time-scales or speeds. This is due to our assumption that $t_{\text{visc}} = \tau_{\text{gw}}$, while Dittmann et al. (2023) assume $t_{\text{visc}} = \tau_{\text{gw}}/4$. Thus, the viscous decoupling scales we derived here are equivalent to those in equations (5) and (6) of Dittmann et al. (2023), for their case of $\xi = 1$ and $\lambda = 2$.

The residence time for the fiducial system at the viscous decoupling stage is

$$\tau_{\text{dec}} = \frac{5}{64} \frac{r_g}{c\eta} \left(\frac{128\eta}{15\bar{v}} \right)^{8/5} \simeq 5.4 \text{ d} \times \left(\frac{M}{10^7 M_\odot} \right) \left(\frac{\eta}{10^{-2}} \right)^{3/5} \left(\frac{\bar{v}}{10^{-3}} \right)^{-8/5}, \quad (7)$$

which corresponds to a time-to-coalescence of $t_{\text{minus}} \simeq 1.4$ d. The orbital period at the viscous decoupling radius is

$$P_{\text{dec}} = 1.2 \text{ h} \times \left(\frac{M}{10^7 M_\odot} \right) \left(\frac{\eta}{10^{-2}} \right)^{3/5} \left(\frac{\bar{v}}{10^{-3}} \right)^{-3/5}. \quad (8)$$

Equation (7) shows the time-scale at which the binary is generally expected to become starved of mass supply. Note the apparent high sensitivity to the viscosity suggests that thinner α -discs might viscously decouple much earlier than warmer discs; if \bar{v} were instead 10^{-5} ($M = 100$), equation (7) indicates the viscous decoupling time-scale would then be of the order of years rather than days. In Section 4.2, we show that this expectation is not supported by our simulations; we find instead that viscous decoupling extends over thousands of τ_{dec} , implying that colder and warmer discs alike could show signs of the inspiral long before the nominal decoupling time.

Measurements of the preferential accretion parameter \dot{M}_2/\dot{M}_1 are reported in a number of studies (e.g. Lubow et al. 1999; Young & Clarke 2015; Duffell et al. 2020; Muñoz et al. 2020; Siwek et al. 2023a), and these generally find values of \dot{M}_2/\dot{M}_1 in the range of 1–10, when the mass ratio is larger than roughly 10^{-3} . All of those studies were based on simulations with a viscously relaxed disc. In Section 4.2, we present calculations revealing how the preferential accretion parameter evolves throughout the long-term viscous decoupling.

2.4 An argument against tidal squeezing

As the disc loses its viscous coupling to the binary, the feeding rates \dot{M}_1 and \dot{M}_2 evolve away from their steady-state values. These long-term trends will largely determine the IMRI brightness and colour evolution before coalescence. However, the feeding rates are not easy to predict without detailed simulations. The main question

here is whether the GW burst is preceded by a period of AGN brightening or dimming. Armitage & Natarajan (2002) computed the evolution of a one-dimensional Keplerian disc with an embedded low-mass secondary black hole ($q = 0.01$) on a GW-driven inspiral. They accounted for the gravitational torque of the secondary on the inner and outer discs, using a prescription for Lindblad torques due to Lin & Papaloizou (1986). That prescription leads to the formation of an evacuated gap around the secondary orbit, which fully insulates the mass flow between the inner and outer discs and prohibits any movement of gas on to the secondary black hole. The result, which can be seen, for example, in fig. 3 in Armitage & Natarajan (2002), is that the inner disc is squeezed towards the primary as the secondary spirals in.

Tidal squeezing would require the secondary to exert sufficient gravitational torque to the inner disc to lower the orbits of gas parcels faster than \dot{a}_{gw} . It implies the positive torque, \dot{J}_{in} , exerted by the inner disc on the secondary, must exceed

$$\dot{J}_{\text{plow}} \equiv \pi a^2 \Sigma(a) \times \Omega(a) a^2 / \tau_{\text{gw}} \simeq \frac{64\eta \dot{M} r_g^3 c}{15\bar{v} a^2}, \quad (9)$$

where the second equality is precise in a steady state with $\dot{M} = 3\pi v(r)\Sigma(r)$. The dependence of this torque on a^{-2} indicates that to squeeze the inner disc would require the torque to diverge leading up to merger. The gravitational torque coming from the inner disc can be written as $\dot{J}_{\text{in}} = \ell_{\text{in}} J_2 \dot{M} / M$, where $J_2 \simeq M_2 \Omega a^2$ is the secondary angular momentum, and ℓ_{in} is a dimensionless ‘eigenvalue’, which is typically not greater in magnitude than about 10, which is confirmed for binary mass ratio as low as $q = 0.01$ (e.g. Duffell et al. 2020). It is then straightforward to see that \dot{J}_{plow} (the torque required to squeeze the inner disc) exceeds \dot{J}_2 some $1/(6\pi\eta\bar{v}\ell_{\text{in}})$ orbits before merger. In the fiducial binary and disc system with $q \approx \eta \simeq 0.01$, $\bar{v} = 10^{-3}$, and torque eigenvalue of order unity, this corresponds to ~ 5000 orbits or ~ 25 yr before merger. Thus, on dynamical grounds, we expect that in the final thousands of orbits, the inner disc gas is either funnelled to the outer disc or falls on to the secondary. It would require an unrealistically large gravitational torque to confine the inner disc inside the secondary’s orbit in the late inspiral. In Section 4.6, we present numerical calculations confirming these estimates.

2.5 Transition from gas to GW driving

For much of the binary lifetime, the orbital contraction is thought to be governed by coupling to the circumbinary disc (e.g. Begelman et al. 1980). As the binary contracts and the GW power increases, there is a transition from gas to GW driving (Escala et al. 2005; Cuadra et al. 2009; Lodato et al. 2009; Muñoz et al. 2020; Bortolas et al. 2021). Once the binary reaches that separation, which we denote here as a_{trans} , an inspiral towards merger is inevitable as GW’s remove orbital energy from the binary at a runaway rate. Note that some authors use different terminology for this separation and time-scale, for example, Sesana et al. (2012) simply calls this ‘decoupling’. We adopted the term ‘transition separation’ for the switch to GW driving, in order to reserve the term decoupling for the process of viscous decoupling.

The transition separation occurs where the rate of orbital contraction by GW radiation (equation 1) is equivalent to the rate of orbital contraction by gas i.e. $\dot{a}_{\text{gas}} = \dot{a}_{\text{gw}}$. Gas-induced changes of the semimajor axis are prescribed here in terms of another dimensionless constant ϵ , as $\dot{a}_{\text{gas}} = -\epsilon a \dot{M} / M$. Equating \dot{a}_{gw} and \dot{a}_{gas} , we find the

separation at which GW's begin to dominate the inspiral,

$$a_{\text{trans}} = r_g \left(\frac{64\eta c^3}{5G\dot{M}\epsilon} \right)^{1/4} \simeq 94 \text{ au} \times \left(\frac{M}{10^7 M_\odot} \right)^{3/4} \left(\frac{\eta}{10^{-2}} \right)^{1/4} \left(\frac{\epsilon}{5.0} \right)^{-1/4} \dot{m}^{-1/4}, \quad (10)$$

where $\dot{m} \equiv \dot{M}/\dot{M}_{\text{edd}}$ is the mass accretion rate normalized by the Eddington accretion rate and $\epsilon \approx 5.0$ is consistent with results from simulations (e.g. Tiede et al. 2020; Dittmann & Ryan 2024; Tiede et al. 2024a). The time-scale for the transition from gas to GW-driven inspiral is $\tau_{\text{trans}} \equiv \tau(a_{\text{trans}}) = t_{\text{sal}}/\epsilon\dot{m}$, where $t_{\text{sal}} = 5 \times 10^7 \text{ yr}$ is the Salpeter time. This nominal time-scale for the end of the gas-driven evolutionary phase holds when the orbital decay follows $\dot{a}/a \sim -\dot{M}/M$ and the mass accretion rate is a fraction of the Eddington accretion rate.

Therefore, *LISA* inspirals will be generally deep in the GW-driven regime, and it justifies the choice of circular binary orbits in this study. Even though a detectable level of orbital eccentricity could be left over from a gas-driven evolutionary phase in *LISA* sources (Armitage & Natarajan 2005; Roedig et al. 2011; Zrake et al. 2021; Siwek et al. 2023b; Garg et al. 2024; Valli et al. 2024), from the point of view of the circumbinary disc interaction years before a merger, the orbit can be considered as nearly circular.

2.6 Post-merger viscous closing

Viscous decoupling implies the binary merges in a relatively low-density ‘hole’ in the disc. Here, we estimate the time-scale for the hole to be refilled following the merger, noting that such effect has been surmised to lead to a late-time brightening, or ‘merger afterglow’ (e.g. Milosavljević & Phinney 2005). The afterglow is predicted to begin a time span t_{refill} following the merger, given by the viscous relaxation time-scale at the radius where the binary became viscously decoupled,

$$t_{\text{refill}} \equiv \frac{3}{4} t_{\text{visc}}(a_{\text{dec}}) = \frac{1}{2} \frac{r_g}{c\bar{v}} \left(\frac{128\eta}{15\bar{v}} \right)^{3/5} \simeq 4.1 \text{ d} \times \left(\frac{M}{10^7 M_\odot} \right) \left(\frac{\eta}{10^{-2}} \right)^{3/5} \left(\frac{\bar{v}}{10^{-3}} \right)^{-8/5}. \quad (11)$$

This estimate is analogous to one given in Milosavljević & Phinney (2005). Note that it suggests the remnant black hole begins refuelling on a time-scale that is quite sensitive to the disc viscosity; in an α -viscosity model, the refuelling time-scale would appear to increase from days for $\mathcal{M} = 10$ to decades for $\mathcal{M} = 100$. We show in Section 4.3 that such sensitivity is not supported by the simulations, instead the remnant begins fuelling over year-like time-scales, with surprisingly little sensitivity to the normalization of the kinematic viscosity.

2.7 Importance of self-gravity

Our simulations do not include the self-gravity of the disc gas, so it is important to check when that assumption could be valid. The importance of self-gravity can be estimated from the Toomre parameter, $Q = c_s \Omega / \pi G \Sigma$ for a Keplerian disc; discs are generally unstable to fragmentation instabilities if $Q \lesssim 1$. When $\alpha = 0.1$ and $\mathcal{M} = 10$ the Toomre parameter is found to be

$$Q = 6 \times 10^5 \left(\frac{M}{10^7 M_\odot} \right)^{1/2} \left(\frac{\bar{v}}{10^{-3}} \right) \left(\frac{r}{10a} \right)^{-3/2} \dot{m}. \quad (12)$$

Colder discs are more unstable to gravitational fragmentation; however equation (12) indicates that even when the Mach number is of the order of hundreds, the disc still has $Q \gg 1$ out to $r \sim 10a$. The effect of gravitational instabilities on long-term binary orbital evolution has been examined in, e.g. Bortolas et al. (2021) and Franchini, Sesana & Dotti (2021). Although Franchini et al. (2021) only simulated equal-mass binaries and Bortolas et al. (2021) only studied $q \gtrsim 0.1$, self-gravity is not likely to influence the EM signatures of the late inspiral regardless of binary mass ratio.

3 NUMERICAL METHODS

Our hydrodynamic simulations are performed using the publicly available *SAILFISH* code. *SAILFISH* is a GPU-accelerated, grid-based hydrodynamics code with physics and post-processing features specifically targeting problems related to binary–disc interactions. *SAILFISH* uses a fixed-mesh, second-order Godunov solver and has been used in many published studies of binary accretion (e.g. Westernacher-Schneider et al. 2022; Krauth et al. 2023a,b; Duffell et al. 2024; Tiede & D’Orazio 2024; Westernacher-Schneider et al. 2024; DeLaurentiis et al. 2025).

3.1 Equations of motion

Our simulations are based on solutions to the vertically averaged, time-dependent mass continuity and Navier–Stokes equations,

$$\frac{\partial \Sigma}{\partial t} + \nabla \cdot (\Sigma \mathbf{v}) = \dot{\Sigma}_{\text{sink}}, \quad (13)$$

$$\frac{\partial \Sigma \mathbf{v}}{\partial t} + \nabla \cdot (\Sigma \mathbf{v} \mathbf{v} + P \mathbf{I} - \mathbf{T}_{\text{vis}}) = \dot{\Sigma}_{\text{sink}} \mathbf{v} + \mathbf{F}_g. \quad (14)$$

Here, Σ is the vertically integrated mass density, \mathbf{v} is the gas velocity, $P \equiv \Sigma c_s^2$ is the vertically integrated gas pressure, and \mathbf{I} is the identity tensor (pressure is isotropic). We adopt a local isothermal equation of state, $c_s^2 = -\nabla/\mathcal{M}^2$, where \mathcal{M} is the orbital Mach number, and the gravitational potential is given by

$$\nabla = -\frac{GM_1}{(r_1^2 + r_s^2)^{1/2}} - \frac{GM_2}{(r_2^2 + r_s^2)^{1/2}}. \quad (15)$$

Here, r_1 and r_2 are the distances to each black hole, and r_s is the gravitational softening length to ensure the potential is finite at the component positions. Use of the local isothermal equation of state means that we do not solve an energy transport equation nor adopt a phenomenological cooling prescription. Certain thermodynamic processes, such as shock-heating, are not accounted for as a result of the isothermal equation of state.

The ‘sink’ term $\dot{\Sigma}_{\text{sink}}$, appearing in equations (13) and (14), is responsible for the direct exchange of mass and momentum between the gas and the binary components,

$$\dot{\Sigma}_{\text{sink}} = -\frac{\Sigma}{\tau_{\text{sink}}} (e^{-r_1^2/2r_{\text{sink}}^2} + e^{-r_2^2/2r_{\text{sink}}^2}). \quad (16)$$

We have checked that our choice of sink rate, τ_{sink}^{-1} , does not significantly affect our measurements of the preferential accretion rate, provided the sink rate is sufficiently large. This seems to be different from Dittmann & Ryan (2024), who reported modest sensitivity to the sink rate. We suspect the difference could be in the use of α versus constant- ν viscosity. In this study, we have adopted a sink radius of $r_{\text{sink}} = 0.05a_0$ and a sink rate of $\tau_{\text{sink}}^{-1} = 10\Omega_0$.

Our chosen sink radius is always smaller than the Hill radius of the secondary black hole, $r_{\text{hill}} = a_0(q/3)^{1/3} \geq 0.07a_0$, at the start of the simulation, even for the lowest mass ratio considered ($q = 10^{-3}$).

However, as the binary separation decreases, the sink size of $0.05a_0$ will become too large to resolve the Hill radius for lower mass ratios. We acknowledge this as a caveat, as our results may be sensitive to sink size, particularly for the smallest mass ratios simulated. To address this issue, we distinguish our accretion rate results based on the sink radius of the secondary black hole. When the sink radius satisfies $r_{\text{sink}} \geq 0.5 r_{\text{hill}}$, we represent the results with dashed, less opaque lines to indicate that gas flows in the circum-secondary disc are not well resolved. For a discussion of the sensitivity of the accretion rates to the sink size and sink rate in our simulations, see Appendix A.

The vertically integrated gravitational force density, associated with the softened gravitational potential, is $\mathbf{F}_g = -\Sigma \nabla \nabla$. We assume the binary mass $M = M_1 + M_2$ is much larger than the disc mass so that self-gravity can be safely ignored (see Section 2.7). The components of the viscous stress tensor \mathbf{T}_{vis} in equation (14) are given by $T_{\text{vis}}^{ij} = \nu \Sigma (\partial^j v^i + \partial^i v^j - \partial_k v^k \delta^{ij})$. SAILFISH supports both ‘constant- ν ’ and ‘constant- α ’ viscosity prescriptions. For this study, we have adopted the constant- α viscosity prescription with kinematic viscosity coefficient

$$\nu = \frac{\alpha c_s^2}{\tilde{\Omega}}, \quad (17)$$

where the modified Keplerian orbital frequency $\tilde{\Omega}$ is

$$\tilde{\Omega} = \sqrt{\frac{GM_1}{r_1^3} + \frac{GM_2}{r_2^3}}. \quad (18)$$

At large radii $r \gg a$, equation (17) reduces to $\nu \simeq \bar{v} \sqrt{GMr}$, where $\bar{v} = \alpha/M^2$.

The black holes are moved according to the Kepler two-body problem, modified to account for the GW-driven inspiral up to the leading post-Newtonian order (Peters 1964). The semimajor axis as a function of time is the solution of equation (1),

$$a(t) = a_0 (1 - 4t/\tau_0)^{1/4}, \quad (19)$$

where τ_0 is the contraction time-scale at the start of the simulation. Positioning the binary components on the grid requires an expression for the orbital phase $\phi(t)$ as a function of the simulation time t . The orbital phase is obtained by integrating in time the instantaneous orbital frequency,

$$\phi(t) = \int_{a_0}^{a(t)} \sqrt{\frac{GM}{a^3} \frac{da}{\dot{a}_{\text{gw}}}} = \frac{2}{5} \tau_0 \Omega_0 \left(1 - \left(\frac{a(t)}{a_0} \right)^{5/2} \right). \quad (20)$$

The gravitational force of the gas is neglected in the binary equation of motion, which is justified because we are only considering binaries deep in the GW-driven regime.

3.2 Initial data, outer boundary condition, and diagnostics

We start our simulations of accretion on to the binary with a circular gas disc: the azimuthal velocity profile is $v_\phi = v_{\text{kep}} = \sqrt{GM/r}$, the radial velocity is inwards at the viscous drift speed, $v_r = -3v(r)/2r$, and the surface density is written as

$$\Sigma_0(r) = \frac{\dot{M}_0}{3\pi v(r)} \propto r^{-1/2}, \quad (21)$$

generally as a steady-state viscous disc in the middle and on the right-hand side approximated for an α -disc with viscosity as in equation (17) and with uniform orbital Mach number. Note that equation

(21) is the surface density of a steady disc with a zero net angular momentum current; this fact is revisited in Section 3.3. After the start of the simulation, it takes several viscous time-scales at $r \sim a_0$ for the flow to relax to a statistically steady state. For reference, when $\alpha = 0.05$ and $\mathcal{M} = 10$, the viscous time-scale at $r = a_0$ is $t_{\text{visc}} \simeq 200$ orbits. For both $q = 0.01$ and $q = 0.1$ binaries, Fig. 2 shows a snapshot of the surface density once the system has viscously relaxed to a quasi-steady state, or in other words, the surface density snapshot is taken after a thousand orbits or five viscous time-scales at $r = a_0$.

Because our code uses Cartesian coordinates with a square-shaped domain, we impose a ‘soft’ outer boundary condition within a circular buffer zone starting at radius $r_{\text{buf}} \simeq 7.5a_0$ (we have confirmed that our results are not changed by an increase of r_{buf} , see Section 3.3). Inside the buffer region, the surface density and gas velocity are driven smoothly towards a zero-torque α -disc (which is also the initial condition). The domain size of our simulations is $20a_0$, and the number of grid zones is 2400×2400 unless mentioned otherwise; the simulations considered in this paper have a grid spacing $\Delta x \simeq 0.008a_0$.

The accretion rate for each black hole is measured by separately integrating the two terms in equation (16) corresponding to the primary and secondary black hole accretion rates. Smoothing is applied to the time series of the accretion rates unless specifically mentioned. The time series of the gravitational torque is computed by integrating the gravitational torque density over the whole surface of the disc,

$$j_{\text{grav}} = \int r \times \mathbf{F}_g \, dA. \quad (22)$$

The gravitational torque is decomposed into the torque on the primary or the secondary and also into contributions from the inner and outer discs. We ignore the accreted spin angular momentum as we have calculated that the accretion torques are negligible.

3.3 A note about accretion rates in the steady state

Ideally, our accretion rates \dot{M}_1 and \dot{M}_2 would be reported relative to the theoretical inflow rate \dot{M}_0 in equation (21). However, we have observed that in our simulations, the mean binary accretion rate $\langle \dot{M} \rangle$ generally exceeds \dot{M}_0 by a significant margin. We understand this ‘overshoot’ to be an artefact of our treatment of the domain outer boundary (Section 3.2).

Near the outer boundary (or the buffer region), the surface density is driven smoothly towards $\Sigma_0(r)$ in equation (21), which corresponds to a zero-torque disc. We make this choice mainly for convenience because the torque on the binary j cannot be controlled. Rather, it needs to be measured from the simulation. When j is positive, a better approximation of the surface density far from the binary is (see e.g. Pringle 1981)

$$\Sigma(r) = \frac{1}{3\pi v(r)} \left(\langle \dot{M} \rangle - \frac{j}{\sqrt{GMr}} \right). \quad (23)$$

Thus, the target surface density $\Sigma_0(r)$ would be generally smaller than $\Sigma(r)$ if we had $\langle \dot{M} \rangle = \dot{M}_0$. However, the buffer zone forces the equality $\Sigma(r_{\text{buf}}) = \Sigma_0(r_{\text{buf}})$, so in general we have $\langle \dot{M} \rangle \gtrsim \dot{M}_0$.

The amount of this ‘overshoot’ can be estimated, noting that when the binary mass ratio is small, the gravitational torque is negligible, and the net angular momentum current is then $j \simeq \dot{M}_2 \Omega a^2$. We check if the gravitational torque is negligible in Section 4.5 and find that for a mass ratio of $q = 0.01$, the gravitational torque is non-negligible. However, as an approximation, we find that the ratio of the measured

Table 1. Initial values for the binary parameters used in the simulation results presented in this manuscript. These values are based on the fiducial binary system with $q = 0.01$, $M = 10^7 M_{\odot}$, and 10^4 orbits until merger, but they can be scaled accordingly for any binary black hole inspiral.

Binary parameter	Initial value
Initial separation:	$a_0 = 3.6 \text{ AU} \left(\frac{M}{10^7 M_{\odot}} \right) \left(\frac{\eta}{0.01} \right)^{2/5} \left(\frac{\tau/P}{10^4 \text{ orbits}} \right)^{2/5}$
Initial time until merger:	$t_{\text{minus}} = 5.4 \text{ years} \left(\frac{M}{10^7 M_{\odot}} \right)^{-1/2} \left(\frac{a_0}{3.6 \text{ AU}} \right)^{3/2} \left(\frac{\tau/P}{10^4 \text{ orbits}} \right)$
Initial primary BH gravitational radius:	$r_{g_1} = 2.7 \times 10^{-2} a_0 \left(\frac{M}{10^7 M_{\odot}} \right) \left(\frac{1+q}{1.01} \right)^{-1}$
Initial secondary BH gravitational radius:	$r_{g_2} = 2.7 \times 10^{-4} a_0 \left(\frac{M}{10^7 M_{\odot}} \right) \left(\frac{\eta}{0.01} \right)$

inflow rate $\langle \dot{M} \rangle$ to the nominal inflow rate \dot{M}_0 is then

$$\frac{\langle \dot{M} \rangle}{\dot{M}_0} \simeq \frac{1}{1 - \sqrt{a/r_{\text{buf}}}}. \quad (24)$$

In equation (24) we have also assumed the secondary dominates a majority of the accretion, i.e. $\dot{M}_2 \simeq \langle \dot{M} \rangle$. In simulations where $r_{\text{buf}} = 7.5a_0$, equation (24) predicts $\langle \dot{M} \rangle / \dot{M}_0 \simeq 1.57$, which is very close to what we observe.

We have checked that the overshoot does not contaminate any of our derived science results, such as the shape of the accretion rate time series. Overshoots of the kind we observed here could be mitigated by guessing the value of j/\dot{M} and then driving the solution in the buffer region towards $\Sigma(r)$ in equation (23). Note that overshoots of the binary accretion rate have also been reported in Farris et al. (2014) and Shi & Krolik (2015) and were also discussed in Rafikov (2016). We suggest here that the finite domain size could have been responsible for those observations.

Note, on the other hand, the accretion overshoot due to the finite outer boundary introduced in this section is not the same as the accretion overshoot described in Miranda, Muñoz & Lai (2017), Dittmann & Ryan (2022), and Duffell et al. (2024). The accretion overshoot described in those works will, over many orbits, slowly approach the theoretical inflow rate \dot{M}_0 as the disc settles to larger radii. The lasting overshoot we observe here is due to the finite radius of the domain outer boundary, and the gradual decline of the accretion rate towards \dot{M}_0 is halted when the disc would need to be viscously relaxed at radii larger than r_{buf} .

4 SIMULATION RESULTS

We simulated the GW-driven inspirals of black hole binaries with mass ratios between $q = 10^{-3}$ and $q = 10^{-1}$, surrounded by discs with orbital Mach numbers $\mathcal{M} = 10, 15, 20$. The binary mass is $M = 10^7 M_{\odot}$; however, the accretion rate time series could be rescaled by an appropriate remapping of the time coordinate to describe a binary of any mass. Results related to observer time-scales, EM light curves, and emission spectra depend on M , but they also scale with it in obvious ways. Simulations generally begin of the order of $\tau/P \sim 10^4$ binary orbits prior to the coalescence. In Section 4.7, we show that this provides sufficient time for the disc to relax and ‘forget’ the simulation’s initial condition well before the viscous decoupling stage. Times are measured in the binary rest-frame, not in the observer time, i.e. we do not adjust the time-scales for cosmological redshift.

We have summarized the initial binary parameters used in our results in Table 1 for a binary with a total mass of $M = 10^7 M_{\odot}$, a mass ratio of $q = 0.01$, and 10^4 orbits until merger. As shown in the

table, these values can be scaled for different total binary masses, mass ratios, and numbers of orbits before merger using equations (1) and 2. Consequently, not all results in this work use the exact initial values given in Table 1, as our study explores a broad range of mass ratios and varying initial orbits before merger.

4.1 Evolution of disc surface density

Fig. 3 shows the surface density radial profiles of the fiducial disc, initially at $t_{\text{minus}} = 1$ yr and then at four other representative times leading up to merger. At radii $r \gtrsim a_0$, the surface density profile is approximately $\Sigma \propto r^{-1/2}$, as for a steady state α -disc in equation (21). The surface density evolution is similar in certain ways to the one-dimensional disc model from Chang et al. (2010). An annular gap is formed around the secondary orbit, and the outer disc follows the secondary inwards up until the viscous decoupling time. However, the gap in our simulation is not fully evacuated; gas flows across the gap and accretes to the secondary. Also, the surface density of the inner disc in our simulation is seen to be modestly decreasing over time, whereas in Chang et al. (2010), the presence of the vacuum gap around the secondary orbit requires the surface density of the inner disc to increase. We show in the next subsection that the decrease of the inner disc mass is driven mainly by accretion to the secondary. In other words, negligible mass passes the secondary’s orbit, and instead, most gas is captured by the secondary, gradually starving the primary.

The surface density radial profiles for varying Mach number discs are shown in Fig. 4 at 4000 orbits before merger (dashed) and at merger (solid). As the Mach number increases, the viscosity decreases, and while the inflow rate $\dot{M}_0 = 3\pi\nu\Sigma$ is held constant across all runs, the surface density magnitude must then increase. This explains the shift in surface density for higher Mach number discs.

4.2 Long-term viscous decoupling

Fig. 5 shows the binary accretion rate \dot{M} (top), the accretion rates to the binary components \dot{M}_1 and \dot{M}_2 (middle), and the preferential accretion rate \dot{M}_2/\dot{M}_1 (bottom). The curves show the time series data for binaries with various mass ratios. Our main results here can be summarized: (1) the secondary generally starves more slowly than the primary, and (2) the binary accretion rate \dot{M} shows a long-term downward trend, being reduced from $\langle \dot{M} \rangle$ by 10 per cent as early as $\sim 500\tau_{\text{dec}}$ (decade-like time span) before the merger.

The second result is surprising, as one could reasonably guess that \dot{M} would be significantly reduced from \dot{M}_0 only within a

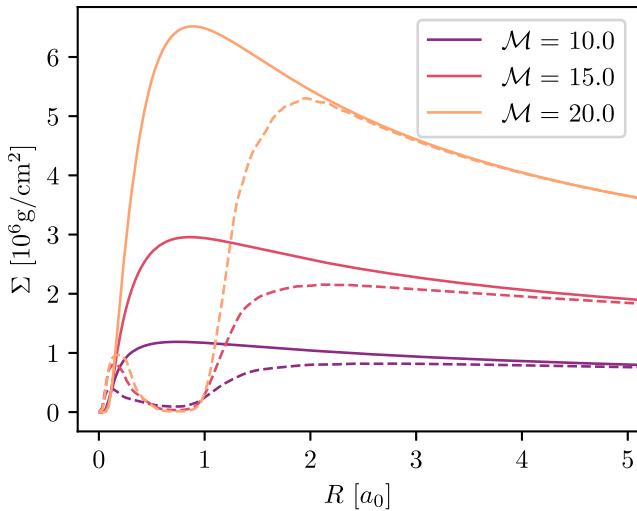


Figure 4. The radial profile of the surface density $\Sigma(r)$ in a binary with mass ratio $q = 0.01$ and orbital Mach numbers in the range $\mathcal{M} = 10, 15, 20$, shown at 4000 orbits before merger (dashed) and at merger (solid). The radial profiles are obtained from the azimuthal average of the two-dimensional disc surface density.

few τ_{dec} of the merger, perhaps following a functional form like $\dot{M} = \dot{M}_0 e^{-\tau_{\text{dec}}/t_{\text{minus}}}$. Instead, our results suggest the quantity $\delta \equiv 1 - \dot{M}/\dot{M}_0$ might grow as a power law in t_{minus} , at least while δ is small. We summarize next a model of the viscous decoupling process, which accounts for the time evolution of the binary-disc torque and predicts such scaling of δ . A full derivation is given in Zrake, Clyburn & Feyan (2025).

Our model characterizes the disc as having two distinct zones: a close zone that is tightly coupled to the binary and a distant zone that is relatively frozen over the remaining time before the merger. The two zones meet at a viscous radius, r_v , satisfying $t_{\text{visc}}(r_v) = -a/\dot{a}_{\text{gw}}$.² In the close zone $r < r_v$, the disc evolves through a sequence of steady states (equation 23), each conducting angular momentum inwards at rate $\dot{J} = \dot{M}\Omega a^2$. At each step n of the sequence, a boundary condition is imposed, $\Sigma_n(r_n) = \Sigma_{n-1}(r_n)$, where r_n is the viscous radius at step n . Such recursion leads to a first-order ordinary differential equation for δ , and the solutions go asymptotically as $\delta \propto t_{\text{minus}}^{-5/24}$ (Zrake et al. 2025). Such weak dependence of δ on t_{minus} could account for the long-term viscous decoupling seen in our simulations. Fig. 6 shows a plot of $\delta(t)$ for the fiducial binary compared to the asymptotic prediction $\delta \propto t_{\text{minus}}^{-5/24}$. With many decoupling time-scales before the merger, the simulations agree with the predicted power-law dependence.

In Section 2.3, we estimated that the viscous decoupling time-scale would be very sensitive to the Mach number of the disc, $\tau_{\text{dec}} \propto \mathcal{M}^{16/5}$. However, given our argument above, where δ scales as a power law in t_{minus} , it seems reasonable to expect that the late-inspiral accretion rate curves might have only modest sensitivity to the Mach number. To test this, we ran a suite of simulations with a range of

²This idea is inspired by section 3.2 of Rafikov (2016), which defines a viscous radius, but in a different setting where the binary orbit is held fixed, while the disc viscously adjusts to a new angular momentum current. The idea is also used in section 5.6 of Duffell et al. (2024) to explain the damping of a startup transient in simulations of binary accretion initiated with a zero-torque disc.

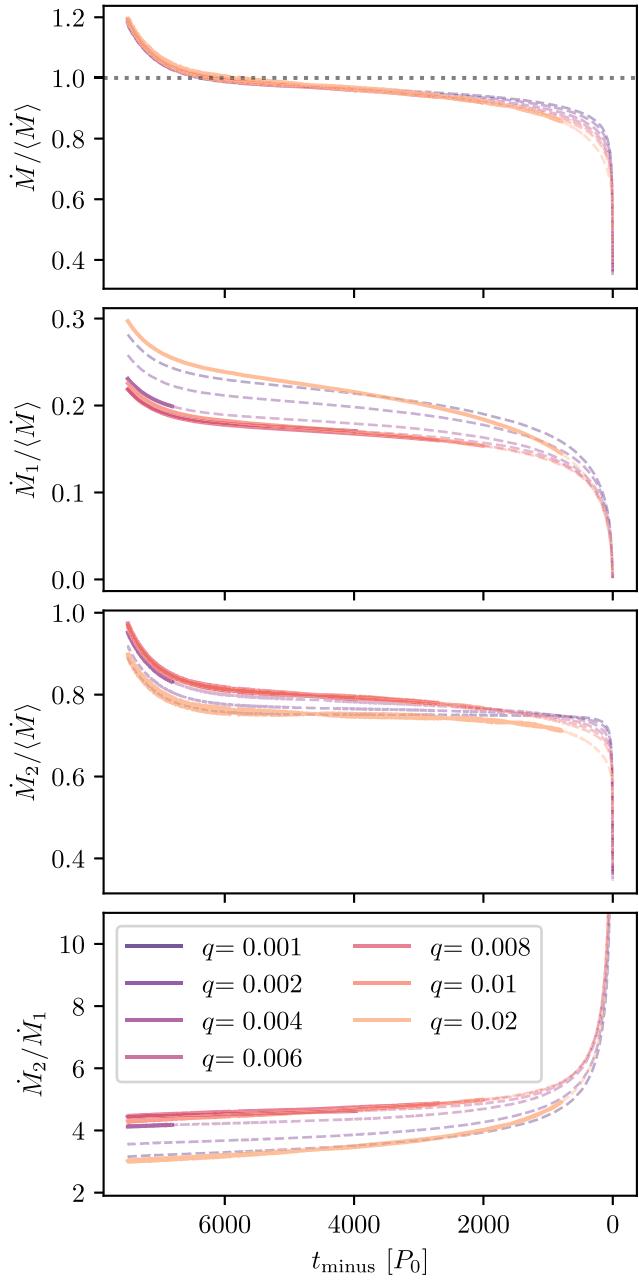


Figure 5. Time series of the mass accretion rates \dot{M} (top), \dot{M}_1 (second), \dot{M}_2 (third), and the preferential accretion parameter \dot{M}_2/\dot{M}_1 (bottom), for mass ratios q in the range of 0.001–0.02. Each line starts as a solid line and transitions to a dashed line when the sink radius of the secondary BH becomes comparable to the size of the Hill radius, indicating that gas flows in the secondary disc are not well resolved for the dashed curves. The gas orbital Mach number is $\mathcal{M} = 10$. The dotted line in the top panel indicates where $\dot{M} = \langle \dot{M} \rangle$.

Mach numbers, $\mathcal{M} = 10, 15, 20$. The results appear in Fig. 7, and do confirm this expectation. The run with Mach 20 has a nominal viscous decoupling time roughly 10 times larger than the run with Mach 10, and yet the curves differ mainly by a vertical offset. The long-term decoupling effect seen here also seems to be present in simulations of equal-mass inspirals in, e.g. fig. 6 in Farris et al. (2014), or fig. 3 of Dittmann et al. (2023).

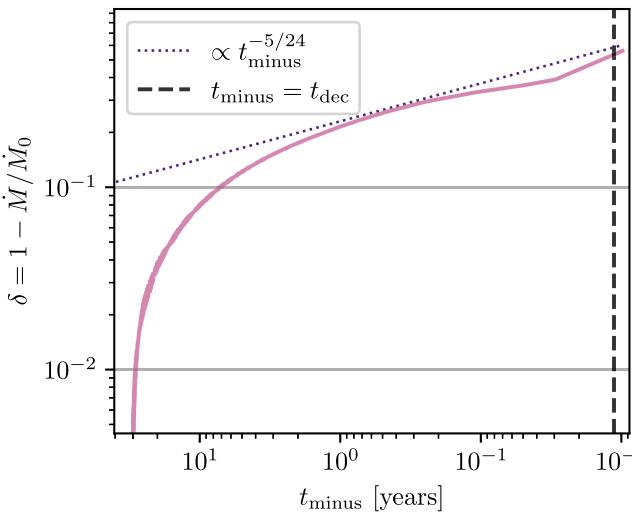


Figure 6. Time series of the $\delta = 1 - \dot{M}/\dot{M}_0$ parameter for the fiducial binary with mass ratio $q = 0.01$ and $\mathcal{M} = 10$. The dotted line represents the predicted power-law dependence, and the dashed vertical line is the decoupling time-scale.

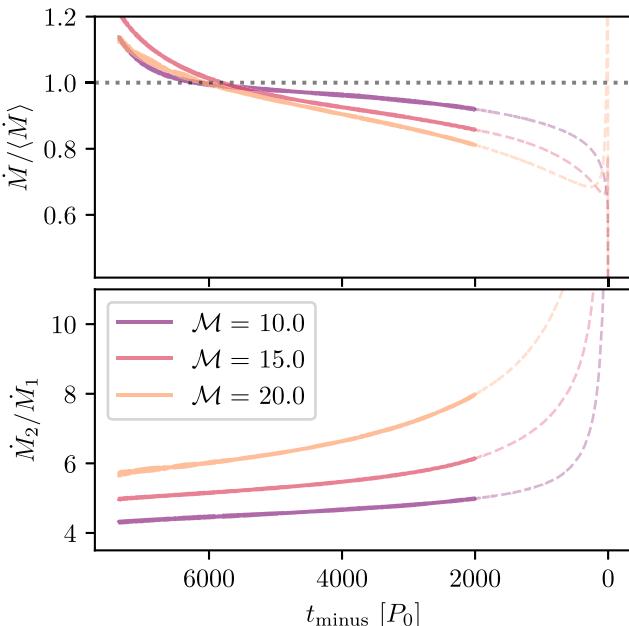


Figure 7. Time series of the mass accretion rates \dot{M} (top) and the preferential accretion parameters \dot{M}_2/\dot{M}_1 (bottom) for a range of Mach numbers, $\mathcal{M} = 10, 15, 20$. The binary mass ratio is $q = 0.01$. Each curve in the top panel is normalized by the empirical average binary accretion rate throughout the inspiral, $\langle \dot{M} \rangle$. Each line starts as a solid line and transitions to a dashed line when the sink radius of the secondary BH becomes comparable to the size of the Hill radius, indicating that gas flows in the secondary disc are not well resolved for the dashed curves. The dotted line in the top panel indicates where $\dot{M} = \langle \dot{M} \rangle$.

4.3 Remnant fuelling

Equally important to the pre-merger accretion evolution is the post-merger remnant refuelling, where the disc viscously refills the cavity and accretes on to the merger remnant. In Section 2.6, we derived the refilling time-scale t_{refill} (equation 11) and found that for the fiducial binary, the disc should refill the cavity, and begin accreting

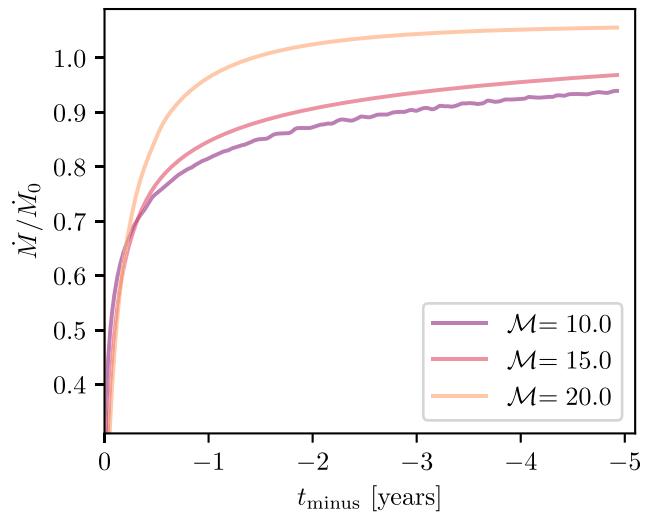


Figure 8. Time series of the post-merger mass accretion rate for the fiducial model, but with a range of Mach numbers $\mathcal{M} = 10, 15, 20$.

to the black hole merger remnant roughly 4 d after the merger. Note, however, that our estimate predicts $t_{\text{refill}} \propto \mathcal{M}^{16/5}$, analogous to the viscous decoupling time-scale τ_{dec} . It would imply that the onset of refuelling could be delayed from days (Mach 10) to decades if the disc had a Mach number of 100.

In Fig. 8, we plot the total binary accretion rate post-merger out to 5 yr after coalescence for a binary with mass ratio $q = 0.01$ and for simulations with different Mach numbers. The result indicates that the onset of refuelling happens over a time span of months to years, with relatively little sensitivity to our derived estimate for t_{refill} . We thus suggest that the refilling time-scale post-merger is not well represented by a time constant such as that given in equation (11), as the recovery of the accretion rate to \dot{M}_0 does not show any exponential characteristics. The exact functional form of the refuelling curve can be predicted from the model, which we summarized in Section 4.2, and comparisons are presented in Zrake et al. (2025).

4.4 Variability of mass accretion rates

Low mass ratio binaries seem to accrete with lower stochastic variability than binaries of near-equal mass (e.g. Farris et al. 2014; Shi & Krolik 2015; D’Orazio et al. 2016; Duffell et al. 2020; Dittmann & Ryan 2024). This suggests that accreting IMRIs could produce ‘cleaner’ light curves, possibly making it easier to identify them based on EM periodicity, or colour evolution, in the years before a merger.³ In this section, we show that the reduced stochasticity seen by Duffell et al. (2020) and Dittmann & Ryan (2024) applies throughout a GW-driven inspiral.

The time series of the secondary accretion rate \dot{M}_2 is shown in Fig. 9 for a $q = 0.01$ binary, and in Fig. 10 for a $q = 0.1$ binary. The amplitude of the stochastic variability in \dot{M}_2 is less than roughly 10 per cent for the case of $q = 0.01$ and greater than roughly 50 per cent for the case of $q = 0.1$, which confirms there is a significant

³Note a caveat of this prediction is that AGN tends to, in general, be highly variable. Therefore, any variability due to a binary could be difficult to detect regardless of binary mass ratio. Additionally, Noble et al. (2012) discovered using 3D GRMHD simulations that the luminosity has a suppressed variability amplitude relative to accretion rates, suggesting that our predictions could be optimistic.

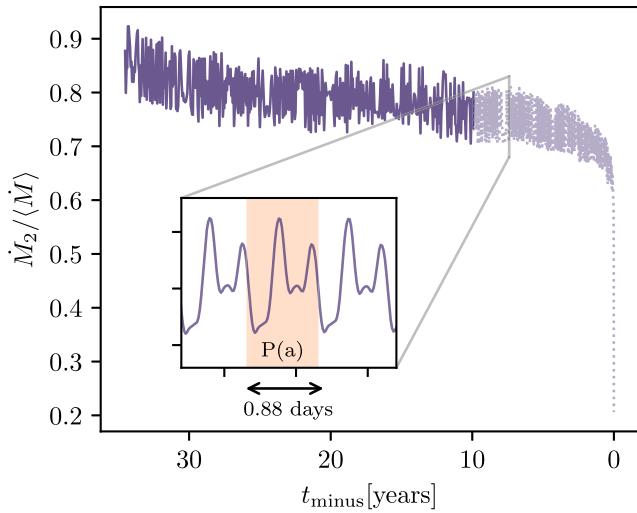


Figure 9. Time series of the secondary black hole mass accretion rate for a binary with a mass ratio of $q = 0.01$ and $\mathcal{M} = 10$. The accretion rate curve starts as a solid line and transitions to a dashed line when the sink radius of the secondary BH becomes comparable to the size of the Hill radius, indicating that gas flows in the secondary disc are not well resolved for the dashed curves. The highlighted region in the inset plot shows the (unsmoothed) accretion rate during one rest-frame binary orbit, which is approximately 0.9 d at $t_{\text{minus}} = 8$ yr when $q = 0.01$.

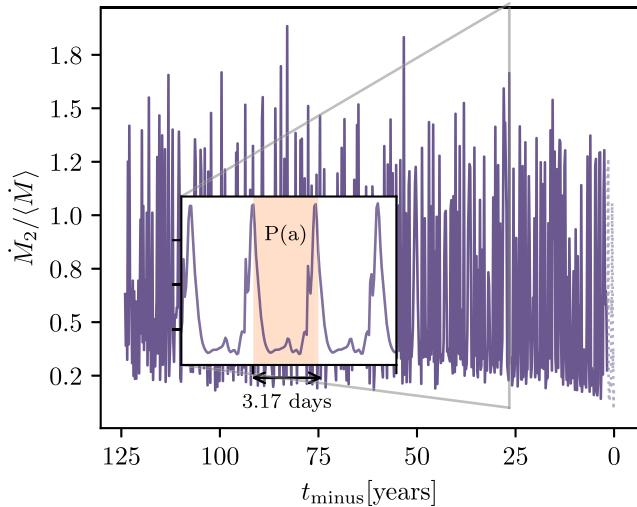


Figure 10. Time series of the secondary black hole mass accretion rate for a binary with a mass ratio of $q = 0.1$ and $\mathcal{M} = 10$. The accretion rate curve starts as a solid line and transitions to a dashed line when the sink radius of the secondary BH becomes comparable to the size of the Hill radius, indicating that gas flows in the secondary disc are not well resolved for the dashed curves. The highlighted region in the inset plot shows the (unsmoothed) accretion rate during one rest-frame binary orbit, which is approximately 3.2 d at $t_{\text{minus}} = 25$ yr when $q = 0.1$.

increase in the variability amplitude somewhere between $q = 0.01$ and $q = 0.1$, and that this change persists also in the late-inspiral phase, at least before the viscous decoupling.

The periodic variability of unequal-mass systems can be seen in the insets of Figs 9 and 10. The secondary accretion rate modulates at the binary orbital period, and we believe this could indicate a slight eccentricity of the outer disc (e.g. Kley & Dirksen 2006). For lower

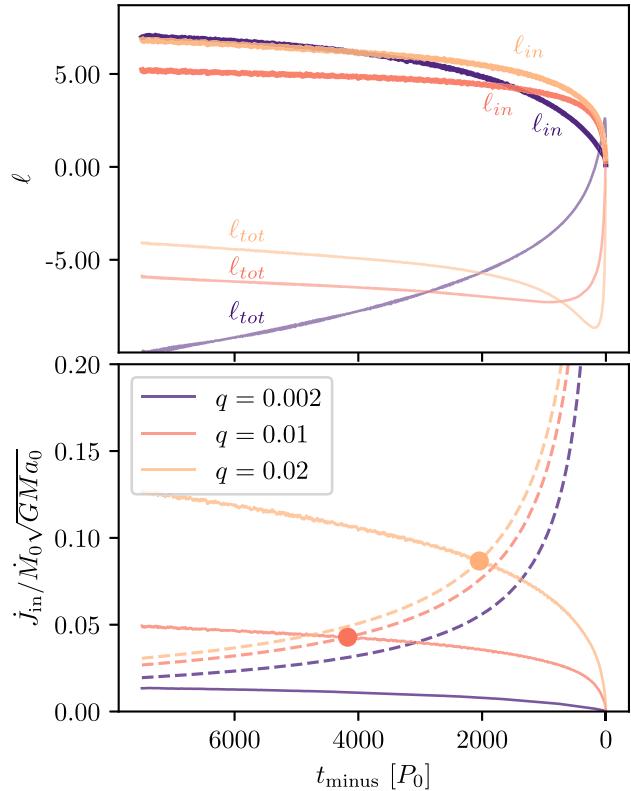


Figure 11. Evolution of the gravitational torque, normalized in different ways, for mass ratios $q = 0.002, 0.01, 0.02$ and Mach number $\mathcal{M} = 10$. In the bottom panel, the solid lines are the measured torque \dot{J}_{in} , applied to the binary by the inner disc. The dashed lines show the minimum torque \dot{J}_{plow} (equation 9) needed to sustain tidal squeezing, and the filled circles are where $\dot{J}_{\text{in}} = \dot{J}_{\text{plow}}$.

mass ratios, a second peak is more predominant in the secondary's accretion rate. This half-orbital periodicity has been seen in other works (e.g. Farris et al. 2014; Cocchiararo et al. 2024; Dittmann & Ryan 2024), and is likely the signature of an eccentric outer disc.⁴ We have measured the eccentricity of the outer disc in our runs by summing the local gas eccentricity vectors and found the eccentricity to be approximately 2 per cent for the $q = 0.01$ binary and around 14 per cent for the $q = 0.1$ binary.

4.5 Evolution of binary torques

Fig. 11 shows the time series of the gravitational torque on the binary for three mass ratios ($q = 0.002, 0.01, 0.02$). The top panel shows the non-dimensional torques $\ell_{\text{tot}} = (\dot{J}_{\text{tot}}/J)/(\dot{M}/M)$ and $\ell_{\text{in}} = (\dot{J}_{\text{in}}/J)/(\dot{M}/M)$, and the bottom panel shows the inner disc gravitational torque \dot{J}_{in} , normalized by $\dot{M}_0 \sqrt{GMa_0}$. Fig. 12 shows the gravitational torque on the binary for simulations of varying Mach numbers. The top panel shows the total gravitational torque exerted on the binary, and the middle and bottom panels show, respectively, the inner and outer disc contributions.

In the top panel of Fig. 12, we find that the net gravitational torque on the binary is negative for all of our runs and that it

⁴Cocchiararo et al. (2024) also reported a second peak in the total accretion rate; the minor peak we see in the secondary accretion rate seems compatible with their results.

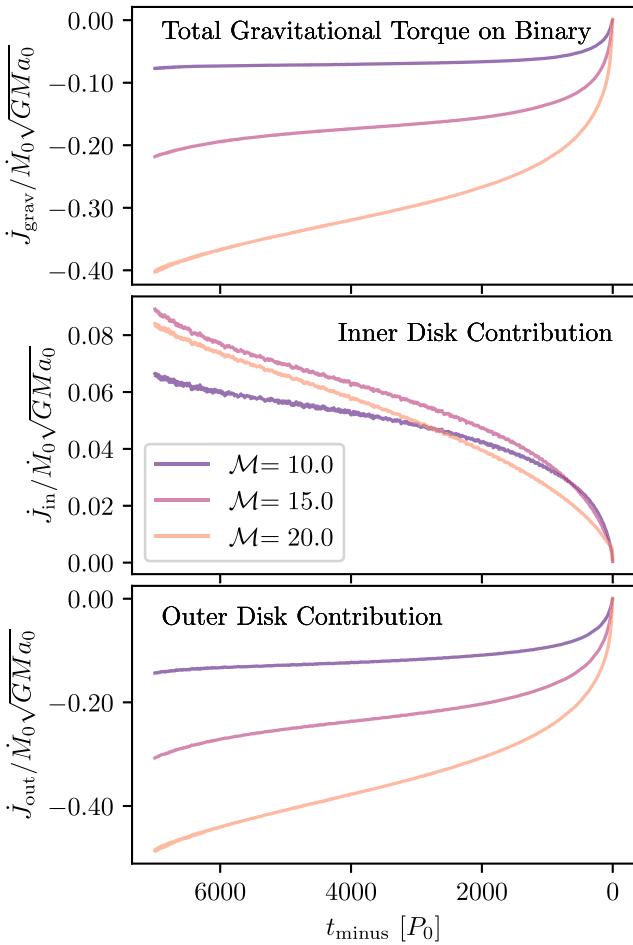


Figure 12. Evolution of the total gravitational torque \dot{J}_{grav} (top), inner disc gravitational torque \dot{J}_{in} (middle), and the outer disc gravitational torque \dot{J}_{out} (bottom) on the binary. The mass ratio is $q = 0.01$ and the Mach numbers are $\mathcal{M} = 10, 15, 20$.

becomes increasingly negative with increasing Mach number. This trend is consistent with other studies, including Tiede et al. (2020) and Penzlin et al. (2022), which also find that the torque on the binary becomes increasingly negative with increasing Mach number, albeit these studies focused on equal-mass or high mass ratio binaries. Dittmann & Ryan (2024), on the other hand, simulated unequal-mass binaries down to $q = 0.01$ and measured the gravitational torque of a $q = 0.01$ binary with $\mathcal{M} = 10$ to be marginally positive with $\dot{J}_{\text{grav}} \simeq 0.1$ and no clear trend for varying Mach number. Our results in Fig. 12, for the same binary system, show a marginally negative total gravitational torque of $\dot{J}_{\text{grav}} \simeq 0.1$ with a clear trend of monotonically decreasing torque with increasing Mach number. We believe the (small) difference might be attributable either to the difference in the viscosity prescription (constant- α versus constant- ν) or differences in the sink rate, as Dittmann & Ryan (2024) found some sensitivity of their torques to the value selected for the sink rate.

The outer disc is responsible for the increasingly negative torque at a higher Mach number; the inner disc actually exerts a slightly positive torque on the binary, which is rather insensitive to the Mach number. The increasing magnitude of torque from the outer disc can be attributed to the accumulation of mass outside the secondary orbit (Fig. 4).

These results could be used to predict how the gravitational waveform of IMRIs surrounded by gas should differ from those evolving in vacuum (e.g. Derdzinski et al. 2019, 2021; Tiede et al. 2024b). Such predictions will be reported in a future study focusing specifically on GW signatures.

4.6 Tidal squeezing of the inner disc

As discussed in Section 2.4, the theory of disc–satellite interactions implies that the orbiting secondary generally exerts a negative gravitational torque to the inner disc. When prescriptions for that torque are utilized in a one-dimensional disc model, they can lead to the formation of a zero-density, insulating gap in the disc surrounding the secondary. If the secondary is on a GW-driven inspiral, the insulating gap (which prohibits the flow of gas across the secondary’s orbit) then implies tidal squeezing of the inner disc (so-called ‘snowplow’ mechanism), and a potentially dramatic enhancement of the accretion rate \dot{M}_1 to the primary in the very late inspiral. In Section 2.4, we referenced other studies that have observed that the insulating gap does not form in two-dimensional hydrodynamics calculations, and we also pointed out that the magnitude of the gravitational torque, denoted as \dot{J}_{plow} , needed to reduce the inner disc radius at the rate \dot{a}_{gw} , diverges in the late inspiral. We report here the time series of gravitational torque \dot{J}_{in} , applied to the binary by the inner disc (as well as \dot{J}_{tot} , the torque from the whole disc), and confirm our estimate from Section 2.4 that in the final thousands of orbits, \dot{J}_{in} is exceeded by \dot{J}_{plow} , meaning that tidal squeezing becomes ineffective well before the merger.

Fig. 11 shows time series data from the gravitational torques on the binary for three mass ratios ($q = 0.002, 0.01, 0.02$). The bottom panel shows \dot{J}_{in} (solid) and \dot{J}_{plow} (dashed) as derived in equation (9). The filled circles are where $\dot{J}_{\text{in}} = \dot{J}_{\text{plow}}$. Notice that \dot{J}_{plow} diverges close to merger; the secondary black hole cannot provide enough torque to squeeze the inner disc in the final ~ 4000 orbits, consistent with the estimate given in Section 2.4. The top panel of Fig. 11 confirms an assumption we made in that estimate: that the non-dimensional torques do not change radically throughout the inspiral.

4.7 Numerical convergence

We ran many simulations to check the sensitivity of our results to numerical parameters, including the grid resolution and the start time of the simulation relative to the binary merger; see Appendix A for additional discussions on the sensitivity of our results to the sink size and sink rate. Fig. 13 presents the accretion rates of the fiducial binary for four simulations with different start times. The startup transients settle after roughly a hundred to a thousand orbits, and henceforth the accretion rates show very little sensitivity to the simulation start time. Also, in Fig. 13, we have plotted the total accretion rate for a fiducial binary that is not inspiralling. This also helps assure us that the long-term viscous decoupling discussed in Section 4.2 is uniquely a consequence of the inspiralling binary.

We also checked that our accretion rate and gravitational torque time series data are well converged with respect to the grid resolution. In Fig. 14, we have plotted the total accretion rate, preferential accretion rate, and gravitational torque for simulations performed with different resolutions and for two different mass ratios. Resolution is quantified by the grid spacing $\Delta x/a_0$ (better resolution means smaller Δx). Some disagreement is seen in both \dot{M}_1 and \dot{M}_2 and the gravitational torques when the grid spacing is greater than $\Delta x \simeq 0.016a_0$, especially in the very late-inspiral stage when the distance between the binary components becomes poorly resolved.

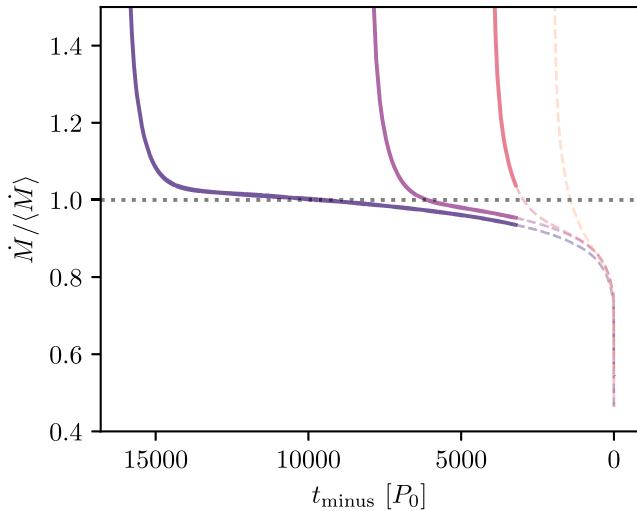


Figure 13. Time series of the binary mass accretion rate for a mass ratio of $q = 0.01$ and an orbital Mach number of $\mathcal{M} = 10$. The curves are from distinct runs initiated progressively earlier in the inspiral. The dotted line indicates where $\dot{M} = \langle \dot{M} \rangle$. Each curve starts as a solid line and transitions to a dashed line when the sink radius of the secondary BH becomes comparable to the size of the Hill radius, indicating that gas flows in the secondary disc are not well resolved for the dashed curves.

However, at early times, the shapes of the accretion rate curves, as well as the torques, are reasonably consistent with one another when the grid spacing is smaller than $\Delta x \simeq 0.016 a_0$. Additionally, we see convergence for both the fiducial mass ratio $q = 0.01$ and for the lowest mass ratio tested $q = 0.001$ in both the accretion rates and torques in Fig. 14.

5 ELECTROMAGNETIC SIGNATURES

The inspiral of unequal-mass MBHBs might produce tell-tale EM signatures. The detections of such EM signatures by ground or space-based observatories, especially alongside a GW detection by *LISA*, would provide an abundance of knowledge on binary evolution, accretion, and environments of MBHBs within galaxies. To maximize the likelihood of an EM detection, we must understand how the hydrodynamical interactions between the binary and disc affect the thermal radiation released from the accretion flow. In this section, we use the spectral toy model from Section 2.2, combined with the measured mass accretion rates, to provide a reasonable first approximation of the brightness and colour evolution of an accretion flow surrounding an inspiralling black hole binary of significantly unequal mass. Note that a full model of the variable AGN emission around the time of a merger will require a more detailed treatment of the radiative transfer, both within the disc and also throughout the larger environment. Nevertheless, we do believe the analysis presented here qualifies as a first approximation of the variable AGN disc emission before and after an MBHB merger.

5.1 Spectral evolution and light curves

In Fig. 15 we plot the composite spectral energy distribution of the accretion flow in the fiducial binary simulation, at different times throughout the inspiral. We used the three-disc toy model from Section 2.2, in which the spectrum is computed by approximating each disc (inner, secondary, and outer) to emit as a blackbody. The inner cut-off radii for the inner and secondary disc are $R_i = 6GM_i/c^2$,

the innermost stable circular orbits of the respective black hole components. The inner radius of the outer disc and the outer radius of the inner disc are set to the instantaneous separation of the binary, while the secondary disc is extended to the instantaneous Hill radius. The result is not sensitive to the outer radius of the outer disc.

In Fig. 16, we plot the soft UV (top), hard UV (second), soft X-ray (third), and hard X-ray (bottom) light curves derived from our three-disc toy model of the disc temperatures. We define the soft UV energy range as 1.5–4 eV, the hard UV energy range as 4–100 eV, the soft X-ray energy range as 0.2–2 keV, and the hard X-ray energy range as 2–12 keV. Note that once the binary reaches the decoupling separation, we have manually imposed in our model to evolve the outer disc cavity wall at the viscous drift rate rather than at the binary merger rate to highlight the difference in the EM signatures pre- and post-viscous decoupling. This is the reason for the abrupt change in the light curves seen around a couple of days prior to the merger, at the vertical dashed line corresponding to the decoupling time-scale in Fig. 16.

Years before merger, the majority of the overall accretion power is released as X-rays from the small secondary disc. However, as the inspiral proceeds, the outer disc reaches deeper into the gravitational potential of the primary, glowing brighter in the UV, while the secondary is gradually starved and gets dimmer in the X-rays. The outer disc follows the inspiralling secondary black hole, and as it does so, the inner edge of the outer disc increases in speed and heats up the gas, leading to a brightening effect. The UV brightening peaks at the time when the disc viscously decouples from the binary as seen in Fig. 15 at $t_{\text{minus}} \simeq 1$ d. For the fiducial disc–binary model, the X-ray emission is surpassed by the UV emission of the order of 1 month before the merger. At the same time as the peak of the UV emission, the X-ray emission from the secondary disc decays because of the fast reduction in the rate of mass supplied to the secondary very late in the inspiral (Krauth et al. 2023b). This is also evident in the measured accretion rates in Fig. 5.

We do not show the post-merger EM signatures associated with the remnant fuelling curves presented in Section 4.3, because the disc is far from being viscously relaxed while the accretion rate recovers to \dot{M}_0 . Analytic estimates for this phase are presented in Milosavljević & Phinney (2005), Schnittman & Krolik (2008), and Shapiro (2010). We also do not show the orbital period variations in the light curves or spectrum, as these cannot generally be captured in the three-disc model. Spectral evolution on the time-scale of individual binary orbits requires a self-consistent thermodynamic treatment and will be presented in future work (see Section 6.2).

5.2 Prospects for EM detections

Here, we consider how our results might influence proposed strategies to discover the EM counterparts of *LISA* GW detections. As discussed in the introduction, EM counterparts from *LISA* sources will likely be needed to confidently identify the host galaxies of MBHB inspirals.

Our results indicate that in the years before a merger, accreting unequal-mass black hole binaries get brighter in the UV and dimmer in the X-ray. They also exhibit an abrupt disappearance in X-rays in the final hours to days before the merger. We further predict that variations at the orbital period should be discernible at the tens-of-per cent level, resulting from orbital-period modulations of the accretion on to the secondary. Observatories carrying out time domain surveys, or space observatories capable of slewing in response to a GW trigger, could thus be instrumental in EM counterpart identification. We discuss below the relevance of these

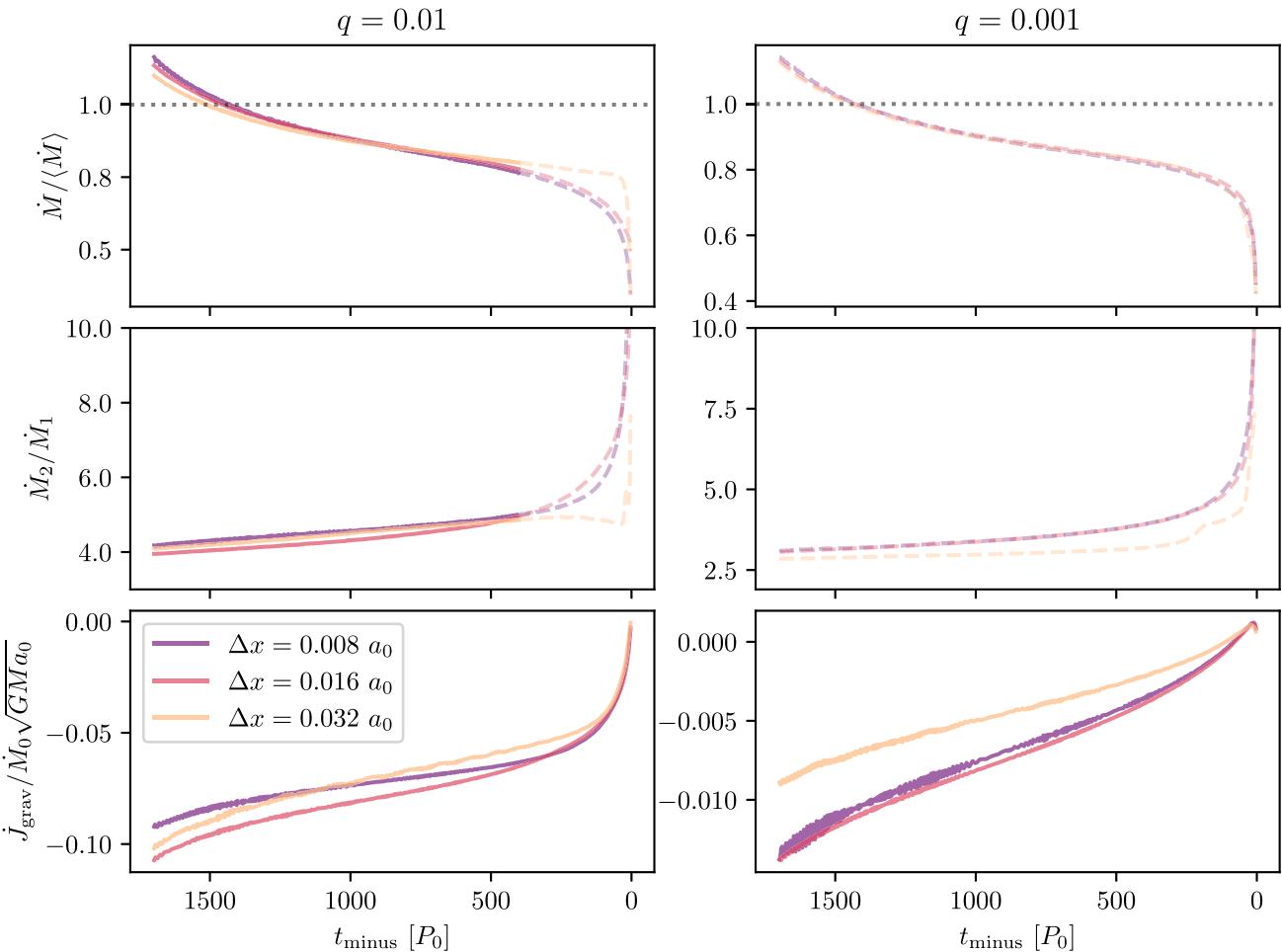


Figure 14. Time series of the mass accretion rates \dot{M} (top row), the preferential accretion rate \dot{M}_2/\dot{M}_1 (middle row), and the total gravitational torque \dot{J}_{grav} (bottom row) for distinct runs with grid spacings Δx in the range of $0.008a_0$ – $0.032a_0$. We show numerical convergence for two mass ratios, $q = 0.01$ (left column) and $q = 0.001$ (right column), with the orbital Mach number set to $\mathcal{M} = 10$ for both studies. The accretion rate curves start as solid lines and transition to dashed lines when the sink radius of the secondary BH becomes comparable to the size of the Hill radius, indicating that gas flows in the secondary disc are not well resolved for the dashed curves. The accretion rates are normalized by the empirical average binary accretion rate throughout the inspiral, $\langle \dot{M} \rangle$, and the dotted line in the top row indicates where $\dot{M} = \langle \dot{M} \rangle$.

three predictions (periodicity, UV brightening, X-ray dimming) to current and upcoming observatories.

With regard to periodicity in the EM signatures, the *Nancy Grace Roman Space Telescope*, a wide-field optical to near-infrared space telescope, could potentially detect these EM signatures in an accreting MBHB, corresponding to the quasi-periodic variability of the mass accretion rates shown in the insets of Figs 9 and 10. Such periodicity detections by *Roman* could become extremely valuable if the system is 10–20 yr away from merging because the merger might then take place during the *LISA* period of operation (Haiman et al. 2023). As discussed in Section 2.1, a $10^7 M_\odot$ binary that is 50 yr away from merging has a rest-frame orbital period of 1.8 d and in Fig. 9, the period is seen to decrease to 0.9 d, around 10 yr before merging. At $z \sim 1$ –2, this corresponds to multiday periodicity, around a factor of 2 shorter than the anticipated ~ 5 d cadence planned for the *Roman* high latitude time domain survey. Our calculations in Section 4.4 only predict the periodic variations of the mass accretion rates, so follow-up calculations will be needed to predict the level of periodicity expected where *Roman* is sensitive. This is an intended area of future work; see Section 6.2.

The primary EM signature predicted in this work is a gradual increase in UV brightness from the circumbinary disc, aligning with the contraction rate of the binary driven by gravitational radiation emission. The Legacy Survey of Space and Time (*LSST*) of the *Vera Rubin Observatory* will observe nearly half the sky, $18\,000 \text{ deg}^2$, and will detect up to 10^6 galaxies per square degree, observing quasars with a cadence of once per day (Xin & Haiman 2021, 2024). *Vera Rubin's LSST* camera will be sensitive in the soft UV energy range up to 4 eV. In the top panel of Fig. 16, we have plotted the light curve for the approximate *LSST* sensitivity. The steady UV brightening we are predicting to accompany a *LISA* inspiral could be evident in *LSST* archival data over the 10 yr prior to a GW detection by *LISA*; however, the UV-brightening is more pronounced and thus easier to detect in the hard UV energy ranges (second panel of Fig. 16). Therefore, telescopes in the hard-UV energy range, such as the recently selected NASA Explorer mission, *UVEX*, could be better suited to detect the predicted UV brightening (Kulkarni et al. 2021). Alternatively, since the characteristic temperature of the disc's inner edge decreases with the black hole mass, it is possible that *LSST* will be sensitive to the variable emission from higher mass AGNs.

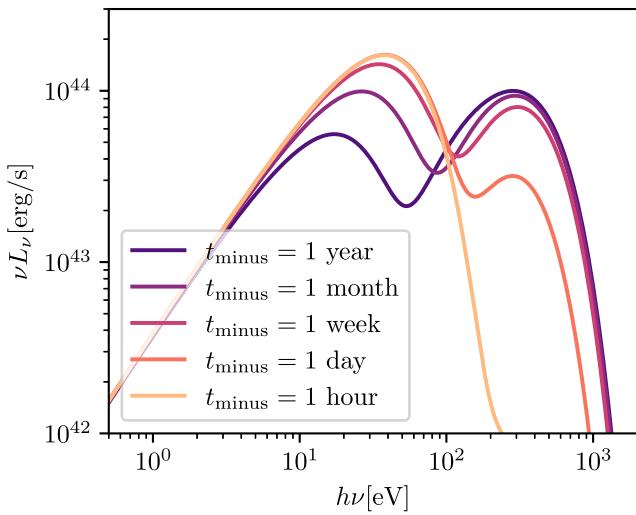


Figure 15. Evolution of the total spectral energy distribution of thermal emission produced by the fiducial system from 1 yr before merger (purple) to 1 h before merger (yellow). The model assumes that the accretion flow can be characterized as a steady state α -disc, with a radial temperature profile given by equation (3).

The X-ray dimming after viscous decoupling predicted in equal-mass binaries by Krauth et al. (2023b) and confirmed for unequal-mass binaries here could be detectable by the *Athena* X-ray telescope and/or *eROSITA*. *Athena* could potentially operate during the lifetime of *LISA* to have simultaneous multimessenger detections, while *eROSITA* is a current X-ray detector that could provide a reference sky template to aid in host galaxy characterization (Merloni et al. 2012). In the bottom two panels of Fig. 16, we have plotted the light curves for both the soft and hard X-ray energy ranges. The soft X-rays will be brighter initially, making them easier to observe; however, both energy ranges exhibit a significant drop in brightness due to viscous decoupling.

6 SUMMARY

6.1 Main results

Using analytic estimates and high-resolution grid-based hydrodynamics simulations, we have explored the gas dynamics and EM signatures of accreting binary black holes with significantly unequal masses, $q = 10^{-3}-10^{-1}$ and $M \sim 10^7 M_\odot$, over decade-like time-scales before a merger. Our aim is to help enable EM counterpart detection and host galaxy identification of binary black hole inspirals observed by *LISA*. Our main results are summarized here:

(i) The viscous decoupling time is generally of the order of $\tau_{\text{dec}} \sim \text{days}$ for $\sim 10^7 M_\odot$ binaries. One might reasonably expect the binary accretion rate \dot{M} to remain steady until $\sim \tau_{\text{dec}}$ before merger; however, our simulations show significant reductions in \dot{M} over decade-like time-scales, $10^3 \tau_{\text{dec}}$. We attribute this effect to the decrease of angular momentum current through the disc, accompanying the binary contraction. A model of this process is the subject of a forthcoming work.

(ii) Low mass ratio binaries show accretion rate variability at the orbital period, at the level of around 10 per cent. Stochastic variability is reduced throughout the inspiral when the mass ratio is $q = 0.01$ compared to $q = 0.1$, consistent with the trend reported in Dittmann & Ryan (2024).

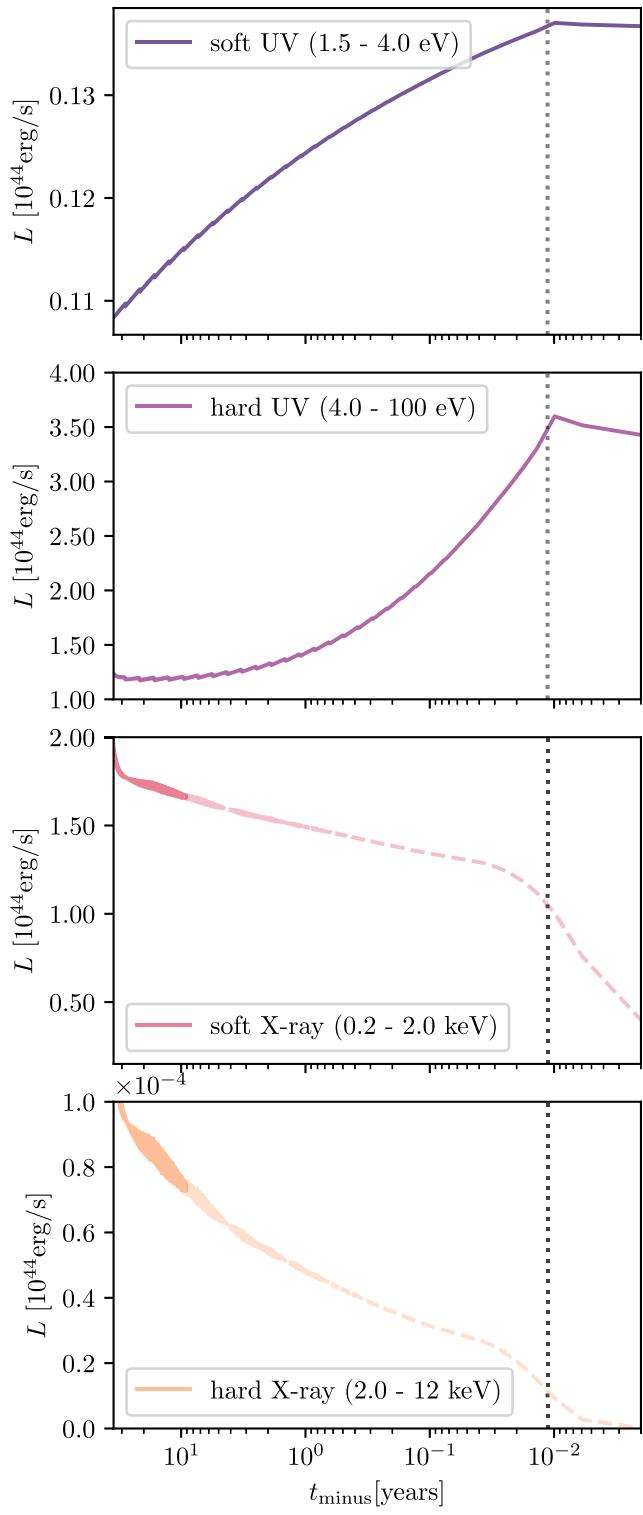


Figure 16. Light curves of the soft UV (1.5 – 4 eV), hard UV (4 – 100 eV), soft X-ray (0.2 – 2 keV), and hard X-ray (2 – 12 keV) emission, obtained by integrating the luminosity found in Fig. 1 over each energy band at each time. Light curves are shown for the fiducial binary where the mass ratio is $q = 0.01$ and the orbital Mach number is $\mathcal{M} = 10$. The X-ray light curves start as solid lines and transition to dashed lines when the sink radius of the secondary BH becomes comparable to the size of the Hill radius, indicating that gas flows in the secondary disc are not well resolved for the dashed curves. The dotted vertical line corresponds to the viscous decoupling time-scale in equation (7).

The secondary generally consumes most of the mass coming to the binary (preferential accretion effect). Both components are gradually starved of mass supply as the inspiral progresses, but the secondary starves more slowly than the primary.

(iii) EM flares produced by tidal squeezing of the inner disc (snowplow effect) are unlikely. This is based on dynamical considerations and confirmed in the simulations; the gravitational torque becomes too weak years before merger to reduce the orbits of inner disc gas parcels faster than \dot{a}_{gw} . Instead, the inner disc gas accretes to the secondary during the inspiral, leading to the observed slower starvation of the secondary.

(iv) We developed a toy model for the quasi-thermal emission spectrum of the accretion flow surrounding a low mass ratio binary black hole based on the assumption that the outer, inner, and secondary discs are viscously relaxed. The simple model predicts an emission spectrum with peaks in the UV and X-ray, reflecting the characteristic temperatures of the outer and circum-secondary discs, respectively. Those discs outshine the inner (primary) disc because of the preferential accretion effect.

(v) Together with the mass accretion rates measured in the simulations, the toy model predicts a decade-like gradual brightening in the UV, and dimming in the X-ray, followed by an abrupt disappearance of the X-rays around time $\tau_{\text{dec}} \sim$ days before the merger (analogous to Krauth et al. 2023b), and finally year-like gradual X-ray re-brightening associated with fuelling of the remnant black hole. The long-term pre-merger UV brightening is the result of the increasing peak temperature of the outer disc, as it extends deeper into the potential well of the primary throughout the inspiral.

6.2 Caveats and future work

Our simulations are based on two-dimensional (vertically averaged) solutions of the Navier–Stokes equations in the local isothermal approximation; the effects of shock heating, radiation, magnetic fields, general relativity, and out-of-plane gas flows, including winds and jets, are all neglected. Other groups are actively exploring these, especially in the context of near-equal mass ratio binaries in the final tens of orbits before a merger (see references given in the introduction). The simplified simulations presented here were necessary to study the long-term dynamics and EM signatures and may be accurate enough to help identify host galaxies of *LISA* events.

In future work, we will present simulations of unequal-mass inspirals with a more realistic thermodynamics treatment to properly account for shock heating, as done with the `SAILFISH` code in Westernacher-Schneider et al. (2022), Krauth et al. (2023a,b), DeLaurentiis et al. (2025), and Westernacher-Schneider et al. (2024). In the context of near-equal-mass binaries, shock-heating does significantly alter the profile of disc surface temperature (e.g. Farris et al. 2015). However, when the mass ratio is small, we expect less of the gas orbital energy to be dissipated in strong shocks, so the result might not be very different from the spectrum toy model we presented here.

We do not expect general relativistic effects to significantly alter the long-term brightness and colour changes of a binary-host quasar in the years before a merger because the long-term changes are driven by large-scale viscous relaxation of the disc far from the gravitational radii of either black hole. Accounting for magnetic fields and radiation could change some of our predictions, especially in a regime where the mass flow to the binary is comparable to the Eddington rate of the primary. In such cases, preferential accretion implies the secondary tries to accrete much faster than its own

Eddington rate and that radiation-driven outflows could significantly change the global mass currents in the system.

ACKNOWLEDGEMENTS

MC acknowledges support from the NASA Future Investigators Program (FINESST) through Award No. 80-NSSC-23K1443. JZ acknowledges support from the *LISA* Preparatory Science Program (LPS) through NASA Award No. 80-NSSC-24K0440. JZ also acknowledges valuable discussions with Alex Dittmann, Sterl Phinney, Julian Krolik, Elena Rossi, and Zoltan Haiman, some of which took place at the 2022 binary accretion workshop hosted by the Kavli Institute for Theoretical Physics (KITP); this research was supported in part by grant NSF PHY-1748958 to KITP. The authors would also like to thank the anonymous referee for their detailed reading of our manuscript and constructive input. All simulations were performed on Clemson University’s Palmetto cluster.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Amaro-Seoane P. et al., 2023, *Living Rev. Relativ.*, 26, 2
 Arca Sedda M., Amaro Seoane P., Chen X., 2021, *A&A*, 652, A54
 Armitage P. J., Natarajan P., 2002, *ApJ*, 567, L9
 Armitage P. J., Natarajan P., 2005, *ApJ*, 634, 921
 Avara M. J., Krolik J. H., Campanelli M., Noble S. C., Bowen D., Ryu T., 2024, *ApJ*, 974, 242
 Baibhav V., Berti E., Cardoso V., 2020, *Phys. Rev. D*, 101, 084053
 Barnes J. E., Hernquist L., 1996, *ApJ*, 471, 115
 Baruteau C., Ramirez-Ruiz E., Masset F., 2012, *MNRAS*, 423, L65
 Begelman M. C., Blandford R. D., Rees M. J., 1980, *Nature*, 287, 307
 Bellotroy J. et al., 2019, *BAAS*, 51, 175
 Bortolas E., Franchini A., Bonetti M., Sesana A., 2021, *ApJ*, 918, L15
 Bowen D. B., Mewes V., Campanelli M., Noble S. C., Krolik J. H., ao M. Z., 2018, *ApJ*, 853, L17
 Bowen D. B., Mewes V., Noble S. C., Avara M., Campanelli M., Krolik J. H., 2019, *ApJ*, 879, 76
 Science Collaboration LSST et al., 2009, preprint (arXiv:0912.0201)
 Chang P., Strubbe L. E., Menou K., Quataert E., 2010, *MNRAS*, 407, 2007
 Cocchiararo F., Franchini A., Lupi A., Sesana A., 2024, *A&A*, 691, A250
 Combi L., Armengol F. G. L., Campanelli M., Ireland B., Noble S. C., Nakano H., Bowen D., 2021, *Phys. Rev. D*, 104, 044041
 Cuadra J., Armitage P. J., Alexander R. D., Begelman M. C., 2009, *MNRAS*, 393, 1423
 D’Orazio D. J., Duffell P. C., 2021, *ApJ*, 914, L21
 D’Orazio D. J., Haiman Z., Duffell P., MacFadyen A., Farris B., 2016, *MNRAS*, 459, 2379
 DeLaurentiis S., Haiman Z., Westernacher-Schneider J. R., Krauth L. M., Davelaar J., Zrake J., MacFadyen A., 2025, *ApJ*, 980, 55
 Derdzinski A. M., D’Orazio D., Duffell P., Haiman Z., MacFadyen A., 2019, *MNRAS*, 486, 2754
 Derdzinski A., D’Orazio D., Duffell P., Haiman Z., MacFadyen A., 2021, *MNRAS*, 501, 3540
 Dittmann A. J., Ryan G., 2022, *MNRAS*, 513, 6158
 Dittmann A. J., Ryan G., 2024, *ApJ*, 967, 12
 Dittmann A. J., Ryan G., Miller M. C., 2023, *ApJ*, 949, L30
 Duffell P. C., Haiman Z., MacFadyen A. I., D’Orazio D. J., Farris B. D., 2014, *ApJ*, 792, L10
 Duffell P. C., D’Orazio D., Derdzinski A., Haiman Z., MacFadyen A., Rosen A. L., Zrake J., 2020, *ApJ*, 901, 25
 Duffell P. C. et al., 2024, *ApJ*, 970, 156

Escala A., Larson R. B., Coppi P. S., Mardones D., 2005, *ApJ*, 630, 152

Farris B. D., Liu Y. T., Shapiro S. L., 2010, *Phys. Rev. D*, 81, 084008

Farris B. D., Liu Y. T., Shapiro S. L., 2011, *Phys. Rev. D*, 84, 024024

Farris B. D., Gold R., Paschalidis V., Etienne Z. B., Shapiro S. L., 2012, *Phys. Rev. Lett.*, 109, 221102

Farris B. D., Duffell P., MacFadyen A. I., Haiman Z., 2014, *ApJ*, 783, 134

Farris B. D., Duffell P., MacFadyen A. I., Haiman Z., 2015, *MNRAS*, 447, L80

Franchini A., Sesana A., Dotti M., 2021, *MNRAS*, 507, 1458

Franchini A., Bonetti M., Lupi A., Sesana A., 2024, *A&A*, 686, A288

Fung J., Shi J.-M., Chiang E., 2014, *ApJ*, 782, 88

Garg M., Tiwari S., Derdzinski A., Baker J. G., Marsat S., Mayer L., 2024, *MNRAS*, 528, 4176

Goldreich P., Tremaine S., 1980, *ApJ*, 241, 425

Gong Y., Luo J., Wang B., 2021, *Nat. Astron.*, 5, 881

Gutiérrez E. M., Combi L., Noble S. C., Campanelli M., Krolik J. H., López Armengol F., García F., 2022, *ApJ*, 928, 137

Haiman Z. et al., 2023, preprint (arXiv:2306.14990)

Hughes S. A., 2002, *MNRAS*, 331, 805

Izquierdo-Villalba D., Sesana A., Bonoli S., Colpi M., 2022, *MNRAS*, 509, 3488

Kley W., Dirksen G., 2006, *A&A*, 447, 369

Krauth L., Davelaar J., Haiman Z., Westernacher-Schneider J. R., Zrake J., MacFadyen A., 2023a, *Phys. Rev. D*, 109, 103014

Krauth L. M., Davelaar J., Haiman Z., Westernacher-Schneider J. R., Zrake J., MacFadyen A., 2023b, *MNRAS*, 526, 5441

Kulkarni S. R. et al., 2021, preprint (arXiv:2111.15608)

Lin D. N. C., Papaloizou J., 1986, *ApJ*, 309, 846

Liu F. K., Wu X.-B., Cao S. L., 2003, *MNRAS*, 340, 411

Lodato G., Nayakshin S., King A. R., Pringle J. E., 2009, *MNRAS*, 398, 1392

Lubow S. H., Seibert M., Artymowicz P., 1999, *ApJ*, 526, 1001

Mangiagli A. et al., 2020, *Phys. Rev. D*, 102, 084056

Merloni A. et al., 2012, preprint (arXiv:1209.3114)

Miller M. C., 2009, *Class. Quantum Gravity*, 26, 094031

Milosavljević M., Phinney E. S., 2005, *ApJ*, 622, L93

Miranda R., Muñoz D. J., Lai D., 2017, *MNRAS*, 466, 1170

Muñoz D. J., Lai D., 2016, *ApJ*, 827, 43

Muñoz D. J., Lai D., Kratter K., Miranda R., 2020, *ApJ*, 889, 114

Noble S. C., Mundim B. C., Nakano H., Krolik J. H., Campanelli M., Zlochower Y., Yunes N., 2012, *ApJ*, 755, 51

Penzlin A. B. T., Kley W., Audiffren H., Schäfer C. M., 2022, *A&A*, 660, A101

Pereira F. A. C., Lodato G., Rodrigues I., Alves M. E. S., Price D. J., 2019, *MNRAS*, 484, 31

Peters P. C., 1964, *Phys. Rev.*, 136, 1224

Pitte C., Baghi Q., Marsat S., Besançon M., Petiteau A., 2023, *Phys. Rev. D*, 108, 044053

Pringle J. E., 1981, *ARA&A*, 19, 137

Rafikov R. R., 2016, *ApJ*, 827, 111

Roedig C., Dotti M., Sesana A., Cuadra J., Colpi M., 2011, *MNRAS*, 415, 3033

Ruan W.-H., Guo Z.-K., Cai R.-G., Zhang Y.-Z., 2018, preprint (arXiv:1807.09495)

Ruiz M., Tsokaros A., Shapiro S. L., 2023, *Phys. Rev. D*, 108, 124043

Schnittman J. D., Krolik J. H., 2008, *ApJ*, 684, 835

Schutz B. F., 1986, *Nature*, 323, 310

Sesana A., 2021, *Front. Astron. Space Sci.*, 8, 7

Sesana A., Roedig C., Reynolds M. T., Dotti M., 2012, *MNRAS*, 420, 860

Shakura N. I., Sunyaev R. A., 1973, *A&A*, 24, 337

Shapiro S. L., 2010, *Phys. Rev. D*, 81, 024019

Shi J.-M., Krolik J. H., 2015, *ApJ*, 807, 131

Siwek M., Weinberger R., Muñoz D. J., Hernquist L., 2023a, *MNRAS*, 518, 5059

Siwek M., Weinberger R., Hernquist L., 2023b, *MNRAS*, 522, 2707

Tang Y., Haiman Z., MacFadyen A., 2018, *MNRAS*, 476, 2249

Tiede C., D’Orazio D. J., 2024, *MNRAS*, 527, 6021

Tiede C., Zrake J., MacFadyen A., Haiman Z., 2020, *ApJ*, 900, 43

Tiede C., Zrake J., MacFadyen A., Haiman Z., 2024a, preprint (arXiv:2410.03830)

Tiede C., D’Orazio D. J., Zwick L., Duffell P. C., 2024b, *ApJ*, 964, 46

Trias M., Sintes A. M., 2008, *Class. Quantum Gravity*, 25, 184032

Valli R. et al., 2024, *A&A*, 688, 128

Westernacher-Schneider J. R., Zrake J., MacFadyen A., Haiman Z., 2022, *Phys. Rev. D*, 106, 103010

Westernacher-Schneider J. R., Zrake J., MacFadyen A., Haiman Z., 2024, *ApJ*, 962, 76

Xin C., Haiman Z., 2021, *MNRAS*, 506, 2408

Xin C., Haiman Z., 2024, *MNRAS*, 533, 3164

Young M. D., Clarke C. J., 2015, *MNRAS*, 452, 3085

Yunes N., Kocsis B., Loeb A., Haiman Z., 2011, *Phys. Rev. Lett.*, 107, 171103

Zrake J., Tiede C., MacFadyen A., Haiman Z., 2021, *ApJ*, 909, L13

Zrake J., Clyburn M., Feyan S., 2025, *MNRAS*, 537, 3620

APPENDIX A: SENSITIVITY TO NUMERICAL PARAMETERS

As mentioned in Section 3.1, the accretion rates of low mass ratio binaries are sensitive to the sink size of the black holes and should be interpreted with caution. In Figs A1 and A2, we plot the primary, secondary, and preferential accretion rates for different sink rates and sink radii, respectively. All simulations discussed in this section use a binary mass ratio of $q = 0.01$ and an orbital Mach number of $\mathcal{M} = 10$.

In Fig. A1, we present simulations with sink rates of $\tau_{\text{sink}}^{-1} = 5, 10, 20 \Omega_0$, each using a sink size of $r_{\text{sink}} = 0.05 a_0$. For the results presented in Section 6.1, we use a sink rate of $\tau_{\text{sink}}^{-1} = 10 \Omega_0$, which, according to Fig. A1, is high enough that the accretion rates are relatively insensitive to the sink rate. However, the accretion rate to the primary BH becomes spuriously large, and the accretion rate to the secondary BH becomes spuriously small when the sink rate is too low, i.e. $\tau_{\text{sink}}^{-1} < 10 \Omega_0$. Note that the preferential accretion rate increases with increasing sink rate. Although the components’ accretion rates are nearly identical at high sink rates, the preferential accretion depends on the ratio \dot{M}_2/\dot{M}_1 . Consequently, even small variations in \dot{M}_1 between runs with different sink rates can lead to appreciable changes in the preferential accretion.

In Fig. A2, we present simulations with sink radii of $r_{\text{sink}} = 0.05, 0.025, 0.0125 a_0$, each with a sink rate of $\tau_{\text{sink}}^{-1} = 10 \Omega_0$. As mentioned in Section 3.1, all simulations in this manuscript begin with $r_{\text{sink}} < r_{\text{hill}}$, but during inspiral, the Hill radius decreases, making the flow of gas around the secondary increasingly difficult to capture. For low mass ratios, the accretion rate to the primary BH becomes spuriously small, and the accretion rate to the secondary BH becomes spuriously large when the sink size is too large. Thus, the preferential accretion in the bottom panel of Fig. A2 is enhanced for smaller BH sinks. Nevertheless, the sensitivity of the accretion rates to the sink size does not affect our key results of a long-term viscous decoupling and a gradual UV brightening during inspiral.

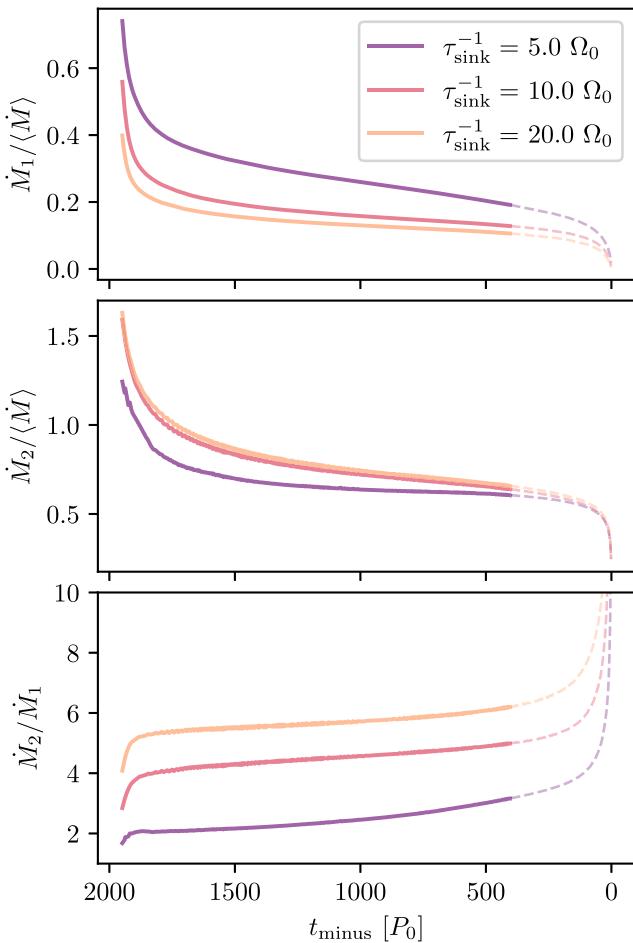


Figure A1. Time series of the mass accretion rates \dot{M}_1 (top), \dot{M}_2 (middle), and \dot{M}_2/\dot{M}_1 (bottom) for a binary mass ratio of $q = 0.01$ and a gas orbital Mach number of $\mathcal{M} = 10$, with varying sink rates and a sink size of $r_{\text{sink}} = 0.05a_0$. Each line starts as a solid curve and transitions to a dashed curve at the time when the secondary BH sink radius becomes comparable to the size of the Hill radius, beyond which the gas flow in the secondary disc is not well resolved for the dashed curves.

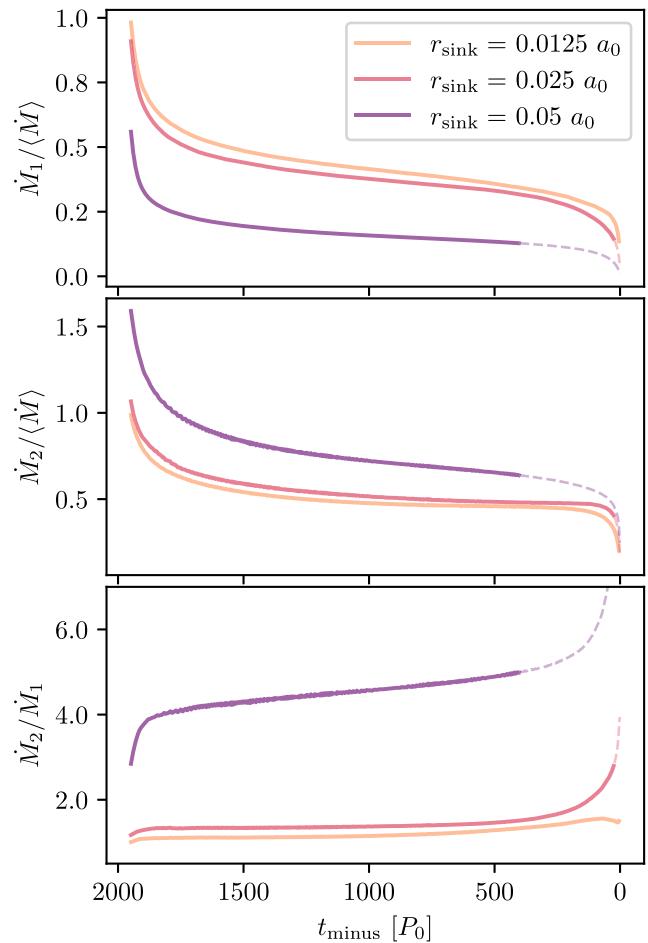


Figure A2. Time series of the mass accretion rates \dot{M}_1 (top), \dot{M}_2 (middle), and \dot{M}_2/\dot{M}_1 (bottom) for a binary mass ratio of $q = 0.01$ and a gas orbital Mach number of $\mathcal{M} = 10$, with varying sink sizes and a sink rate of $\tau_{\text{sink}}^{-1} = 10 \Omega_0$. Each line starts as a solid curve and transitions to a dashed curve at the time when the secondary BH sink radius becomes comparable to the size of the Hill radius, at which point the gas flow in the secondary disc is not well resolved for the dashed curves.

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