The Astrophysical Journal, 922:175 (29pp), 2021 December 1 https://doi.org/10.3847/1538-4357/ac2229

Mass-ratio and Magnetic Flux Dependence of Modulated Accretion from Circumbinary Disks

Scott C. Noble[1], Julian H. Krolik[2], Manuela Campanelli[3], Yosef Zlochower[3], Bruno C. Mundim[4], Hiroyuki Nakano[5], and Miguel Zilhão[6]
[1] Gravitational Astrophysics Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA; scott.c.noble@nasa.gov
[2] Physics and Astronomy Department, Johns Hopkins University, Baltimore, MD 21218, USA
[3] Center for Computational Relativity and Gravitation, Rochester Institute of Technology, Rochester, NY 14623, USA
[4] SciNet High Performance Computing Consortium, University of Toronto, Toronto, ON M5G 1M1, Canada
[5] Faculty of Law, Ryukoku University, Kyoto 612-8577, Japan
[6] Centro de Astrofísica e Gravitação—CENTRA, Departamento de Física, Instituto Superior Técnico—IST, Universidade de Lisboa—UL, Avenida Rovisco Pais 1, 1049-001, Portugal
Received 2021 March 23; revised 2021 August 23; accepted 2021 August 27; published 2021 November 29

Abstract

Accreting supermassive binary black holes (SMBBHs) are potential multimessenger sources because they emit both gravitational-wave and electromagnetic (EM) radiation. Past work has shown that their EM output may be periodically modulated by an asymmetric density distribution in the circumbinary disk, often called an "overdensity" or "lump;" this modulation could possibly be used to identify a source as a binary. We explore the sensitivity of the overdensity to SMBBH mass ratio and magnetic flux through the accretion disk. We find that the relative amplitude of the overdensity and its associated EM periodic signal both degrade with diminishing mass ratio, vanishing altogether somewhere between 1:2 and 1:5. Greater magnetization also weakens the lump and any modulation of the light output. We develop a model to describe how lump formation results from internal stress degrading faster in the lump region than it can be rejuvenated through accretion inflow, and predicts a threshold value in specific internal stress below which lump formation should occur and which all our lump-forming simulations satisfy. Thus, detection of such a modulation would provide a constraint on both mass ratio and magnetic flux piercing the accretion flow.

Unified Astronomy Thesaurus concepts: Supermassive black holes (1663); Astrophysical black holes (98); Magnetohydrodynamics (1964); Accretion (14); Black hole physics (159); Active galactic nuclei (16); Relativistic disks (1388); High energy astrophysics (739)

1. Introduction

Somewhere in the universe, several pairs of supermassive black holes (SMBHs) should merge every year, leaving behind a still more massive single black hole at the centers of the galaxies where this occurs (Klein et al. 2016; Katz et al. 2020). These events are extremely challenging to observe, but are of great interest because they are the most distant gravitational-wave sources we can hope to detect, and complementary photon and gravitational-wave data could provide uniquely powerful diagnostics of these events (Baker et al. 2019; Kelley et al. 2019; Mangiagli et al. 2020). In addition, the consequences of such mergers for galactic evolution are profound, including strong correlations between the galaxies and the (merged) central black holes. Moreover, the physics of accretion onto a binary is by no means limited to relativistic systems—it may also be applied to protoplanetary disks in binary stellar systems (Keppler et al. 2020).

A short time before the merger, these systems are in the post-Newtonian (PN) regime, in which the binary loses energy and inspirals rapidly due to the emission of gravitational radiation. Because of their small separations (a fraction of a parsec) and the greater chance of a relatively rare supermassive binary black hole (SMBBH) system being outside our local extragalactic neighborhood, spatially resolving the two BHs is unlikely. Hence, electromagnetic identification of SMBBH systems requires matching theoretical expectations to observed phenomena in their light curves, spectra, or polarization.

Accreting binaries whose mass ratios $q \equiv M_2/M_1 \gtrsim 0.01$ generically exhibit a gap within a radius $\sim 2a$, where a is the binary's semimajor axis. Matter travels across this gap in a pair of streams, which convey the accreted mass to small accretion disks ("mini-disks") around each member of the binary. The binary can also break the axisymmetry of the circumbinary accretion disk (MacFadyen & Milosavljević 2008; Noble et al. 2012; Shi et al. 2012), causing it to concentrate much of its inner rings' mass in a limited range of azimuthal angles, a feature we refer to as the "lump," following Noble et al. (2012) and Shi et al. (2012). As a result, the accretion rate onto the binary is modulated at a frequency $\simeq 0.2\Omega_{\rm bin}$ (here $\Omega_{\rm bin}$ is the binary orbital frequency).

Periodic modulation of the accretion rate can lead to a corresponding modulation of the system luminosity if the residence time of matter in the mini-disks is short compared to the modulation period (Bowen et al. 2018, 2019; d'Ascoli et al. 2018). Moreover, because the mass falling through the gap onto a black hole's mini-disk is expected to shock against the outer edge of the mini-disk, the hard X-rays radiated by that shocked gas should also be modulated on the same timescale as the accretion rate (Sesana et al. 2012; Roedig et al. 2014).

The heating rate directly associated with the lump is also modulated; when $q = 1$, the frequency is twice the beat frequency between the binary's orbital frequency and the lump's orbital frequency (Noble et al. 2012). The lump is bright for the same reason it exists. As shown by Shi et al. (2012), it forms because some of the matter leaving the circumbinary disk's inner edge gains enough angular momentum from the binary's gravitational torques to travel back out to the circumbinary disk—rendezvousing there with the matter it left shortly before. The shock associated with its return drives

heating within the lump. The modulation frequency has this beat frequency because passing close to a black hole triggers inflow from the disk's inner edge. In Noble et al. (2012), we put forth the idea that such a periodic signal could be used to identify SMBBHs photometrically and possibly constrain a SMBBH's mass ratio: the signal frequency would be less than twice the beat frequency for unequal-mass binaries because the secondary BH predominantly interacts with the lump in those cases (Noble et al. 2012).

Purported observations of periodic emission from AGN have been reported (Graham et al. 2015; Liu et al. 2015; Charisi et al. 2016). In principle, they might originate from periodic phenomena similar to those just described. However, evidence of a periodic signal in the candidate, PG 1302-102, identified in the first two of these studies, diminished with additional data (Liu et al. 2018);[7] larger surveys have turned up additional results that are marginal at best (Liu et al. 2019; Chen et al. 2020; Liao et al. 2021). Because well-established periodic variation might yet be found, it is critical for us to understand the particular conditions from which such a signature arises, as well as the relation between the period of such a modulation and the binary orbital period.

Circumbinary black hole accretion simulations have been conducted in numerous ways, and each way or method has demonstrated the development of a lump or nonaxisymmetric overdense feature under the right conditions and measured the effect of mass ratio on accretion flow properties. Most of the simulations have been performed using 2D α-model viscous hydrodynamics (2D VH) and Newtonian gravity using Eulerian grid-based codes (MacFadyen & Milosavljević 2008; D'Orazio et al. 2013, 2016; Farris et al. 2014, 2015; Muñoz & Lai 2016; Miranda et al. 2017; Derdzinski et al. 2019, 2021; Muñoz et al. 2019, 2020; Moody et al. 2019; Mösta et al. 2019; Duffell et al. 2020; Muñoz & Lithwick 2020; Tiede et al. 2020; Zrake et al. 2021). Others have used α-viscosity hydrodynamics with Newtonian gravity in 3D, either using SPH (Ragusa et al. 2016, 2017; Fontecilla et al. 2020; Heath & Nixon 2020; Ragusa et al. 2020) or Eulerian grid-based methods (Moody et al. 2019).

Unfortunately, VH simulations use ad hoc internal stress models that poorly represent the expected angular momentum transport mechanism in real systems: internal magnetic stress. Further, the vast majority of VH work uses unrealistic isothermal equations of state or neglects the vertical extent of the system. These approximations have real consequences to predictions; for instance, VH simulations are never turbulent, are laminar, and exhibit relatively steady accretion flows, which is very different from the red-noise-dominated variability always found in AGN light curves. 3D MHD simulations, such as ours here, eliminate these approximations. For example, MHD simulations of single-BH accretion disks generically show large amplitude variations in both their mass accretion rate and luminosity, and the Fourier power spectra of these variations are, like the observations, dominated by red noise (Noble & Krolik 2009; Reynolds & Miller 2009; Hogg & Reynolds 2016; Shiokawa et al. 2017; Bollimpalli et al. 2020). There have been a few MHD binary disk simulations using Newtonian gravity and an isothermal equation of state (Shi et al. 2012; Bankert et al. 2015; Shi & Krolik 2015, 2016), those using approximate GR spacetimes to describe the binary's gravitational influence and our thermodynamic model (Lopez Armengol et al. 2021), those using high-order PN gravity (such as ours here; Noble et al. 2012; Zilhão et al. 2015; Bowen et al. 2018, 2019), and those using full numerical relativity (NR) techniques (though not always evolving the spacetime in order to hold the binary to fixed separations) and our thermodynamic model (Farris et al. 2012; Gold et al. 2014a, 2014b; Paschalidis et al. 2021; see also the review by Gold 2019).

Only those using PN gravity and full NR have used more realistic thermodynamics and MHD. The full NR simulations, however, are often too expensive to run for O(100) binary orbits that preclude them from reaching relaxed conditions, demonstrating significant lump development, and covering a sufficiently vast temporal dynamic range needed for accurate variability analysis. Our 3D GRMHD simulations with PN gravity therefore fall into a particularly useful niche that allows one to include the most realistic physical assumptions for the lengths of time needed to explore the lump and variability in the circumbinary disk region.

In this paper, we begin to explore how the parameters characterizing circumbinary disks using 3D GRMHD and PN gravity affect the nature of the signal. Along the way, we will also attempt to further elucidate how the overdensity feature arises. Our approach will be to use the simulation called RunSE in Noble et al. (2012) as a benchmark, contrasting it with new simulations having different parameters, but all sharing the identical high-order PN spacetime—a binary black hole system with a nonevolving circular orbit of separation $20M$. These new simulations can be grouped into two sets. One is a survey of mass ratios: 1:1, 1:2, 1:5, and 1:10. The second studies the effects of differing amounts of mass and magnetic flux in the accretion disk. Our work here represents the first time anyone explores how mass ratio and magnetic field conditions affect the circumbinary disk with PN gravity and GRMHD techniques.

The specifications of our simulations are given in Section 2. So that we can cleanly separate the lump from smooth behavior, we begin our presentation of results with a description of axisymmetric features (Section 3). In the following section (Section 4), we report how distinctly lump behavior depends on parameters. With these results in hand, we are able to discuss the dynamics controlling the growth of the lump (Section 5.1). All the results are discussed together and summarized in Section 6. Animations of the runs discussed in this paper may be found online here: https://youtube.com/playlist?list=PLNaEA0qwDBaeApzLr2oarVKiO3AFnklTV.

2. Simulation Details

In this section, we explain the methods we use to model our circumbinary accretion disk system and the parameters specifying its state. The gravitational aspects of the simulation are described in Section 2.1, while the aspects regarding the circumbinary magnetized gas are given in Section 2.2. We discuss how we selected each run's parameters in Section 2.3.

We use geometrized units in which $G = c = 1$. We will use Greek letters (e.g., μ, ν, λ, κ) to represent spacetime indices [0, 1, 2, 3], and Roman letters (e.g., i, j, k, l) to represent spatial indices [1, 2, 3].

Although the unit of time in our calculations is M, the most important physical unit of time is the binary orbit's dynamical time $\Omega_{\rm bin}^{-1}$. This timescale depends on mass ratio q, but only weakly, due to post-Newtonian physics. For instance, the relative difference between the two most extreme orbital

[7] A Bayesian analysis shows that periodic signals with red noise may not always strengthen with additional time-series data, and that in PG 1302-102 the data weakly favor a quasiperiodic oscillation over the strictly periodic one expected from a binary (Zhu & Thrane 2020).

frequencies considered, $\Omega_{\rm bin}(q=1)$ and $\Omega_{\rm bin}(q=1/10)$, is 0.8%. In this paper, when we present data all taken from the same run, we will use the q-specific value for $\Omega_{\rm bin}$; when we discuss data in runs with different mass ratios, we will ignore this distinction.

2.1. Binary Black Hole Spacetime

Because our investigation focuses on dynamics close to the SMBBH, we can safely ignore the gravitational influence of the disk's gas and assume that gravity is entirely dictated by the binary. As in our previous work (Noble et al. 2012; Zilhão et al. 2015), we use a 2.5PN closed-form expression for the spacetime metric as described in Mundim et al. (2014). Only the so-called "Near Zone" (NZ) metric is used because our numerical domain does not extend either close enough to the black holes or far enough from the binary for the other zones to be needed. In the present work, unlike our previous, we will consider spacetimes in which the mass ratio, $q=M_2/M_1<1$, where the primary mass M_1 is always expected to be larger than that of the secondary BH (M_2), and we set $M_1+M_2=M=1$. We therefore concentrate in this section on the spacetime's dependence on q.

It is easiest to write down the metric in PN harmonic coordinates, a Cartesian basis system. Using the work of Blanchet et al. (1998), we find that the leading-order, nontrivial components of the NZ metric are

$$\begin{aligned} g_{00}^{\rm NZ} &= -1 + \frac{2M_1}{r_1} + \frac{2M_2}{r_2} + \mathcal{O}(v^4), \\ g_{0i}^{\rm NZ} &= -\frac{4M_1}{r_1}v_1^i - \frac{4M_2}{r_2}v_2^i + \mathcal{O}(v^5), \\ g_{ij}^{\rm NZ} &= \delta_{ij} + \frac{2M_1}{r_1}\delta_{ij} + \frac{2M_2}{r_2}\delta_{ij} + \mathcal{O}(v^4), \end{aligned} \tag{1}$$

where v_N^i denotes the coordinate velocity of the Nth BH, and $r_N=|\boldsymbol{x}-\boldsymbol{y}_N|$, i.e., computed from the Cartesian PN harmonic coordinates. Even though the metric is most simply represented in this Cartesian basis, in our simulation it is transformed to a spherical basis for use in our spherical coordinate system.

The NZ metric is valid only at distances more than $10M_N$ from the Nth black hole (Yunes & Tichy 2006; Yunes et al. 2006; Johnson-McDaniel et al. 2009). This constraint means that $r_{\rm in}$, the inner radial coordinate of our numerical domain, is bound from below to ensure the metric's validity. It can be estimated through the following argument. The positions of the two BHs relative to the center of mass are

$$\begin{aligned} \boldsymbol{y}_1 &= \frac{M_2}{M}\boldsymbol{a} + \mathcal{O}(v^4) = \frac{q}{1+q}\boldsymbol{a} + \mathcal{O}(v^4), \\ \boldsymbol{y}_2 &= -\frac{M_1}{M}\boldsymbol{a} + \mathcal{O}(v^4) = -\frac{1}{1+q}\boldsymbol{a} + \mathcal{O}(v^4), \end{aligned} \tag{2}$$

where $\boldsymbol{a}$ denotes the separation vector from the secondary BH to the primary BH, and we have ignored the 2PN order correction. The constraint on $r_{\rm in}$ is then

$$r_{\rm in} \gtrsim a\frac{1+q/2}{1+q}. \tag{3}$$

The ratio $r_{\rm in}/a$ increases from 0.75 for $q=1$ to 0.95 for $q=0.1$.

The NZ metric itself also depends on q:

$$\begin{aligned} g_{00}^{\rm NZ} &= -1 + \frac{2}{(1+q)r_1} + \frac{2q}{(1+q)r_2} \\ &= -1 + \frac{2}{r} + \frac{qa^2}{(1+q)^2 r^3}[3(\boldsymbol{n}\cdot\hat{a})^2 - 1], \\ g_{0i}^{\rm NZ} &= \frac{4q}{(1+q)^2\sqrt{a}}\left(-\frac{1}{r_1} + \frac{1}{r_2}\right)(\boldsymbol{\lambda})^i \\ &= \frac{4q}{(1+q)^2\sqrt{a}}\left(-\frac{a}{r^2}\,\boldsymbol{n}\cdot\hat{a}\right. \\ &\quad \left. + \frac{a^2(1-q)}{2(1+q)r^3}[3(\boldsymbol{n}\cdot\hat{a})^2 - 1]\right)(\boldsymbol{\lambda})^i, \\ g_{ij}^{\rm NZ} &= \left(1 + \frac{2}{(1+q)r_1} + \frac{2q}{(1+q)r_2}\right)\delta_{ij} \\ &= \left(1 + \frac{2}{r} + \frac{qa^2}{(1+q)^2 r^3}[3(\boldsymbol{n}\cdot\hat{a})^2 - 1]\right)\delta_{ij}, \end{aligned} \tag{4}$$

where $\boldsymbol{n}=\boldsymbol{r}/r$, $\hat{a}=\boldsymbol{a}/a$, $\boldsymbol{\lambda}$ is the unit vector tangent to the black hole's orbit, and we have included only the nontrivial leading-order PN terms. For each component, the second line shows a series expansion with respect to a/r up to $O((a/r)^2)$ beyond the leading-order nontrivial contribution.

The quantity $q/(1+q)^2$, which appears in every metric component, decreases monotonically as q falls from 1 to 0. The quadrupolar component in $g_{00}^{\rm NZ}$ and $g_{ij}^{\rm NZ}$, which has an angular dependence $[3(\boldsymbol{n}\cdot\hat{a})^2-1]$, therefore diminishes for smaller mass ratios. On the other hand, the quadrupolar component in $g_{0i}^{\rm NZ}$, which is a higher-order PN contribution to the equations of motion, has a maximum at $q\approx 0.27$.[8] For further details about the range of validity for the NZ metric, see Appendix A.

2.2. Matter

We use the flux-conservative code HARM3D to evolve the GRMHD equations on the dynamical spacetime (Gammie et al. 2003; Noble et al. 2006, 2009, 2012). It uses a piecewise parabolic reconstruction of the primitive variables at each cell interface, a Lax–Friedrichs-type flux, and a 3D version of the 2D FluxCT constrained transport scheme originally described in Tóth (2000). The stationary spacetime version, the same MHD methods as we use here, was described in Noble et al. (2010), while the modifications for handling dynamic spacetimes were explained in Noble et al. (2012). We refer the reader to these papers for further details on our numerical algorithms, and only briefly describe the equations of motion (EOM) for purposes of definition.

The EOM dictating magnetized gas's evolution are the Euler–Lagrange–Maxwell equations on a curved background spacetime with metric $g_{\mu\nu}$. The entire set may be written in the manifestly conservative form

$$\partial_t \boldsymbol{U}(\boldsymbol{P}) = -\partial_i \boldsymbol{F}^i(\boldsymbol{P}) + \boldsymbol{S}(\boldsymbol{P}), \tag{5}$$

[8] In the PN approach, g_{00} becomes the dominant contribution in the geodesic equation, as $\partial_i g_{00}$ and g_{0i} will be 1PN higher order. Therefore, in the Newtonian limit, the different q-dependence in g_{0i} can be ignored when considering its effect on the accretion flow.

where $\boldsymbol{U}$ is a vector of "conserved" variables, $\boldsymbol{F}_i$ is a vector of fluxes, and $\boldsymbol{S}$ is a vector of source terms:

$$\boldsymbol{U}(\boldsymbol{P}) = \sqrt{-g}\,[\rho u^t,\ T^t{}_t + \rho u^t,\ T^t{}_j,\ B^k]^T \tag{6}$$

$$\boldsymbol{F}^i(\boldsymbol{P}) = \sqrt{-g}\,[\rho u^i,\ T^i{}_t + \rho u^i,\ T^i{}_j,\ (b^i u^k - b^k u^i)]^T \tag{7}$$

$$\boldsymbol{S}(\boldsymbol{P}) = \sqrt{-g}\,[0,\ T^\kappa{}_\lambda \Gamma^\lambda{}_{t\kappa} - \mathcal{F}_t,\ T^\kappa{}_\lambda \Gamma^\lambda{}_{j\kappa} - \mathcal{F}_j,\ 0]^T. \tag{8}$$

Here $B^i = {}^*F^{it}/\sqrt{4\pi}$ is the magnetic field, ${}^*F^{\mu\nu}$ is the Maxwell tensor, u^μ is the fluid's 4-velocity, and $b^\mu = \frac{1}{u^t}(\delta^\mu{}_\nu + u^\mu u_\nu)B^\nu$ is the magnetic 4-vector (which is the magnetic field when projected into the fluid's comoving frame). In addition, $W = u^t/\sqrt{-g^{tt}}$ is the fluid's Lorentz function, $\mathcal{L}$ is the fluid-frame bolometric energy loss rate, and $\mathcal{F}_\mu = \mathcal{L}u_\mu$ is the radiative flux 4-vector. Lastly, g is the determinant of the metric, and $\Gamma^\lambda{}_{\mu\kappa}$ is the metric's affine connection. The stress-energy tensor $T_{\mu\nu}$ is defined as

$$T_{\mu\nu} = (\rho h + 2p_m)u_\mu u_\nu + (p + p_m)g_{\mu\nu} - b_\mu b_\nu, \tag{9}$$

where $p_m = b^\mu b_\mu/2$ is the magnetic pressure, p is the gas pressure, ρ is the rest-mass density, $h = 1 + \epsilon + p/\rho$ is the specific enthalpy, and ϵ is the specific internal energy.

Often, we look at reduced properties of the accretion flow and examine spherical shell-averaged or shell-integrated quantities. Shell averages are made using

$$\langle X \rangle \equiv \frac{\int X\ \sqrt{-g}\ d\theta d\phi}{\int \sqrt{-g}\ d\theta d\phi}, \tag{10}$$

while shell-integrated quantities are defined by

$$\{X\} \equiv \int X\sqrt{-g}\ d\theta\ d\phi = \langle X \rangle \int \sqrt{-g}\ d\theta\ d\phi. \tag{11}$$

Mass-weighted shell-averaged quantities are denoted as

$$\langle X \rangle_\rho \equiv \frac{\{X\rho\}}{\{\rho\}}. \tag{12}$$

We also sometimes smooth time-dependent quantities to highlight trends over longer timescales, and these are designated by overbars:

$$\overline{X}(t, \Delta t) = \frac{1}{\Delta t}\int_{t-\Delta t/2}^{t+\Delta t/2} X(t')dt'. \tag{13}$$

2.3. Run Details

The parameters defining run configurations are stated in Table 1. Measured quantities resulting from the runs are given in Table 2.

Our goal is to define how "lump" behavior (and a few other aspects of circumbinary disk dynamics) depend on mass ratio and magnetic content of the disk. To that end, we made every new simulation as similar as possible to our fiducial simulation, RunSE of (Noble et al. 2012). However, technical considerations demanded small departures in certain instances.

The binary was held at a fixed separation of $20M$ in all cases, like RunSE, so that we can explore the development of the lump and the quality of the periodic EM signal from it. Also like RunSE, in all runs, the binary orbit was circular, and the black holes were nonspinning.

We endeavored to make the problem spacetime volume very nearly the same as in RunSE, extending radially from an inner excision radius of $0.75a$ to an outer boundary radius of $13a$. However, in the three runs with $q < 1$, the inner excision radius was moved outward to a in order to accommodate the orbit of the secondary black hole, and in two runs the outer boundary was pushed out to $50a$ in order to contain a larger disk. The larger value of $r_{\rm in}$ is within the limit—$r_{\rm in}/a = 1.1$—at which Shi et al. (2012) began to see it significantly alter the structure of the circumbinary disk's inner edge. In every case, we used the same angular excision around the polar axis as in RunSE, and, as is necessary for binary dynamics, all runs covered a full 2π in azimuth. All the new runs had durations at least as long as the original RunSErun, from $1.25\times$ to more than $2\times$ that of RunSE.

The angular grid in every run was the same as in RunSE, with $N_\theta = 160$ cells in polar angle and $N_\phi = 400$ cells in azimuthal angle. The disks always satisfied the MRI quality conditions of Hawley et al. (2011) well, so there was no need to increase the number of points. The ratio $\Delta r/r$ was the same in every run, and the number of cells $N_r = 300$ was also the same in all but two. The outer radius of our numerical grid, $r_{\rm out}$, was chosen to lie beyond the extent of the initial distribution of gas. With $\Delta r/r$ fixed, $\mathrm{Run}_{\rm med}$ and $\mathrm{Run}_{\rm lrg}$ required more radial zones to reach their extended outer boundaries.

The initial gas distribution of each run was an equilibrium state with respect to the time average of the run's spacetime, in which the gas was supported against gravity by pressure gradients and rotation. Appendix A of Noble et al. (2012) provides the details of how the hydrostationary solutions were found. The parameters that constrain the solution are the radius of the inner edge of the disk $R_{\rm disk\ in}$, the radius of the pressure maximum R_p, and the initial aspect ratio (H/r) of the disk at the pressure maximum. All of these were identical to the values of RunSE, except in the two runs studying larger disks, where R_p was larger by 20%–25% and $R_{\rm disk\ out}$ was larger by a factor

Table 1
Parameters Determining the Simulations of the Mass Ratio Series and Magnetic Flux Series

Name	q	N_r	$r_{\rm in}$	$r_{\rm out}$	$R_{\rm disk\ in}$	R_p	$R_{\rm disk\ out}$
RunSE	1	300	0.75	13	3	5	11.7
$\mathrm{Run}_{q=1/2}$	0.5	300	1	13	3	5	11.7
$\mathrm{Run}_{q=1/5}$	0.2	300	1	13	3	5	11.7
$\mathrm{Run}_{q=1/10}$	0.1	300	1	13	3	5	11.7
$\mathrm{Run}_{\rm med}$	1	400	0.75	50	3	6	23.4
$\mathrm{Run}_{\rm lrg}$	1	420	0.75	50	3	6.5	39.1
$\mathrm{Run}_{\rm inj}$	1	300	0.75	13	3	5	11.7

Note. All radii are given in units of the binary separation, $a = 20M$.

Table 2
Measured Characteristics of the Simulations of the Mass Ratio Series and Magnetic Flux Series

Name	Σ_0	$\omega_{\rm lump}[\Omega_{\rm bin}]$	$T_{\rm lump}[10^3M]$	$t_{\rm end}[10^3M]$
RunSE	0.096	0.27 ± 0.05	52	76
$\mathrm{Run}_{q=1/2}$	0.083	0.26 ± 0.05	77	107
$\mathrm{Run}_{q=1/5}$	0.082	⋯	⋯	95
$\mathrm{Run}_{q=1/10}$	0.082	⋯	⋯	97
$\mathrm{Run}_{\rm med}$	0.085	0.25 ± 0.09	91	143
$\mathrm{Run}_{\rm lrg}$	0.087	0.26 ± 0.12	133	158
$\mathrm{Run}_{\rm inj}$	0.096	0.25 ± 0.26	113	126

Note. The circumbinary disk's initial peak surface density, Σ_0, is in units of code units for density times M^2.

2–3. In all cases, $H/r = 0.1$. The disks were initially isentropic, with entropy $K = p/\rho^{\Gamma} = 10^{-2}$ and adiabatic index $\Gamma = 5/3$. All simulations used an ideal-gas equation of state with $p = (\Gamma - 1)u$ for internal energy density u.

In every run, the initial magnetic field was given a poloidal distribution in the same way as in Noble et al. (2012). The magnetic field amplitude was normalized so that the ratio of the volume-integrated magnetic and gas pressure, an approximation to the volume-averaged plasma $\beta = p/p_m$, was 100.

So that we can tally the amount of energy dissipated during evolution, all runs cool to the same target entropy, which we choose to be the initial entropy, uniform throughout the flow. We parameterize this entropy by a proxy $K \equiv p/\rho^{\gamma}$. Its value in the initial state $K_0 = 10^{-2}$. A fluid element is cooled if its entropy is above K_0, and neither cooled nor heated if $K < K_0$. Writing $K = K_0 + \Delta K$, the cooling rate is

$$\mathcal{L} = \frac{\rho\epsilon}{T_{\rm cool}}\left(\frac{\Delta K}{K_0} + \left|\frac{\Delta K}{K_0}\right|\right)^{1/2}. \tag{14}$$

Note that Noble et al. (2012) erroneously omitted the exponent of $1/2$ in the paper, though they used it in the simulations reported there. The cooling time is $T_{\rm cool} = 2\pi(r/M)^{3/2}$, the period of a circular equatorial orbit at radius r without any quadrupolar contributions to the potential.

The cooling rate is recorded as 3D data for the duration of each simulation, so that it may serve as a proxy for the gas's bolometric emissivity.

With all these quantities fixed, we performed two series of parameter exploration runs, one varying the black hole mass ratio, the other varying the initial magnetic flux given the disk.

2.3.1. Mass Ratio Series

Real supermassive black hole binaries can have a variety of mass ratios. To measure the effect the mass ratio has on the circumbinary flow, we performed a series of simulations labeled RunSE, $\text{Run}_{q=1/2}$, $\text{Run}_{q=1/5}$, and $\text{Run}_{q=1/10}$, having mass ratios $q = M_2/M_1 = \{1, 1/2, 1/5, 1/10\}$, respectively, but all other physical parameters the same. This set was chosen in the hope of covering the whole range relevant to gas accretion. When $q \to 0$, the secondary black hole acts only as a mild perturber, producing little effect on the circumbinary flow. In fact, well before it reaches that limit, as we will show later in the paper, small q leads to weaker overdensities: even at $q = 1/5$, the lump amplitude is quite small if detectable at all. Simulations RunSE, $\text{Run}_{q=1/2}$, $\text{Run}_{q=1/5}$, and $\text{Run}_{q=1/10}$ ran until approximately 76 kM, 107 kM, 95 kM, and 97 kM, respectively.

2.3.2. Magnetic Flux Series

The idea that parts of an accretion disk can have regions of low angular momentum transfer, like the lump, is not a new one. For instance, in protoplanetary disks, the midplane of the accretion disk may be so shielded from cosmic radiation and its central source that it may be too cold to be adequately ionized and magnetized. Because accretion manifests from angular momentum transport mediated by magnetic stresses, such "dead zones" will be uncoupled from the rest of the disk (Gammie 1996; Gole et al. 2016). The transition from an actively accreting region to inactivity leads to a buildup of matter, similar to the development of our overdensity. Just as a protoplanetary dead zone may be revived by a local heating event thereby turning active again, so may our "dead" overdensity be eroded away if given an injection of additional magnetic flux. We present new runs here designed to see whether added magnetic flux may reignite activity in the dead zone or overdensity.

The control run for this series is also RunSE from Noble et al. (2012). From this run, we learned that the growth in the amplitude of the overdensity feature was coincident in space and time with a decline in how well the simulation can resolve the MHD turbulence, as defined by the quality factor $Q^i \equiv v_A^i/(\Delta x^i \Omega_K)$, where v_A^i is the Alfvén speed associated with the ith magnetic field component, Ω_K is the local circular orbit frequency, and Δx^i is the cell size in the ith direction (Noble et al. 2010; Hawley et al. 2011). Here v_{Ai} is the Alfvén speed for the magnetic field component in the ith direction and Δx^i is the cell size in that dimension. Two mechanisms, logical converses of each other, may explain this effect:

1. A decline in local MHD stress per unit mass (signaled by decreasing Q) fosters the growth of the overdensity.
2. The lump's increasing density decreases v_A, degrading the effective resolution of the simulation; this numerical effect then retards, or even eliminates, magnetic field growth.

It is also entirely possible that both act, reinforcing one another. As we will see, Mechanism 1, i.e., that a natural decline in MHD stress accommodates lump growth, is supported by our simulations.

The first possibility implies the correlation is physical and our results are potentially predictive. The second possibility implies that our simulations have little predictive power since we cannot say whether higher resolution (i.e., what nature uses) would yield an overdense feature.

With three additional runs, we aim to test whether either of these mechanisms operates. In $\text{Run}_{\rm inj}$, we ask whether a late-time strengthening of the magnetic field can, by restoring resolution quality, sustain magnetic stress despite increasing gas density in the lump. In $\text{Run}_{\rm med}$ and $\text{Run}_{\rm lrg}$, we increase the total magnetic flux available to the disk to test whether a stronger field retards lump growth.

For $\text{Run}_{\rm inj}$, at $t = 5 \times 10^4 M$ we added to the existing magnetic field an additional poloidal field whose geometry matches that of the initial field, i.e.,

$$B^i = B^i_{O2} + f\, B^i_{O1}\frac{\sqrt{-g}_{O1}}{\sqrt{-g}_{O2}}, \tag{15}$$

where the subscript "O1" ("O2") means the quantity comes from RunSE at $t = 0M$ ($t = 5 \times 10^4 M$). This procedure automatically preserves the solenoidal character of the field, while also minimizing significant transient behavior. Adding a poloidal field is also desirable because, for equal field intensity, the poloidal field leads to more rapid MRI growth than the toroidal field Hawley & Krolik (2002), Beckwith et al. (2008). We set the constant factor $f = 2$ to make the field dynamically significant after an orbital timescale. The ratio of $\sqrt{-g}$ at different times is necessary in our case because $\sqrt{-g}$ is time dependent, and the determinant is included in the covariant form of the magnetic field's constraint equation:

$$\nabla_\mu B^\mu = \frac{1}{\sqrt{-g}}\partial_i(\sqrt{-g}\, B^i) = 0. \tag{16}$$

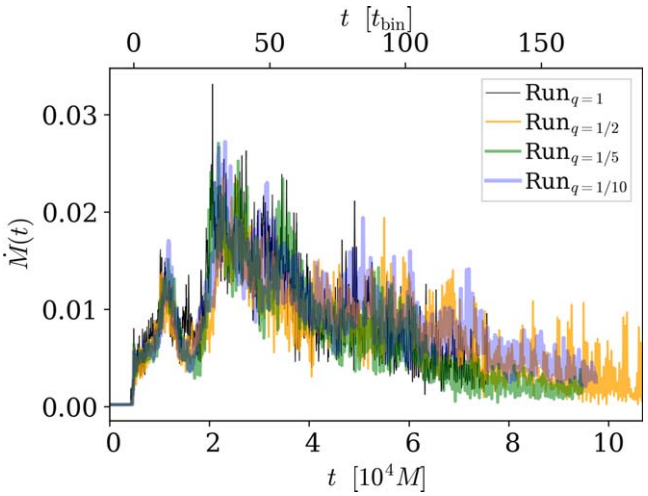

Figure 1. Mass accreted $r = a$ as a function of time, $\dot{M}(r = a, t)$, for the mass ratio series.

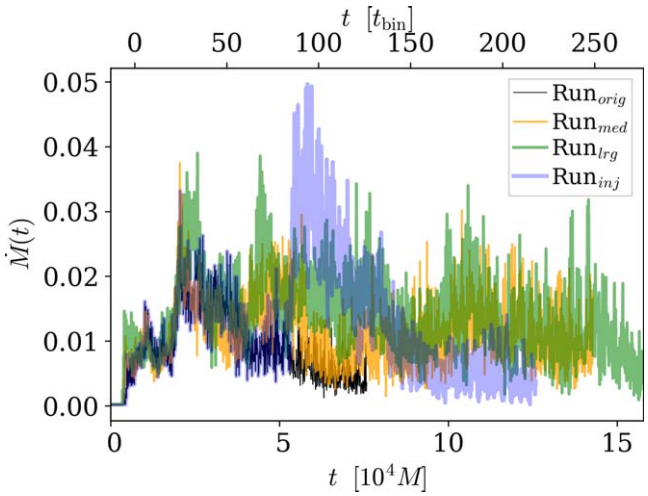

Figure 2. Mass accreted $r = a$ as a function of time, $\dot{M}(r = a, t)$, for the magnetic flux series.

Because the magnetic field in the late-time snapshot of RunSE is turbulent, the magnetic field has a large dynamic range; consequently, the added ordered magnetic field may change the field locally by an amount $\mathcal{O}(1)$. However, the total magnetic field energy added to the system through this procedure is less than 7% of the existing energy. Once the magnetic field is added, the disk is allowed to evolve for an additional $7.5 \times 10^4 M$ in time.

In $\mathrm{Run}_{\rm med}$ and $\mathrm{Run}_{\rm lrg}$ we increased the initial reservoir of magnetic flux available by increasing the size of the hydrostationary torus that encompasses the initial poloidal magnetic field distribution. The initial extent of the magnetic field distribution and its integrated flux content all scale with the size of the disk. Because these disks also have a larger mass reservoir, we expect them to sustain longer periods of accretion than what we observed with the smaller disk (Section 3.1). This longer run time also helps to eliminate a concern that the overdensity develops because diminishing mass accretion at late times also leads to diminishing magnetic flux delivery, and therefore might permit a longer-lived lump. $\mathrm{Run}_{\rm med}$ ($\mathrm{Run}_{\rm lrg}$) included 37% (78%) more magnetic flux and 40% (72%) more mass than RunSE.

3. Axisymmetric Structure

In order to justify our reliance on time averages and make comparisons to steady-state disk theory, we need to evaluate how well our simulations have reached an equilibrium with respect to the accretion of mass. This is important to observables (e.g., electromagnetic luminosity) because the emissivity is proportional to the local rest-mass density (Noble et al. 2010, 2012; d'Ascoli et al. 2018). Further, if the system fluctuates with $\mathcal{O}(1)$ fluctuations on timescales comparable to our simulation's duration, then our results have little predictive power. To assess the degree to which a simulation has entered a state of mass inflow equilibrium we will measure the accretion history, the time evolution of surface density, and the history of the integrated mass at sample radii. These are all most efficiently evaluated using poloidal- and azimuthally integrated quantities, the focus of this section.

3.1. Mass Accretion Rate

The accretion rate history through the spherical surface at $r = a$ is shown in Figures 1 and 2. All runs show a gradual decline of the accretion rate after a period of rapid accretion early on. The peak is due to the burst of inflow that occurs after the linear growth of the MRI saturates. The decline is slow and has only minor fluctuations, allowing us to scale out this slowly varying secular trend when necessary.

Mass ratio does not appear to have a large effect on the accretion history compared to the accretion rate variability at any one instant. The similarity between the curves highlights the fact that the long timescale trend is primarily dictated by the initial conditions of the torus.

The accretion rate's dependence on the initial physical state of the torus is emphasized in the magnetic flux series. Although the accretion rates of all the runs in this group are very similar within the first $3 \times 10^4 M$ of time, at later times they develop larger fluctuations and no longer mimic one another so closely. Nonetheless, all but $\mathrm{Run}_{\rm inj}$ may be fairly described as having a late-time accretion rate that fluctuates within the range $\approx$0.01–0.02, a rate several times greater than the long-term accretion rate for the runs in the mass ratio series. The larger mass reservoirs of $\mathrm{Run}_{\rm med}$ and $\mathrm{Run}_{\rm lrg}$ explain the greater sustained rates of accretion.

The addition of the ordered magnetic field at the run's start, $t = 5 \times 10^4 M$, makes $\mathrm{Run}_{\rm inj}$ different from the others. The burst of accretion seen in $\mathrm{Run}_{\rm inj}$ at $t \approx 5.5 \times 10^4 M$ is triggered by the added ordered magnetic field; the delay is the time required for MRI growth to amplify the MHD turbulence to the saturation level associated with the larger magnetic flux.

A key question to answer in regard to circumbinary accretion is how much of the mass accretion rate in the outer disk can penetrate into the domain of the binary despite the binary's gravitational torque. This quantity, most often called the "leakage fraction," gauges the degree to which the overall system is in inflow equilibrium. Most of the effort on this topic hitherto used hydrodynamics simulations employing a phenomenological "α" viscosity to transfer angular momentum in 2D Eulerian codes (MacFadyen & Milosavljević 2008; D'Orazio et al. 2013; Farris et al. 2014; Muñoz & Lai 2016; Miranda et al. 2017; Duffell et al. 2020; Tiede et al. 2020) or in 3D

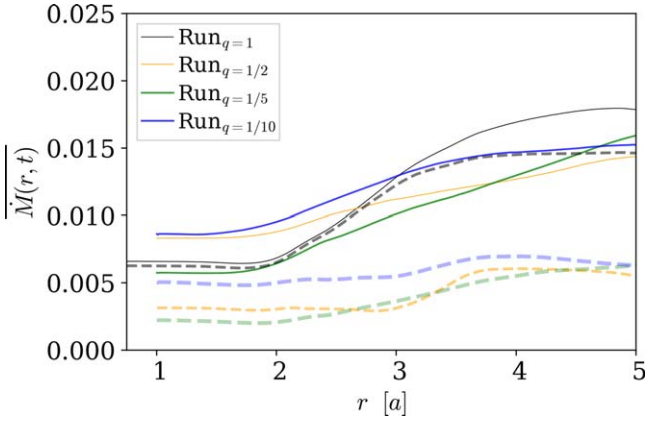

Figure 3. Time average of the mass accretion rate as a function of r for the mass ratio series. The time intervals used for the averages is $40{,}000 < t/M < 76{,}000$ (solid curves), the secularly evolving period of RunSE, and the last $3 \times 10^4 M$ of each run (dashed curves).

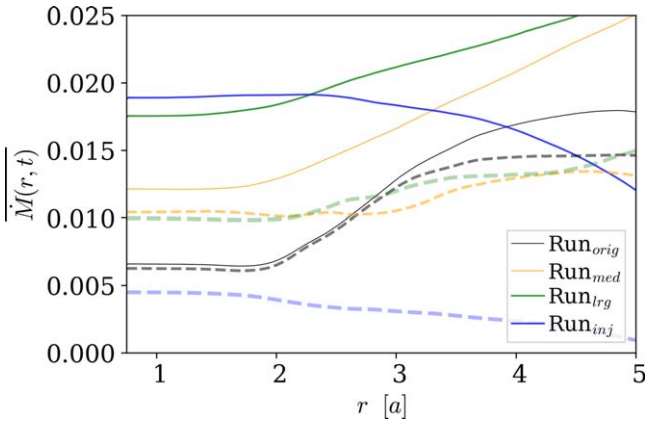

Figure 4. Time average of the mass accretion rate as a function or r for the magnetic flux series. The time intervals used for the averages are $40{,}000 < t/M < 76{,}000$ (solid curves), the secularly evolving period of RunSE, and the last $3 \times 10^4 M$ of each run (dashed curves).

Smoothed Particle Hydrodynamics (SPH) codes (Ragusa et al. 2016; Heath & Nixon 2020); much less attention has been given to inflow dynamics resulting from genuine MHD stresses (Noble et al. 2012; Shi et al. 2012; Gold et al. 2014a; Shi & Krolik 2015; Zilhão et al. 2015). The consensus from both hydrodynamic and MHD work for any explored mass ratio is that the "leakage fraction" is essentially unity; in other words, the system reaches inflow equilibrium.[9] The ratio between the mass accretion rate and the peak surface density near the inner edge may, however, depend on the disk aspect ratio (Tiede et al. 2020).

Figures 3 and 4 portray our results on this question. For all the new runs in the mass ratio series, the extended duration of our simulations led to a significant improvement in the quality of inflow equilibrium. Cases with smaller q generally have smaller departures from inflow equilibrium, but by the end of all the new simulations, the accretion rate became reasonably close to constant as a function of radius out to $\simeq 5a$.

The time-averaged accretion rate as a function of radius is displayed for the magnetic flux series in Figure 4. Again, we see the same flattening of $\overline{\dot{M}(t, r)}$ over time in these runs. Both Run$_{\rm med}$ and Run$_{\rm lrg}$ asymptote to similar profiles at late times.

[9] Only Ragusa et al. (2016) and Heath & Nixon (2020) dissent from this view.

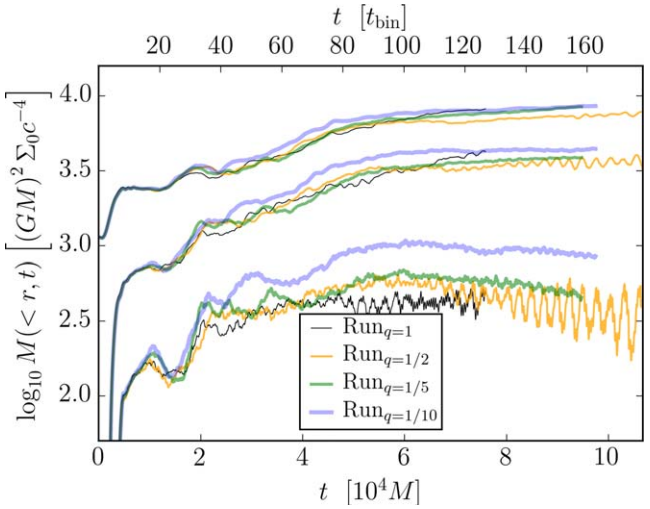

Figure 5. Mass enclosed within a set of radii as a function of time for the mass ratio series. The sample radii are at $r/a = 2$, 3, 4 from bottom to top.

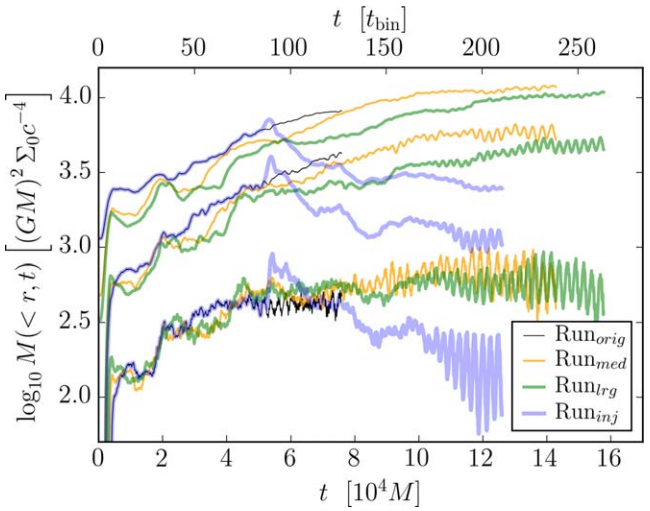

Figure 6. Mass enclosed within a set of radii as a function of time for the magnetic flux series. The sample radii are at $r/a = 2$, 3, 4 from bottom to top.

Because they were both run to longer times, this fact supports the notion that eventually $\overline{\dot{M}(t, r)}$ asymptotes to a flat profile in all cases.

The accretion rate profile from Run$_{\rm inj}$ contrasts strongly with all the others, curving downward with the radius. As we will see in further analyses, the magnetic field perturbation of Run$_{\rm inj}$ leads to a sudden accretion episode that drains a majority of the available mass in the domain. After this happens, the torus in this run no longer has a mass reservoir able to sustain mass inflow equilibrium.

3.2. Enclosed Mass

The next means by which we evaluate mass inflow equilibrium is through the mass enclosed within a given radius versus time, $M(<r, t)$:

$$M(<r, t) \equiv \int_{r_{\rm in}}^{r} dr' \int d\Omega \sqrt{-g}\ \rho. \tag{17}$$

In Figure 5, we plot these trends at sample radii $r/a = 2$, 3, 4 for each mass ratio run. The $M(<r, t)$ curves for disks in

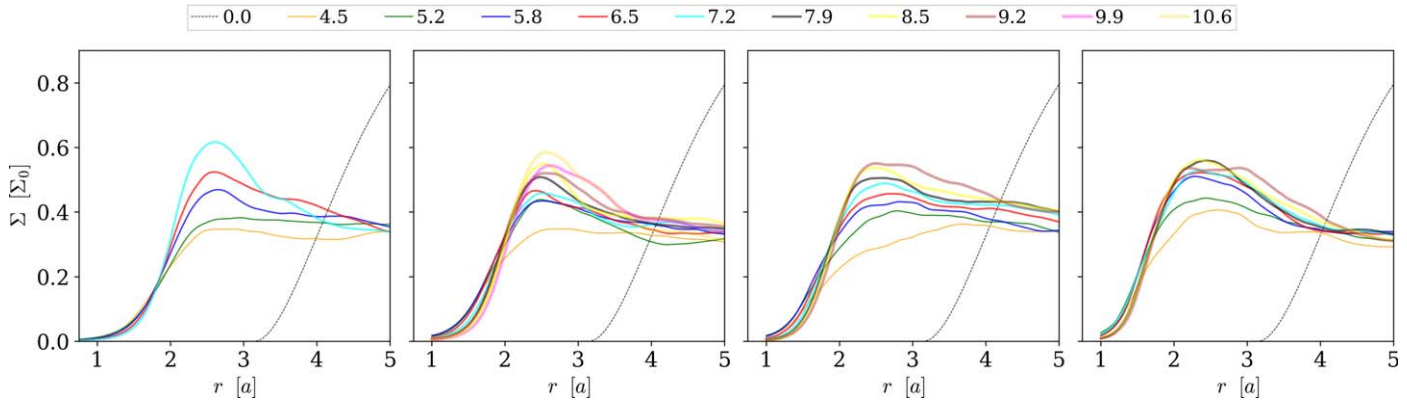

Figure 7. Azimuthally averaged surface density as a function of radius at various times. Radial distance is in units of a, surface density in units of Σ_0. Time averages were performed over windows $\Delta t = 2000M$ in size centered on different times, given in units of 10^4M in the legend. (Left to right) RunSE, $\mathrm{Run}_{q=1/2}$, $\mathrm{Run}_{q=1/5}$, $\mathrm{Run}_{q=1/10}$.

perfect equilibrium at all radii should all be flat. Therefore, deviations from constancy indicate departures from inflow equilibrium. As we found previously for RunSE, the mass within $r = 2a$ hardly changes throughout the post-transient, secularly evolving period (i.e., $t \simeq 5 \times 10^4 M$); this is also true to within 10% for all the other mass ratios. We also find that all the $q < 1$ runs exhibit flatter $M(<r, t)$ trends in time at the largest enclosed radii, with all becoming nearly flat by $t = t_{\rm end}$; we provide the values of $t_{\rm end}$ for all runs in Table 2. This implies that each simulation in the mass ratio series is in approximate mass inflow equilibrium out to these radii. Fluctuations in enclosed mass for all runs are strongest at the smallest radii and grow weaker with decreasing mass ratio, which is expected since the magnitude of the binary's gravitational torque decreases with distance and mass ratio.

The small decline of $M(<r, t)$ from peak to $t_{\rm end}$ decreases in magnitude as the mass ratio decreases, with only $\sim$3% change for $\mathrm{Run}_{q=1/10}$. We also see that $\mathrm{Run}_{q=1/10}$ has significantly more mass enclosed at radii $r/a = 1.5$, 2 for all times. This implies that there is a more massive distribution of steady gas within these smaller annuli. As we will see, this is consistent with the fact $\mathrm{Run}_{q=1/10}$ has a fuller "cavity" region.

These measures taken together suggest that all the mass ratio simulations exhibit a (weakly) secularly evolving steady state of mass flow within $2 < r/a < 4$ for $t \gtrsim 5 \times 10^4 M$.

For the magnetic flux series, the trends in enclosed mass are shown in Figure 6. For the larger radii, $r/a = 3$, 4, the mass-enclosed curves for RunSE rise steadily all the way to $t = t_{\rm end}$; the curves for $\mathrm{Run}_{\rm med}$ and $\mathrm{Run}_{\rm lrg}$ follow closely that of RunSE until its end, but plateau, and to the same level, at $t \simeq 1.2 \times 10^5 M$. In fact, $M(<r, t)$ for $\mathrm{Run}_{\rm med}$ and $\mathrm{Run}_{\rm lrg}$ are also nearly identical at late times at $r/a = 1.5$, 2 suggesting that these runs have likewise achieved steady state at these smaller radii. Because their initial distributions of mass and magnetic flux are quite different, this late-time matching suggests that their mass inflow equilibrium is generic for this thermodynamic model.

The $\mathrm{Run}_{\rm inj}$ $M(<r, t)$ trends are significantly different from those of any other run, however. After the time of injection, we see a dramatic increase in mass followed by a rapid decline for all probed radii $r < 4a$. In contrast, all runs but $\mathrm{Run}_{\rm inj}$ had growing $M(<r, t)$ at $r/a = 3$, 4. Apparently, the magnetic field injection resulted in a sudden redistribution of mass inward that left the entire circumbinary disk with significantly less mass than in the other runs.

3.3. Surface Density Evolution

One of the key differences between circumbinary disks and disks around a single black hole is the very low surface density gap carved out of the accretion flow when there is a binary at its center. The surface density is conventionally defined as the rest-mass density integrated along its "vertical height." For our relatively thin disks aligned with the binary's angular momentum, the vertical integral is approximated well by an integral along the poloidal direction:

$$\Sigma(r, \phi, t) = \frac{\int d\theta \sqrt{-g}\ \rho}{\sqrt{g_{\phi\phi}}|_{\theta=\pi/2}}. \tag{18}$$

Often it is useful to examine azimuthal averages of the surface density, which we calculate as:

$$\Sigma(r, t) = \frac{\{\rho\}}{\int d\phi \sqrt{g_{\phi\phi}}|_{\theta=\pi/2}}. \tag{19}$$

It can be convenient to measure the surface density in units of the initial peak surface density, Σ_0; the values of Σ_0 for all runs are given in Table 2.

As time progresses through each run, in all but RunSE there is a strong convergence in $\Sigma(r)$. Much of this convergence takes place at times earlier than the duration of RunSE, indicating that time steadiness is achieved more rapidly with smaller q. The steadiness of $\Sigma(r)$ at late times in all runs is consistent with the near steady state of enclosed mass shown in Section 3.2.

During this convergence, the shape of $\Sigma(r)$ gradually changes, and the nature of the converged shape is a function of q. We show $\Sigma(r) = \overline{\Sigma(r, t)}(\Delta t = 2000M)$ at a number of times in each run in Figure 7. Several properties show clear trends as the mass ratio decreases. First, the surface density's peak broadens and its contrast with the surface density at a larger radius diminishes. Second, the inner edge of the gap moves inward in terms of r/a, consistent with the decline of the binary's quadrupole moment, which destroys closed orbits. This effect may also be viewed as a weakening of the binary torques, which can repel material outward. Previous Newtonian work in 2D α-viscosity hydrodynamics (D'Orazio et al. 2013; Farris et al. 2015; Miranda et al. 2017), and in MHD

(Shi & Krolik 2015), found a similar trend, while full GRMHD simulations around closer binaries (Gold et al. 2014a) showed little dependence of the circumbinary disk edge position on mass ratio.

The magnetic flux series show a greater variety of behavior and, by this measure, are slower to approach inflow equilibrium (Figure 8). The larger and more distant mass reservoirs of $\mathrm{Run}_{\mathrm{med}}$ and $\mathrm{Run}_{\mathrm{lrg}}$ result in flatter $\Sigma(r)$ profiles at earlier times as mass more slowly redistributes itself to smaller radii. For instance, the local maximum in $\Sigma(r)$ is apparent in $\mathrm{Run}_{\mathrm{med}}$ and $\mathrm{Run}_{\mathrm{lrg}}$ only after $t \gtrsim 8 \times 10^4 M$, and appears to converge to a steady value by $t \simeq 1.1 \times 10^5 M$ in both runs. Also, the surface density's local maximum in $\mathrm{Run}_{\mathrm{med}}$ seems to be significantly narrower and larger compared to its value at larger radii. This last distinction may be the result of the inner portion of $\mathrm{Run}_{\mathrm{lrg}}$'s accretion flow equilibrating with its outer part; the difference between $\Sigma(r = 2.5a)$ and $\Sigma(r = 5a)$ is much larger for $\mathrm{Run}_{\mathrm{med}}$, and $\mathrm{Run}_{\mathrm{lrg}}$'s enclosed mass at $r = 3a$, $4a$ appears to be growing faster than that of $\mathrm{Run}_{\mathrm{med}}$ at late times.

The effect of the magnetic flux injection is apparent in $\Sigma(r, t)$ of $\mathrm{Run}_{\mathrm{inj}}$. The perturbation creates a broad peak in $\Sigma(r)$ just after the time of injection, $t = 5.2 \times 10^4 M$. As the perturbation enhances redistribution of gas and angular momentum, the local maximum's relative amplitude decays over time—as does the absolute magnitude of $\Sigma(r, t)$. Thus, as already noted, this run does not come particularly close to inflow equilibrium.

3.4. Torque Density

Several different mechanisms transport angular momentum within the circumbinary disk. Because angular momentum conservation is broken by nonaxisymmetry in the gravitational spacetime, we discuss this issue here.

The total angular momentum J is the integral over the spatial volume of the time component of its associated current, j^μ: $J = \int j^t \sqrt{-g}\, dV$, where dV is the spatial volume component in the spacelike hypersurface (e.g., $dr d\theta d\phi$). We are interested in the azimuthal component of the momentum, so the desired current is $j^\mu = T^\mu{}_\nu \phi^\nu$, and $\phi^\nu = (\partial_\phi)^\nu = \partial x^\nu / \partial \phi = [0, 0, 0, 1]$ in spherical coordinates, which is what we use.

Radial transport of angular momentum can be traced through examination of the several mechanisms contributing to the local rate of change of angular momentum density, $d^2J/dtdr$:

$$\begin{aligned} \partial_r \partial_t J &= dT/dr - \{\mathcal{F}_\phi\} - \partial_r \{T^r{}_\phi\} \\ &= dT/dr - \{\mathcal{F}_\phi\} - \partial_r \{M^r{}_\phi\} \\ &\quad - \partial_r \{R^r{}_\phi\} - \partial_r \{A^r{}_\phi\}. \end{aligned} \tag{20}$$

The radial density of gravitational torque is

$$\frac{dT}{dr} = \frac{1}{2} \int T^{\mu\nu} (\partial_\phi g_{\mu\nu}) \sqrt{-g}\, d\theta\, d\phi. \tag{21}$$

The quantities $M^r{}_\phi$, $R^r{}_\phi$, and $A^r{}_\phi$ are—respectively—the Maxwell (MHD) stress, Reynolds stress, and advected flux of angular momentum. The Maxwell stress $M^\mu{}_\nu = 2p_m u^\mu u_\nu + p_m \delta^\mu{}_\nu - b^\mu b_\nu$ is the EM part of the stress-energy tensor, while the Reynolds stress and the advected angular momentum flux sum to the hydrodynamic part: $(R^\mu{}_\nu + A^\mu{}_\nu) = T^\mu_{H\nu} = \rho h u^\mu u_\nu + p \delta^\mu{}_\nu$.

The quantities $\{R^\mu{}_\nu\}$ and $\{A^\mu{}_\nu\}$ can be separated by defining the advected flux in terms of the mean radial flow and then subtracting it from the total hydrodynamic angular momentum flux:

$$\{A^r{}_\phi\} \simeq \frac{\{\rho \ell\}\{\rho h u^r\}}{\{\rho\}}. \tag{22}$$

$$\{R^r{}_\phi\} = \{\rho h\, \delta u^r\, \delta u_\phi\} \simeq \{T^r_{H\phi}\} - \{A^r{}_\phi\}. \tag{23}$$

Here $\ell \equiv -u_\phi / u_t$.

We show each of these contributions separately in Figure 9. The contributions are displayed averaged over two different epochs: an earlier period $(4 \times 10^4 < t/M < 7.6 \times 10^4)$ to compare with the final part of RunSE, and the last $3 \times 10^4 M$ of each run in order to illustrate how the system evolves as it nears inflow equilibrium.

Several qualitative conclusions can be drawn from these figures. First, all the runs of the mass ratio series reach a steady state with respect to angular momentum transport at $r \gtrsim 3a$ by the time of the later period, evidenced by the total angular momentum gradient lying close to zero. The magnitudes of all the contributions diminish slightly in time, with the largest decrease in the advected and magnetic contributions.

Second, there is only weak dependence on mass ratio. The peak of the gravitational torque density, dT/dr, moves inward as q decreases. With only a few exceptions where both are small, the radial gradient of the Reynolds stress contributes to the total torque so as to cancel the gravitational torque; in other words, when gravitational torque adds angular momentum to the fluid in a grid cell, Reynolds stress carries it away.

This figure also illustrates the transition from linear gravitational torques to nonlinear. At late times (bottom row of this figure), as q rises from 0.1 to 1, the damped sinusoidal oscillations in both gravitational torque density and Reynolds stress as functions of radius flatten out into low-amplitude plateaus at the third extremum. This plateau feature resides at the location of the lump and demonstrates that the response of the circumbinary disk matter to external torques can no longer be described by linear perturbation theory when $q \gtrsim 1/4$.

On the other hand, the Maxwell stress is consistently close to uniform spatially for all q; the magnitude of this spatially uniform stress is almost the same for all $q \leqslant 0.5$. The degree to which the Maxwell stress maintains the same constant value over all mass ratios suggests that its magnitude reflects the asymptotic behavior of the MRI. As shown in Noble et al. (2012) for RunSE, and found in the other runs but not shown here, the plasma $\beta = p/p_m$ exhibits a local maximum at $r \simeq 2.5a$; this maximum is more pronounced for larger q. The two trends together imply that larger q leads to higher pressure at the radius of the surface density maximum. This greater pressure may be the end result of the stronger gravitational torques associated with higher q doing more mechanical work, and the eventual dissipation of this work into heat, rather than a loss of magnetic field intensity.

The angular momentum budget for the magnetic flux series is shown in Figure 10. Like the mass ratio series, the fact that $d^2J/dtdr$ settles toward the zero line at late times gives strong evidence of an approach to a steady state. There is also a resemblance to the mass ratio series in the sense that, like RunSE and $\mathrm{Run}_{q=1/2}$, the magnetic flux series show an absence of the third peak in the Reynolds stress at late times.

However, there is more contrast between these runs than those in the mass ratio series. During the earlier averaging period, $d^2J/dtdr$ is far from zero throughout the circumbinary disk in RunSE, $\mathrm{Run}_{\mathrm{med}}$, and $\mathrm{Run}_{\mathrm{lrg}}$, indicating that this is a transient

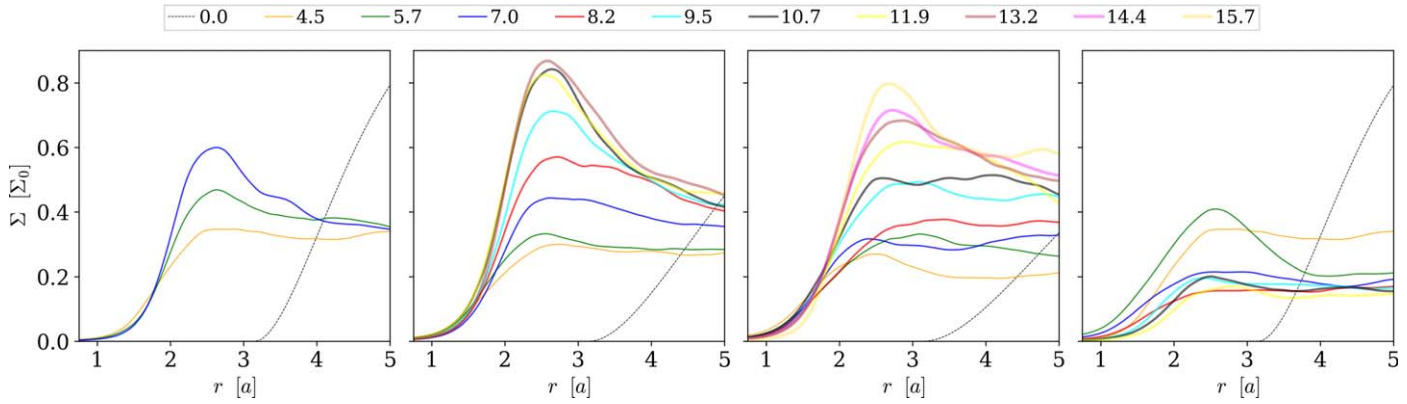

Figure 8. Temporal and azimuthal averaged surface density as a function of radius in units of a. Time averages are performed over windows $\Delta t = 2000M$ in size centered on different times, specified in units of 10^4M in each legend. (Left to right) RunSE, $\mathrm{Run}_{\mathrm{med}}$, $\mathrm{Run}_{\mathrm{lrg}}$, $\mathrm{Run}_{\mathrm{inj}}$.

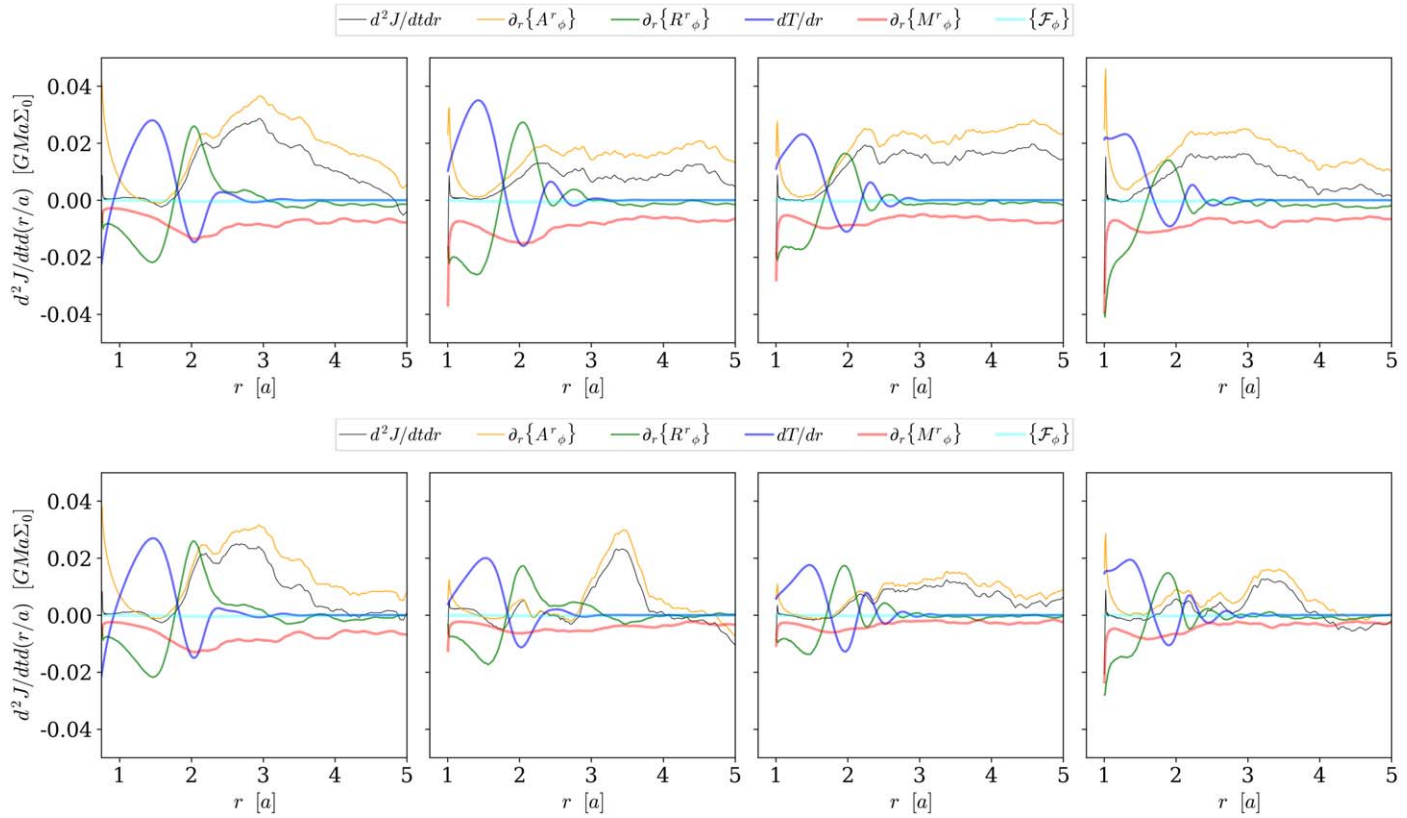

Figure 9. Contributions to the time-averaged radial distribution of $\partial_r\partial_t J$ (black) in the mass ratio series. Shown are the radial derivatives of the Maxwell stress in the Eulerian frame ($\{M^r{}_\phi\}$, red), the angular momentum flux due to shell-integrated Reynolds stress in the Eulerian frame ($\{R^r{}_\phi\}$, green), and advected angular momentum ($\{A^r{}_\phi\}$, gold). Also shown are the torque densities per unit radius due to the actual binary spacetime (dT/dr, blue) and radiative cooling source term ($\{\mathcal{F}_\phi\}$, cyan). The net rate of change of angular momentum $\partial_r\partial_t J$ (solid black). All quantities in the top (bottom) row plots are time-averaged over $40{,}000 < t/M < 76{,}000$ (last $30{,}000M$ of evolution per run). (Left to right) RunSE, $\mathrm{Run}_{q=1/2}$, $\mathrm{Run}_{q=1/5}$, $\mathrm{Run}_{q=1/10}$. Note that $\partial_r\{M^r{}_\phi\}$, $\partial_r\{R^r{}_\phi\}$, $\partial_r\{A^r{}_\phi\}$, and $\{\mathcal{F}_\phi\}$ have all been multiplied by a factor of -1 to match the sign they have in Equation (20) so their curves add up to that of $\partial_r\partial_t J$.

phase in mass/magnetic flux redistribution for all three. In $\mathrm{Run}_{\mathrm{inj}}$, $d^2J/dtdr \simeq 0$ for $r \lesssim 3.5a$, but grows rapidly at larger radii, showing that this run reached a steady state in its inner regions more rapidly than the others, but in this time span is evolving rapidly at larger radii. At later times, all three new runs come much closer to equilibrium in their angular momentum evolution.

Another way to evaluate each run's proximity to a steady state w.r.t. the angular momentum flux distribution is to plot the radial-integrated $d^2J/dtdr$ as is commonly done in Newtonian viscous hydrodynamics studies (e.g., Miranda et al. 2017). We perform this analysis in Appendix B.

To close this section on axisymmetric properties, we remark on how the nonaxisymmetric lump can influence azimuthally averaged properties such as the vertically integrated magnetic stress and the MRI quality factors. The largest value of the former over the entire radial extent of the circumbinary disk is found at the radial location of the lump, even though the minima for the latter are found at the (r, ϕ) locations of the lump (see Appendix C). To explain this diminution in MRI quality, we point out that the magnetic stresses of RunSE, $\mathrm{Run}_{\mathrm{med}}$, and $\mathrm{Run}_{\mathrm{lrg}}$ all agree at $r = 5a$, suggesting that the variations between those runs neither strengthen nor weaken

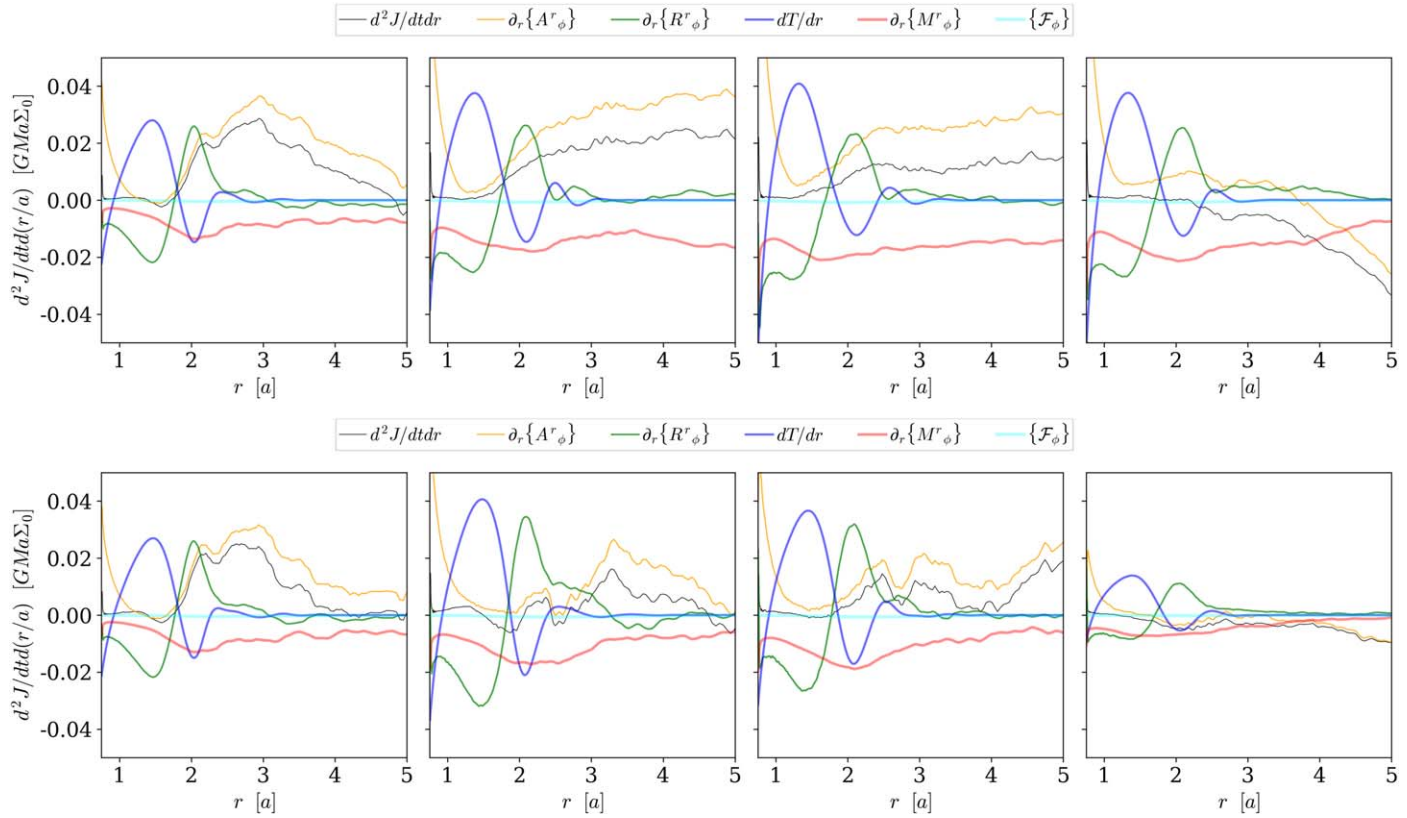

Figure 10. Contributions to the time average radial distribution of $\partial_r\partial_t J$ (black) in the magnetic flux series. Shown are the radial derivatives of the Maxwell stress in the Eulerian frame ($\{M^r{}_\phi\}$, red), the angular momentum flux due to shell-integrated Reynolds stress in the Eulerian frame ($\{R^r{}_\phi\}$, green), and advected angular momentum ($\{A^r{}_\phi\}$, gold). Also shown are the torque densities per unit radius due to the actual binary potential (dT/dr, blue) and radiative cooling source term ($\{\mathcal{F}_\phi\}$, cyan). The net rate of change of angular momentum $\partial_r\partial_t J$ (solid black). All quantities in the top (bottom) row plots are time-averaged over $40{,}000 < t/M$ 76,000 (last 30,000M of evolution per run). (Left to right) RunSE, $\mathrm{Run_{med}}$, $\mathrm{Run_{lrg}}$, $\mathrm{Run_{inj}}$. Note that $\partial_r\{M^r{}_\phi\}$, $\partial_r\{R^r{}_\phi\}$, $\partial_r\{A^r{}_\phi\}$, and $\{\mathcal{F}_\phi\}$ have all been multiplied by a factor of -1 to match the sign they have in Equation (20) so their curves add up to that of $\partial_r\partial_t J$.

the field in the outer disk. Nonetheless, in $\mathrm{Run_{med}}$ and $\mathrm{Run_{lrg}}$, the stresses at $r \simeq 2a$, i.e., the lump region, are even larger than in RunSE. This fact suggests that most of the degradation in MRI quality in these runs must be due to increased density in the lump region.

In order to explore how magnetic stress may influence lump dynamics and evolution, it is useful to define a measure of the magnitude of the magnetic stress per unit mass, which we will call $W^r{}_\phi$ following Balbus & Hawley (1998):

$$W^r{}_\phi = \frac{\{M^r{}_\phi\}}{\{\rho\}}. \tag{24}$$

This quantity for the mass ratio series and the magnetic flux series is shown in Figures 11 and 12, respectively. In every run of the mass ratio series, $W^r{}_\phi$ at radii $r \gtrsim 2a$ drops abruptly by about a factor of 4 at a time $\approx$40,000M. Particularly for low q, this drop begins at a large radius and only then extends inward. The evolution of $W^r{}_\phi$ in the magnetic flux series is very different because we deliberately manipulated the magnetic flux available.

For those runs with a lump, we find that once the specific magnetic stress drops to $W^r{}_\phi \lesssim 10^{-4}$ the lump appears when one uses the criteria described in Section 4. The significance of this value will be discussed in Section 5.1.

4. Nonaxisymmetric Structure

4.1. Lump Amplitude

Although Figures 7 and 8 illustrate well the azimuthally averaged pileup of material at the edge of the gap for all runs, they lack information about the nonaxisymmetric structure. In particular, they say nothing about the lump feature (MacFadyen & Milosavljević 2008; Shi et al. 2012; Noble et al. 2012), which can affect the electromagnetic signal both by the dissipation associated with it directly and by its modulation of the accretion rate. In order to investigate the nonaxisymmetric structure of the flow, in Figures 13 and 14 we plot the distribution of the surface density, $\Sigma(r, \phi)$ at evenly spaced intervals over the secularly evolving period of the runs.

In all cases with a lump, we find that the overdensity region spans $\pi/3 \lesssim \delta\phi_{\mathrm{lump}} \lesssim \pi$ in azimuthal angle, and a radial extent of $0.1 \lesssim \delta r_{\mathrm{lump}} \lesssim a$. In many cases, the density contrast between the lump and its surroundings is quite large.

In the mass ratio series, the development of an azimuthally asymmetric overdensity is obvious only in runs RunSE and $\mathrm{Run}_{q=1/2}$, with the latter run showing weaker development at all the times shown in Figure 13. The lump does not appear at any time in either $\mathrm{Run}_{q=1/5}$ or $\mathrm{Run}_{q=1/10}$.

On the other hand, all four runs in the magnetic flux series show lumps in at least one of the snapshots shown. The lumps develop more slowly in tori extending to larger radius such as

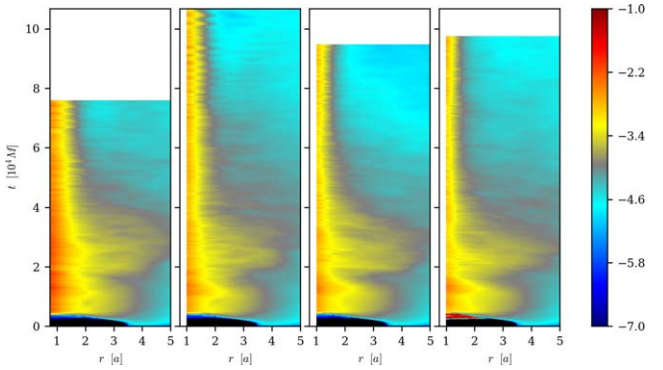

Figure 11. Log_{10} of the ratio of shell-averaged Maxwell stress, $M^r{}_\phi$, to shell-averaged mass density, $\langle\rho\rangle$, vs. radius and time, shown in units of binary separation and total BH mass, respectively. The scale is shown in the color bar. The ranges of time and radius used in the plots cover the full extent of each simulation. (Left to right) RunSE, $\mathrm{Run}_{q=1/2}$, $\mathrm{Run}_{q=1/5}$, $\mathrm{Run}_{q=1/10}$.

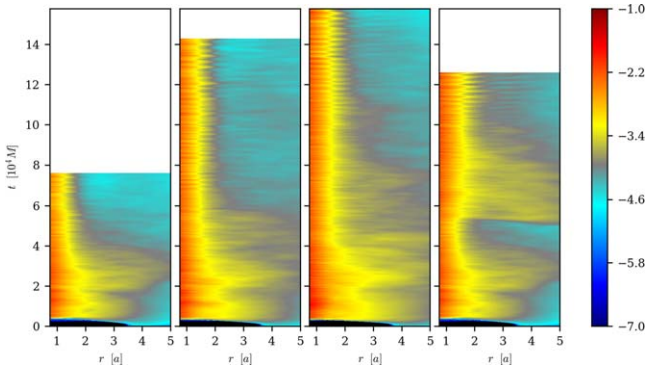

Figure 12. Log_{10} of the ratio of shell-averaged Maxwell stress, $M^r{}_\phi$, to shell-averaged mass density, $\langle\rho\rangle$, vs. radius and time, shown in units of binary separation and total BH mass, respectively. The scale is shown in the color bar. The ranges of time and radius used in the plots cover the full extent of each simulation. (Left to right) RunSE, $\mathrm{Run}_{\mathrm{med}}$, $\mathrm{Run}_{\mathrm{lrg}}$, $\mathrm{Run}_{\mathrm{inj}}$.

$\mathrm{Run}_{\mathrm{med}}$ and $\mathrm{Run}_{\mathrm{lrg}}$. When their lumps form, however, the peak surface density in each is significantly greater than in RunSE. The impact of additional mass supply is evident.

The images of $\mathrm{Run}_{\mathrm{inj}}$ look different than the others, but this is a visual artifact of the burst of accretion triggered by the injected magnetic field. As a result, substantially less mass remains in the disk. Nonetheless, the contrast between the surface density of the lump and the azimuthally averaged surface density at late times is comparable to that in the other runs (see Figure 15 for a clearer view of this contrast). Note that at $t = 5.8 \times 10^4 M$ one may still see the remnant of the $m = 1$ lump structure created in RunSE before the extra magnetic flux was injected in $\mathrm{Run}_{\mathrm{inj}}$.

To quantify the surface density contrast between the lump and its surroundings, we compute the Fourier transform $\mathcal{B}_m$ of θ-integrated $\rho\sqrt{-g}$ with respect to ϕ in the coordinate frame. We call its absolute magnitude A_m, with the definitions

$$A_m(r, t) = |\mathcal{B}_m(r, t)|, \quad \mathcal{B}_m(r, t) = \{\rho e^{im\phi}\}. \tag{25}$$

The phase of the Fourier modes indicates the azimuthal location of the lump:

$$\varphi_m(r, t) = \mathtt{ArcTan}(-\mathrm{Im}(\mathcal{B}_m(r, t)), \mathrm{Re}(\mathcal{B}_m(r, t))) \tag{26}$$

where our $\mathtt{ArcTan}(y, x)$ function returns the angle between the $y = 0$, $x > 0$ line and the line connecting the point (x, y) and the origin.

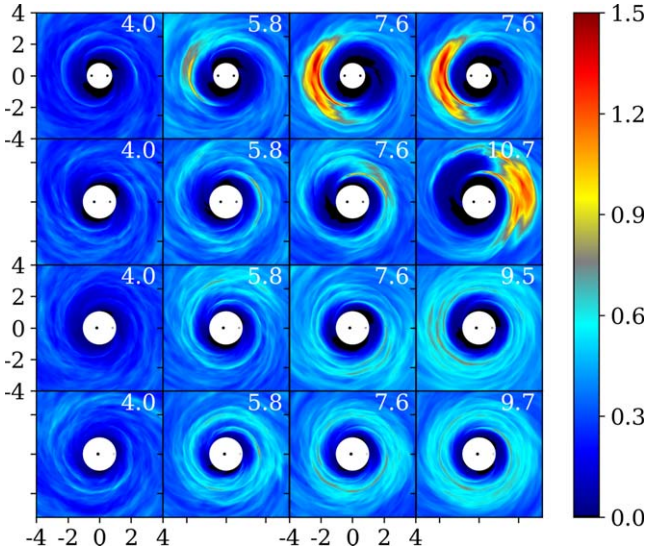

Figure 13. Color contours of surface density in units of Σ_0 as a function of radius and azimuthal angle, i.e., $\Sigma(r, \phi)/\Sigma_0$, at four different times indicated by the number in the upper-right corner of each frame in units of $10^4 M$. The first three times were chosen to span RunSE's secularly evolving state; the time of the right-most column is the last time in each run. (Top to bottom) RunSE, $\mathrm{Run}_{q=1/2}$, $\mathrm{Run}_{q=1/5}$, $\mathrm{Run}_{q=1/10}$. (Right) the linear color scale used in all frames.

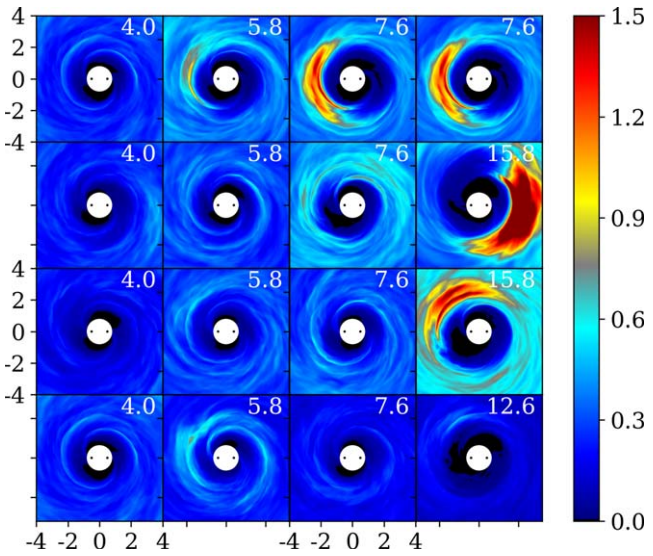

Figure 14. Color contours of surface density in units of Σ_0 as a function of radius and azimuthal angle, i.e., $\Sigma(r, \phi)/\Sigma_0$, at four different times indicated by the number in the upper-right corner of each frame in units of $10^4 M$. The first three times were chosen to span RunSE's secularly evolving state; the time of the right-most column is the last time in each run. (Top to bottom) RunSE, $\mathrm{Run}_{\mathrm{med}}$, $\mathrm{Run}_{\mathrm{lrg}}$, $\mathrm{Run}_{\mathrm{inj}}$. (Right) the linear color scale used in all frames.

As shown in Figure 16, $A_1(r, t)/A_0(r, t)$ for RunSE and $\mathrm{Run}_{q=1/2}$ increases substantially over time. The $m = 1$ mode amplitude is strongest in the accretion stream region but grows significantly in the region of the lump proper, $r \approx 2.5a$, as well. That both the inner region of the circumbinary disk and the stream region develop the same sort of asymmetry is no coincidence. If there were no disk asymmetry, the stream region would be modulated strongly for $m = 2$, not $m = 1$; that the streams also have $m = 1$ character is a sign that the accretion streams originate in the lump. The amplitude of the $m = 1$ mode in the gap is larger than in the disk because the

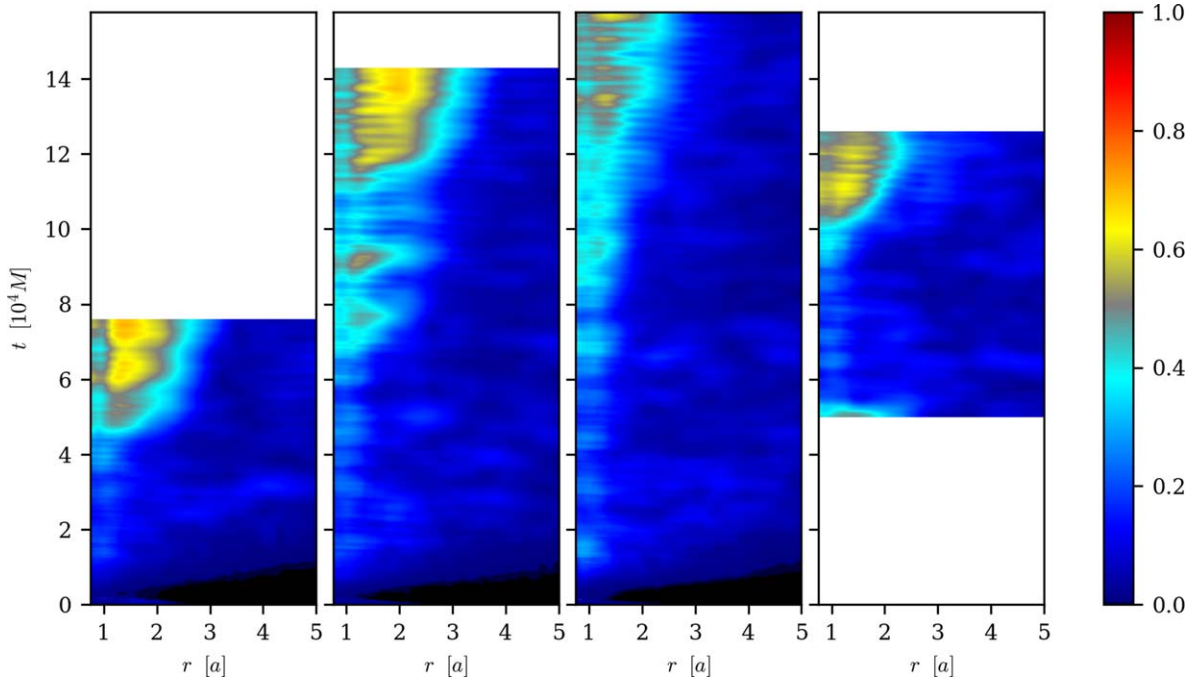

Figure 15. The ratio of the $m = 1$ mode to the $m = 0$, $\overline{A_1}(r, t, \Delta t)/\overline{A_0}(r, t, \Delta t)$, as a function of radius and time for the magnetic flux series. Here the smoothing period slides along t and has duration $\Delta t = 2t_{\rm lump}$. (Left to right) RunSE, $\rm Run_{med}$, $\rm Run_{lrg}$, $\rm Run_{inj}$.

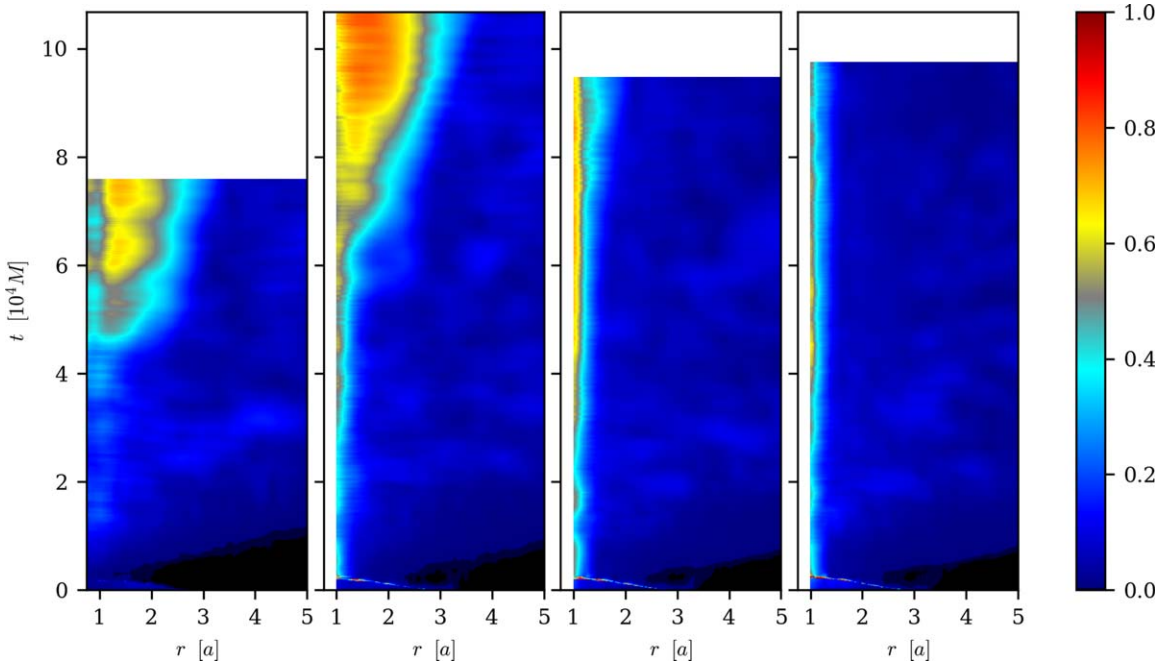

Figure 16. The ratio of the $m = 1$ mode to the $m = 0$, $\overline{A_1}(r, t, \Delta t)/\overline{A_0}(r, t, \Delta t)$, as a function of radius and time for the mass ratio series. Here the smoothing period slides along t and has duration $\Delta t = 2t_{\rm lump}$. (Left to right) RunSE, $\rm Run_{q=1/2}$, $\rm Run_{q=1/5}$, $\rm Run_{q=1/10}$.

streams grow narrower and denser as they fall toward the nearest black hole and the remainder of the gap has such low density. By contrast, $\rm Run_{q=1/5}$ and $\rm Run_{q=1/10}$ show almost no signs of growth in the relative amplitude of the $m = 1$ mode.

All the magnetic flux series runs show enhancements of A_1/A_0 similar to those seen in RunSE and $\rm Run_{q=1/2}$, but at rather later times (the brief appearance of significant A_1/A_0 at the beginning of $\rm Run_{inj}$ is the remnant of the lump in RunSE as it is destroyed by the injection of magnetic flux).

Interpreting a large relative mode strength of the $m = 1$ mode as a signature of the lump, we determine the onset of the lump using the criterion

$$C_{\rm lump}(t) = \frac{\int_{2a}^{3a} \overline{A_1(r, t)}(2t_{\rm lump})\, dr}{\int_{2a}^{3a} \overline{A_0(r, t)}(2t_{\rm lump})\, dr} > 0.2. \quad (27)$$

We plot $C_{\rm lump}(t)$ for each run in both series, along with the $m = 0$, 1, 2 mode amplitudes integrated over $r \in [2a, 3a]$, in

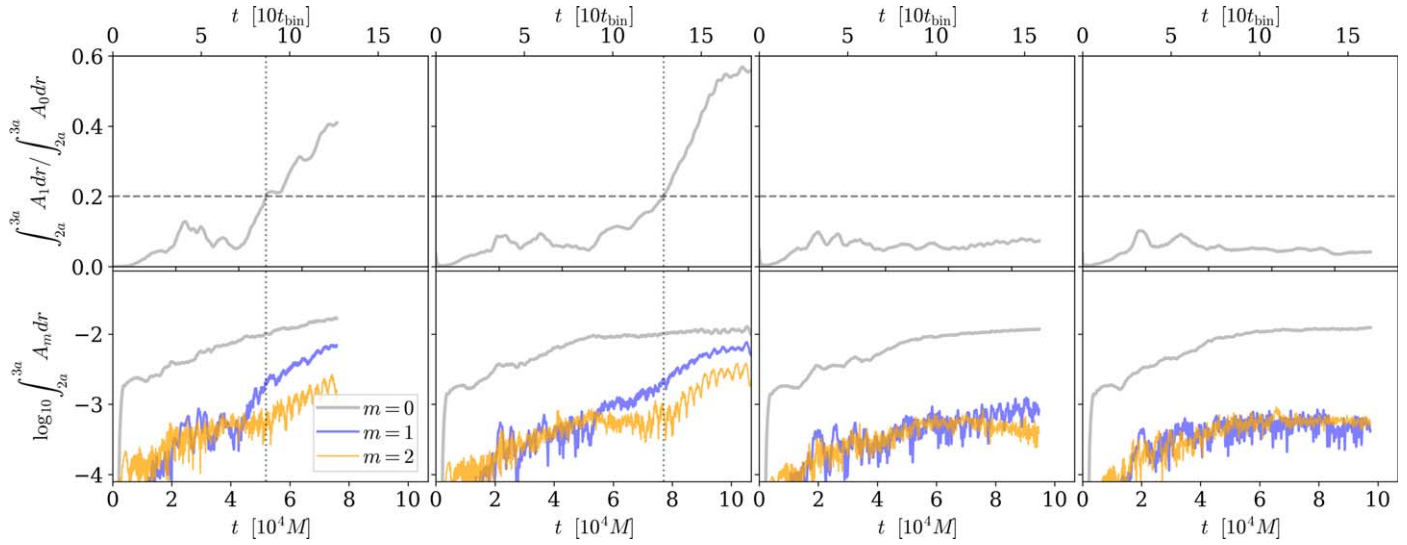

Figure 17. Lump criterion, $C_{\rm lump}(t)$, (top rows) used to determine $T_{\rm lump}$, and the $m = 0$, 1, 2 mode amplitudes integrated over the lump region (bottom rows) for the mass ratio simulations. The horizontal dashed line indicates the threshold $C_{\rm lump} = 0.2$ above which we recognize the presence of an overdensity; the vertical lines denote the time, $T_{\rm lump}$, at which it first satisfies this criterion. (Left to right) RunSE, $\rm Run_{q=1/2}$, $\rm Run_{q=1/5}$, $\rm Run_{q=1/10}$.

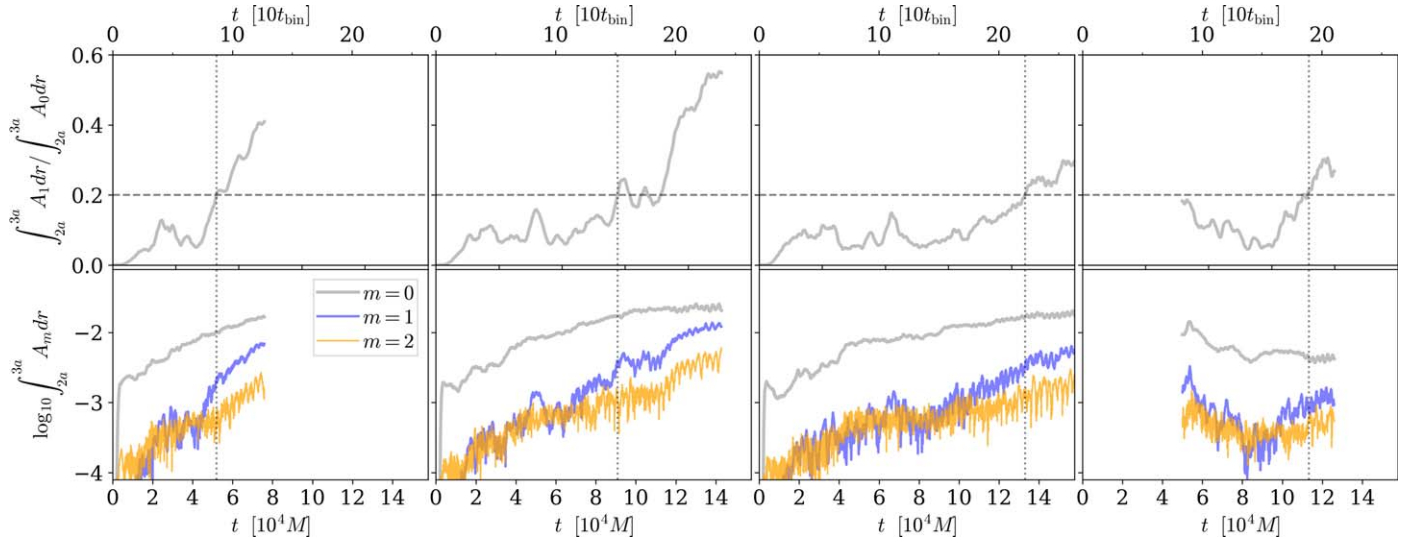

Figure 18. Lump criterion, $C_{\rm lump}(t)$, (top rows) used to determine $T_{\rm lump}$, and the $m = 0$, 1, 2 mode amplitudes integrated over the lump region (bottom rows). The horizontal dashed line indicates the threshold $C_{\rm lump} = 0.2$ above which we recognize the presence of an overdensity; the vertical lines denote the time, $T_{\rm lump}$, at which it first satisfies this criterion. (Left to right) RunSE, $\rm Run_{med}$, $\rm Run_{lrg}$, $\rm Run_{inj}$.

Figures 17 and 18. Satisfaction of the criterion coincides with the time when the amplitude of the $m = 1$ mode begins to rise above that of the $m = 2$ mode and approaches the $m = 0$ mode strength. We define $T_{\rm lump}$ to be the time the lump criterion is first satisfied. The measured values of $T_{\rm lump}$ for all lump-forming runs are presented in Table 2.

Similar behavior is seen in the magnetic flux series. We find that each run in this series satisfies the criterion within its duration. Consistent with the surface density plots, the lump forms later for $\rm Run_{med}$ and $\rm Run_{lrg}$. The $m = 1$ amplitude of $\rm Run_{med}$ reaches more than half that of its $m = 0$ mode, while in $\rm Run_{lrg}$ this ratio crosses the threshold of 0.2 near the end of the simulation and reaches 0.3 at the very end. In $\rm Run_{inj}$, the run begins with the decay of the existing lump from RunSE, but the lump recovers and crosses the threshold later.

An important characteristic of the lump is its phase coherence. In order to quantify the instantaneous phase of the lump, $\varphi_{\rm lump}(t)$, and its associated orbital frequency, $\omega_{\rm lump}(t)$, we rely on our means of calculating the amplitude of the overdensity Equation (25). As we want to track the lump, which resides close to the cavity's edge, we first integrate the density's $m = 1$ Fourier amplitude over the radial extent of the lump:

$$\tilde{\mathcal{B}}_1(t) \equiv \int_{2a}^{4a} \mathcal{B}_1(r, \phi, t)\, dr \tag{28}$$

which is then immediately used to find $\varphi_{\rm lump}(t)$:

$$\varphi_{\rm lump}(t) \equiv \texttt{ArcTan}(-\mathrm{Im}(\tilde{\mathcal{B}}_1(t)), \mathrm{Re}(\tilde{\mathcal{B}}_1(t))) \quad . \tag{29}$$

The instantaneous orbital frequency of the lump, $\omega_{\rm lump}(t)$, is simply the time derivative of the phase:

$$\omega_{\rm lump}(t) = \frac{d\varphi_{\rm lump}(t)}{dt}. \tag{30}$$

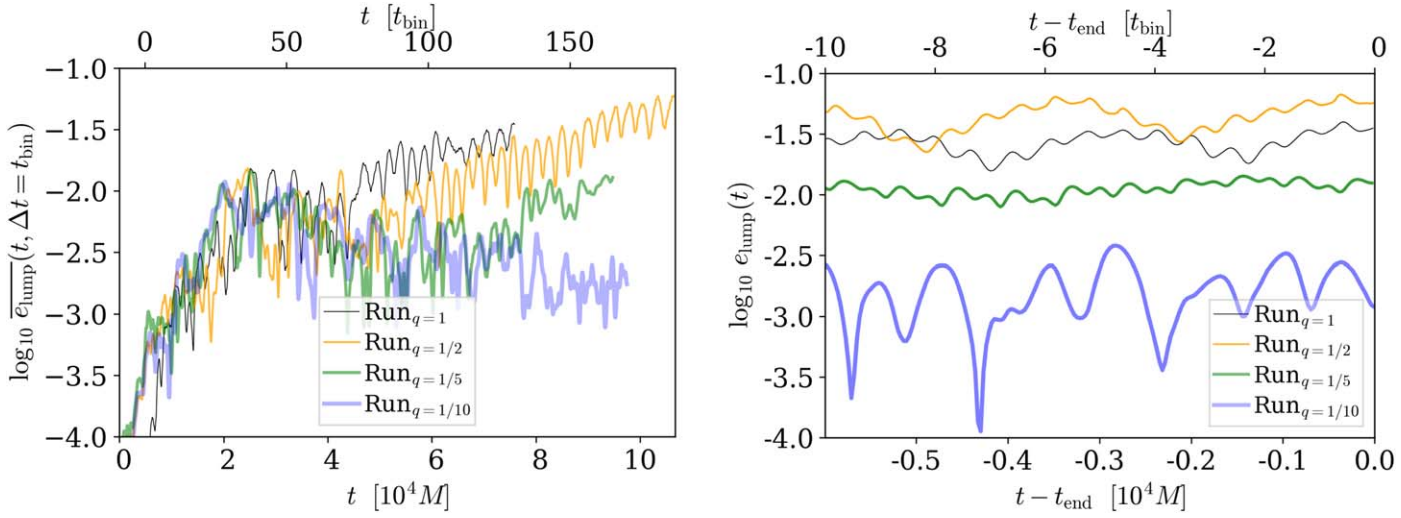

Figure 19. The lump's eccentricity smoothed over $\Delta t = t_{\rm bin}$, $\overline{e_{\rm lump}}(t, t_{\rm bin})$ (left); $e_{\rm lump}(t)$ over the last 10 orbits of each simulation (right); $t_{\rm end}$ represents the final time of each simulation of the mass ratio series.

When referenced without an argument, $\omega_{\rm lump}$ is to be interpreted as the time average of $\omega_{\rm lump}(t)$ over the period $T_{\rm lump} < t < t_{\rm end}$ for the specified run. Table 2 shows the values of $\omega_{\rm lump}$ for each run.

4.2. Eccentricity

Even if the circumbinary gas begins the simulation on circular orbits, previous investigations reported that it acquires nontrivial levels of eccentricity over time (MacFadyen & Milosavljević 2008; Shi et al. 2012; Farris et al. 2014; D'Orazio et al. 2016; Miranda et al. 2017). At late times, the azimuthally averaged eccentricity rises sharply just inside the cavity wall and decays exponentially outward $\approx \exp(-r/a)$ (MacFadyen & Milosavljević 2008; Shi et al. 2012). Previous MHD simulations reported smaller eccentricities than viscous hydrodynamics simulations, but it is unclear if the MHD simulations reached a true steady state in terms of eccentricity. We explore here how our simulations compare with previous work and how eccentricity is associated with the growth and strength of the lump.

We define the eccentricity of a disk annulus at radius r in a way analogous to the Newtonian expressions used in MacFadyen & Milosavljević (2008) and Shi et al. (2012), but expressed in terms of the 4-velocity and the metric:

$$e(r, t) = \frac{|\{\rho u^r e^{i\phi}\}|}{\{\rho r u^{\phi}\}}. \tag{31}$$

We further define the quantity $e_{\rm lump}$ to be the eccentricity of the region $2a < r < 4a$:

$$e_{\rm lump}(t) = \frac{\left|\int_{2a}^{4a} dr\,\{\rho u^r e^{i\phi}\}\right|}{\int_{2a}^{4a} dr\,\{\rho r u^{\phi}\}}. \tag{32}$$

In order to accentuate variability or trends occurring at longer timescales, we display $e_{\rm lump}(t)$ in Figures 19 and 20 smoothed over a binary orbit: $\overline{e_{\rm lump}}(t, t_{\rm bin})$, calculated using Equation (13).

As shown in Figures 19 and 20, in all cases the inner disk eccentricity grows exponentially during the early development of the circumbinary disk. However, once it reaches $\sim 10^{-2}$, further growth is a function of mass ratio and magnetic flux.

For fixed magnetic properties, but varying mass ratio, the eccentricity decreases slightly during the first $\sim 10^4 M$ after rising to $\sim 10^{-2}$. When the mass ratio takes its maximum value, i.e., $q = 1$, the eccentricity then renews its exponential growth, but at a slower rate. For smaller values of q, the growth is delayed longer, and therefore begins from a lower level. However, once begun, growth proceeds at roughly the same rate for all values of q. There is, however, one possible exception: for $q = 0.1$, the slow decline in eccentricity runs all the way to the end of the simulation. We cannot say whether the eccentricity might begin growing at still later times.

On the other hand, for fixed mass ratio and varying magnetic properties, the development of eccentricity is very similar in all the runs of this series. The smoothed form of the eccentricity history, $\overline{e_{\rm lump}}(t, t_{\rm bin})$, ultimately reaches the same value in each run to within $\sim$20%. The only significant contrast between them occurs in Run$_{\rm inj}$, where the eccentricity drops sharply when the additional magnetic flux is added and the lump temporarily dissolves. When the lump returns, this run, too, returns to the common path.

The range of eccentricities we find all lie within the range of eccentricities observed in another GRMHD series (Lopez Armengol et al. 2021) and about 50% smaller than that reported in a Newtonian MHD study (Shi et al. 2012). Our results are also in agreement with 2D VH results (D'Orazio et al. 2016), and some 3D SPH studies show a similar trend with q (Ragusa et al. 2020).

Generally speaking, in all the runs that exhibit a lump as determined by our lump criterion, $\overline{e_{\rm lump}}(t, t_{\rm bin})$, grows as the $m = 1$ density mode amplitude, A_1, grows. In both series, in each run exhibiting a lump, $\overline{e_{\rm lump}}(t, t_{\rm bin})$ begins its second period of exponential growth approximately $10^4 M$ before satisfying the lump criterion. Moreover, the radial profiles are also quite similar to one another: the time-averaged radial eccentricity profile for all cases declines exponentially with increasing radius from $r \approx 1$ to $r \approx 3$ with an e-folding scale consistently $\approx a$.

The timescales of variability seen in $e_{\rm lump}(t)$ for each run of the magnetic flux series are also similar. Each run exhibits a

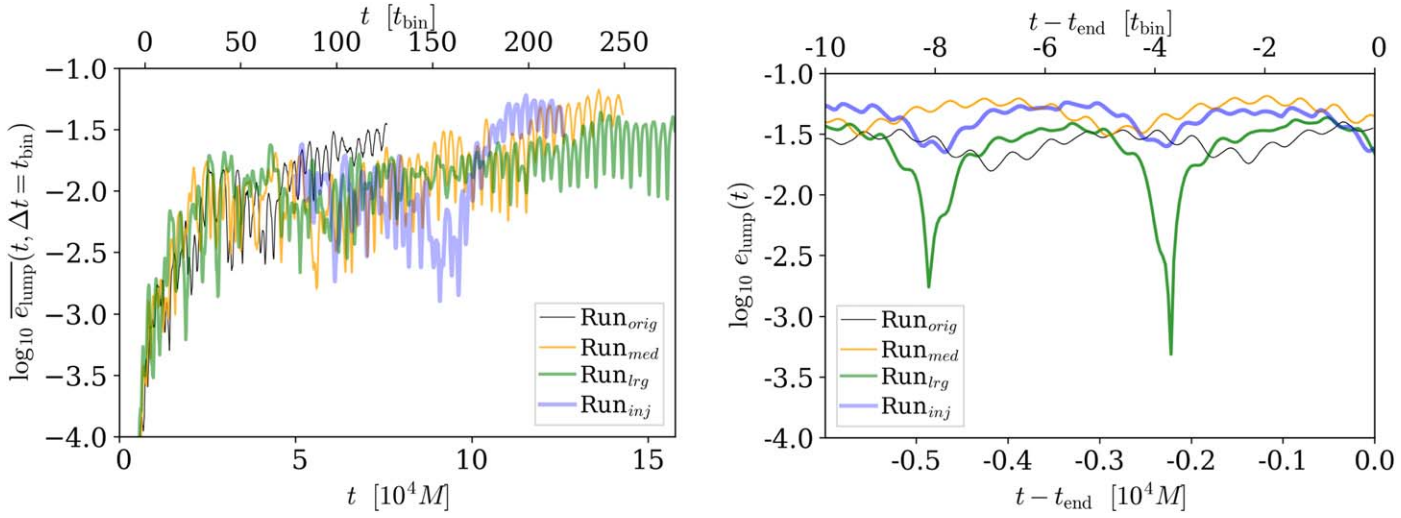

Figure 20. The lump's eccentricity smoothed over $\Delta t = t_{\rm bin}$, $\overline{e_{\rm lump}}(t, t_{\rm bin})$ (left); $e_{\rm lump}(t)$ over the last 10 orbits of each simulation (right); $t_{\rm end}$ represents the final time of each simulation of the magnetic flux series.

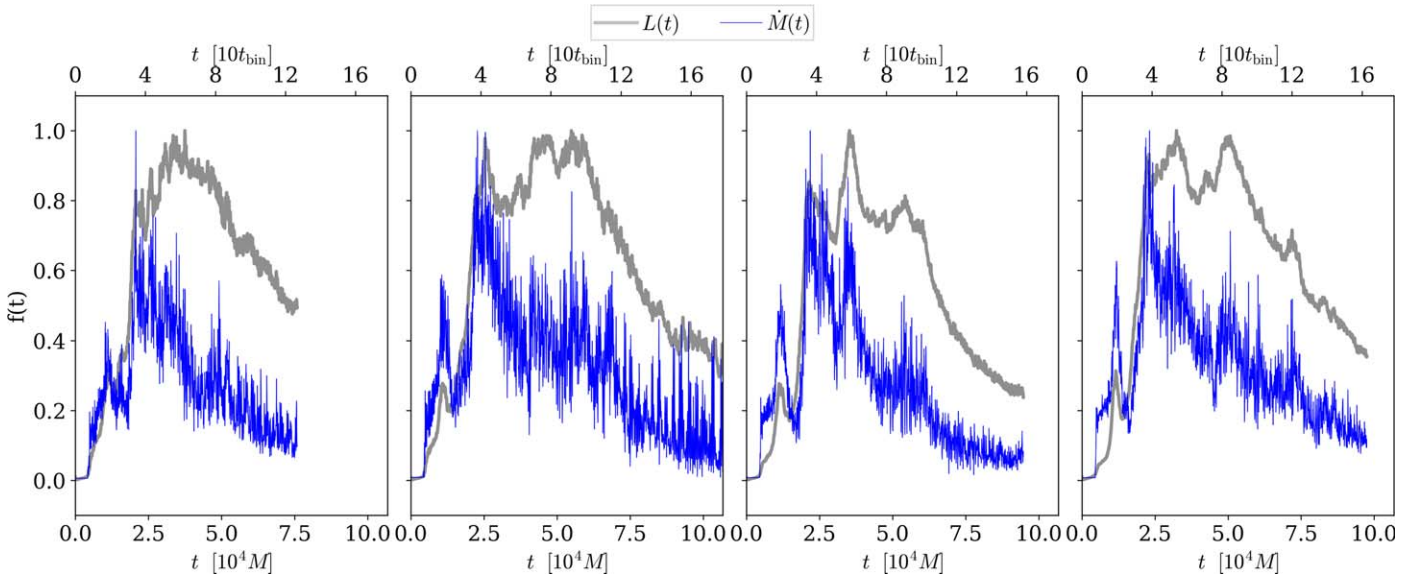

Figure 21. Light curves and accretion rates, $\dot{M}(r = a, t)$, over each simulation's full extent. Each curve is normalized by its peak value. (Left to right) RunSE, Run$_{q=1/2}$, Run$_{q=1/5}$, Run$_{q=1/10}$.

low-frequency oscillation at $\omega_{\rm lump}$, the orbital frequency of the lump, and a carrier signal at twice the beat frequency, $2(\Omega_{\rm bin} - \omega_{\rm lump})$. The mass ratio series show more differences in their $e_{\rm lump}(t)$ variability than do the magnetic flux series. Whereas RunSE and Run$_{q=1/2}$ exhibit variability at these two frequencies, $e_{\rm lump}(t)$ in Run$_{q=1/5}$ fluctuates at $2\Omega_{\rm bin}$ and $e_{\rm lump}(t)$ in Run$_{q=1/10}$ only at $\Omega_{\rm bin}$. The weakness of the lump in Run$_{q=1/5}$ and Run$_{q=1/10}$ explains they are not varying at the beat frequency; the lower frequency variability of Run$_{q=1/10}$ is due to the fact that the secondary BH dominates the gravitational torque in the would-be lump region at this small mass ratio.

4.3. Accretion Streams and Variability

We ultimately aim to provide a quantitative model of how electromagnetic emission depends on q so that system parameters may be derived from observables. The bolometric luminosity L is the simplest of observable measures. Just as in Noble et al. (2012), we calculate L by integrating the local cooling rate, $\mathcal{L}$, in the Eulerian frame over the numerical domain:

$$L(t) = \int \mathcal{L}\, u_t\, \sqrt{-g}\, dr\, d\theta\, d\phi. \tag{33}$$

In such a method, Doppler and gravitational shifts are ignored, but the magnitude of their effect is smaller than other uncertainties. Each simulation is cooled toward the same target entropy, at the same cooling timescale, starting from tori at the same initial scale height ($H/R = 0.1$ at the pressure maximum) and target entropy. Hence, any changes in L should be the result of the mass ratio or magnetic flux distribution, modulo statistical fluctuations.

In Figure 21 we show the light-curve and accretion rate for each simulation in the mass ratio series. After the initial burst of accretion, i.e., $t > 2.5 \times 10^4 M$, the luminosity in each run generally tracks the run's accretion rate, implying that the

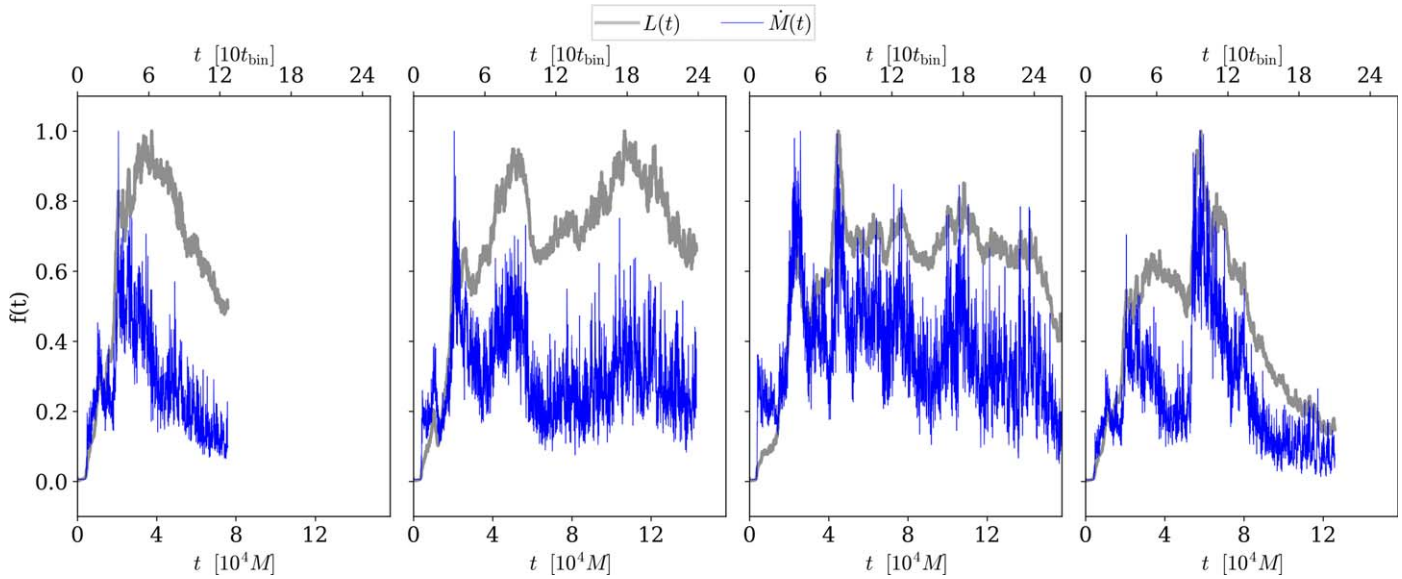

Figure 22. Light curves and accretion rates, $\dot{M}(r = a, t)$, over each simulation's full extent. Each curve is normalized by its peak value. (Left to right) RunSE, $\rm Run_{med}$, $\rm Run_{lrg}$, $\rm Run_{inj}$.

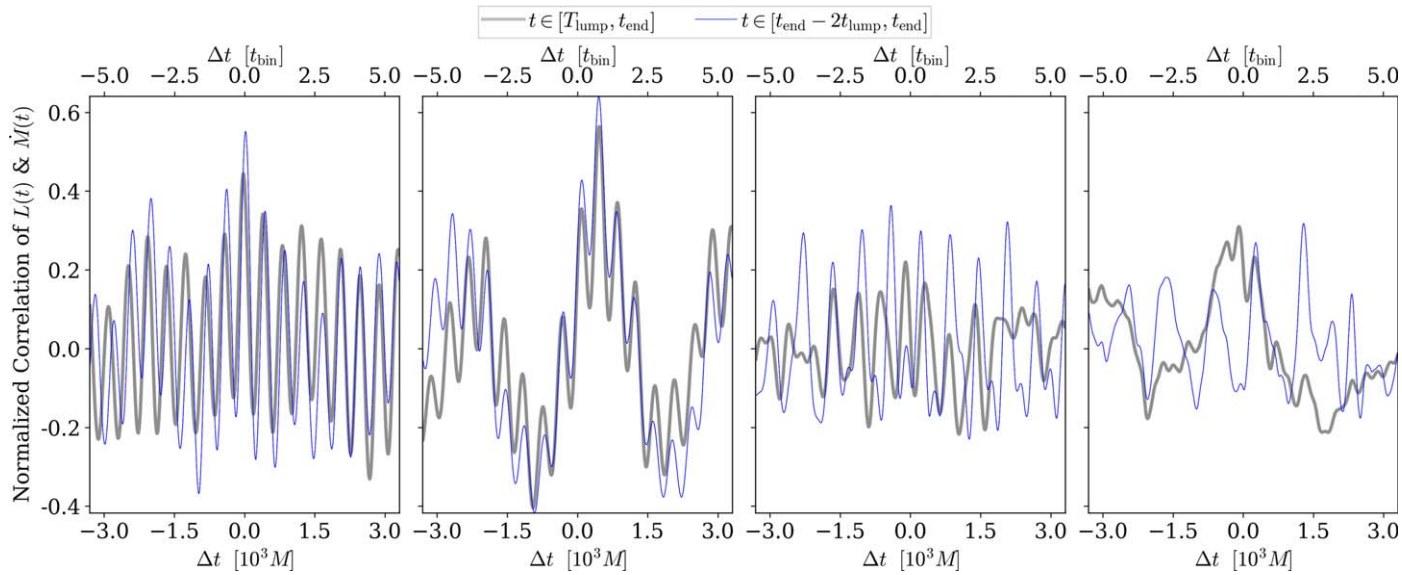

Figure 23. Normalized correlations between the light curves, $L(t)$, and accretion rates, $\dot{M}(r = a, t)$, calculated since the onset of the lump (gray curves) or the last two periods of the lump's orbit in a simulation (blue curves). The correlations are plotted vs. the lags, and are calculated using Equation (34). A 5th-order polynomial fit to each curve has been removed prior to calculating the correlation. Each plot is displayed over a span of lag time approximately equal to $3t_{\rm lump}$. (Left to right) RunSE, $\rm Run_{q=1/2}$, $\rm Run_{q=1/5}$, $\rm Run_{q=1/10}$.

radiative efficiency, $\eta \equiv L/\dot{M}$, remains nearly constant. Closer inspection, however, shows a sharp increase in η at the time of maximum luminosity; after this time, η declines slowly from this elevated value. Similar trends are also seen in the magnetic flux series, in Figure 22. We note that η is a nontrivial, time-dependent result from our simulations and is not immediately derived from our target temperature, $T_{\rm cool}$. The radiation efficiency of any one simulation depends on how efficient the gas is at turning the free energy of orbital motion into heat, and this is in and of itself a nonlinear process that depends on the disk's thermodynamics, motion, mass distribution, and local gravitational field.

In all cases, the accretion rate exhibits larger relative amplitude and higher frequency fluctuations than the luminosity does, an observation we found with single-BH accretion disk simulations using similar thermodynamics (Noble & Krolik 2009). The luminosities of the mass ratio series are all quite similar, demonstrating that the mass inflow provided by the outer disk is the dominant regulator of light output.

Although the magnitude of the luminosity in the magnetic flux series is quite similar to that of the mass ratio series, their time dependence is quite different. $\rm Run_{med}$ and $\rm Run_{lrg}$ both exhibit a plateau period in their luminosities following the initial peak. The perturbation imparted in $\rm Run_{inj}$ is apparent in both $L(t)$ and $\dot{M}(t)$. In effect, it causes the disk to go through two cycles of rise and fall, rather than the single one of the other simulations. Interestingly, there is a short delay ($\sim$2000M) between the second peak in accretion and the subsequent peak in luminosity.

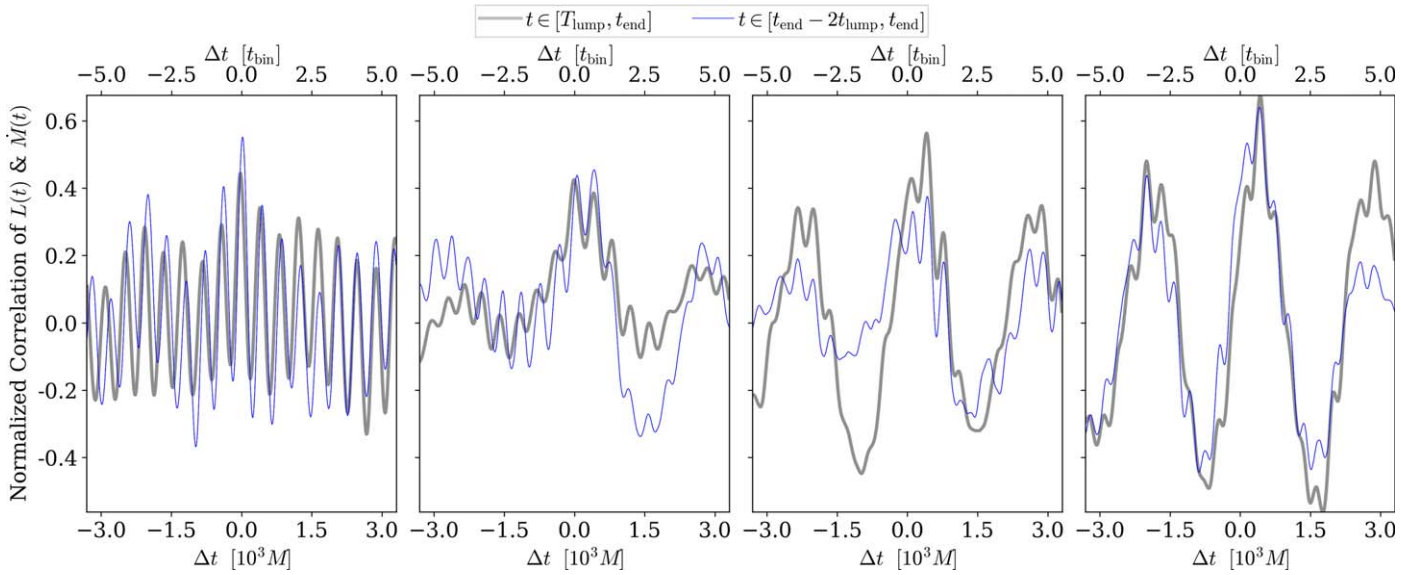

Figure 24. Normalized correlations between the light curves, $L(t)$, and accretion rates, $\dot{M}(r = a, t)$, calculated since the onset of the lump (gray curves) or the last two periods of the lump's orbit in a simulation (blue curves). The correlations are plotted vs. the lags, and are calculated using Equation (34). A 5th-order polynomial fit to each curve has been removed prior to calculating the correlation. Each plot is displayed over a span of lag time approximately equal to $3t_{\rm lump}$. (Left to right) RunSE, $\rm Run_{med}$, $\rm Run_{lrg}$, $\rm Run_{inj}$.

Phase alignment in the fluctuations of $L(t)$ and $\dot{M}(t)$ for the mass ratio series and magnetic flux series is best seen through their normalized correlations, which we show in Figures 23 and 24, respectively. We calculate the normalized correlations between $L(t)$ and $\dot{M}(t)$ using

$$\texttt{Corr}[L, \dot{M}](\Delta t) = \frac{\sum_i L(t_i)\dot{M}(t_i - \Delta t)}{||L||_2\ ||\dot{M}||_2}, \tag{34}$$

where $||L||_2$ ($||\dot{M}||_2$) is the l^2-norm of the luminosity (accretion rate) time series minus a 5th-order polynomial fit to the raw data to remove secular trends. For runs clearly exhibiting a lump (RunSE and $\rm Run_{q=1/2}$), the correlations show larger peak amplitudes than for those without ($\rm Run_{q=1/5}$ and $\rm Run_{q=1/10}$). In addition, the lump runs show variability on both the $t_{\rm lump}(\simeq 4t_{\rm bin})$ and twice the beat frequency ($1.5\Omega_{\rm bin}$) timescales, though RunSE shows relatively weak lump period variability. Because lump-producing runs all show correlations that peak at positive lags, the luminosity variation follows that of the accretion rate. This reflects the sequence of events in which an accretion stream leaves the inner edge of the circumbinary disk; part of it feels sufficient torque to return to the circumbinary disk, and once it arrives, dissipates some of its energy in a shock, whose heat is then radiated (Noble et al. 2012; Shi & Krolik 2015). Further evidence supporting the picture of how the lump region in lump-forming runs undergoes excessive heating that can be seen in the vertical scale height of the disk as measured from the gas's pressure (Appendix D).

Using the detrended and normalized functions of time, we plot the normalized Fourier power distributions in Figure 25 and Figure 26 for the mass ratio series and magnetic flux series, respectively. The power spectra are calculated using the period $t > T_{\rm lump}$ if the run exhibits a lump, and $t > t_{\rm end} - 2.5 \times 10^4 M$ if not.

The character of the variability changes with mass ratio. Neglecting the peak at very low frequency, which could be an artifact of the detrending, the nature of the strongest peak in the luminosity power spectrum is different in each case. In RunSE, it is at $\simeq 1.5\Omega_{\rm bin}$, twice the beat frequency between the lump orbital frequency and the binary orbital frequency. In $\rm Run_{q=1/2}$, there is also a strong peak at twice the beat frequency for $\dot{M}$, but the peak for L is much smaller than in RunSE, and there is a comparable peak at $\simeq 0.75\Omega_{\rm bin}$, the actual beat frequency. In $\rm Run_{q=1/5}$, there is a very strong peak at almost exactly $\Omega_{\rm bin}$. Lastly, in $\rm Run_{q=1/10}$, there is no significant periodic behavior in the light output at all. The presence of a small peak at $\Omega_{\rm bin}$ in $\rm Run_{q=1/2}$ and a larger one in $\rm Run_{q=1/5}$ may be interpreted as due to the closer approach of the secondary to the inner edge of the circumbinary disk as q decreases, and the consequent enhancement of modulation at the secondary's orbital frequency; the disappearance of this peak in $\rm Run_{q=1/10}$ is likely a sign that when the mass ratio is this small, the secondary has hardly any effect on the accretion. In all cases, there are contrasts between the power spectra of the accretion rate and the luminosity; in other words, there are significant contributions to the rate of heat dissipation that are not due immediately to mass accretion.

In Table 3 we provide the relative standard deviations of fluctuations in the accretion rate and luminosities. The most striking feature is that the fractional variation in accretion rate is consistently an order of magnitude larger than the fractional variation in the luminosity. This fact, too, strongly indicates that the luminosity is not directly related to the accretion flow. No clear trends exist in the relative variabilities within each series, though if a run is more variable than another in one quantity it typically is more variable in the other quantity as well.

We plot the PSDs of different quantities related to the lump in Figures 27 and 28. The functions analyzed are the $m = 1$ mode amplitude of the density integrated over the lump region ($\int_{2a}^{4a} A_1 dr$), the phase difference between the binary and the lump's phase ($|\phi_{\rm lump} - \phi_{\rm bin}|$), the orbital frequency of the lump ($\omega_{\rm lump}$), and $e_{\rm lump}(t)$. The phase of the lump is found by

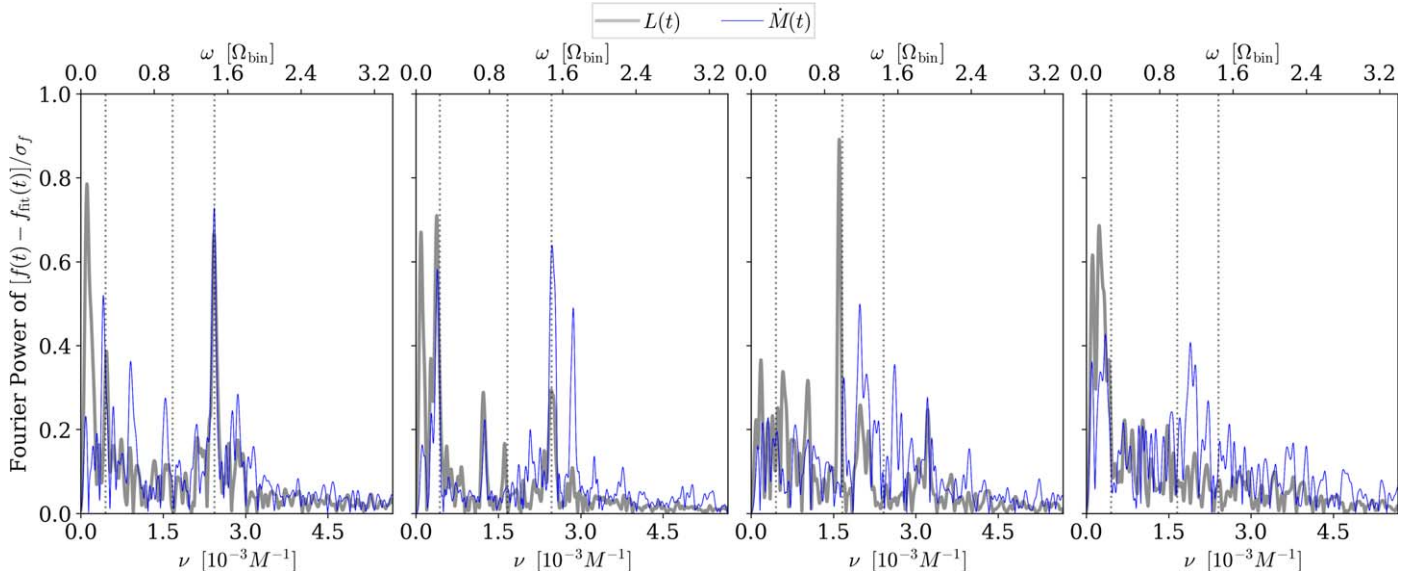

Figure 25. Fourier power spectra of the light curves and accretion rates, $\dot{M}(r = a, t)$, including only times $t > T_{\rm lump}$. For those runs with no observed lump, we use the simulation's last $2.5 \times 10^4 M$ of time. Before performing the Fourier power spectrum, the function is conditioned by subtracting a 5th-order polynomial fit and then applying a normalization factor equal to the curve's standard deviation. Vertical dotted lines in each plot lie, from left to right, at $\omega = \omega_{\rm lump}$, $\Omega_{\rm bin}$, and $2(\Omega_{\rm bin} - \omega_{\rm lump})$; for those runs without a lump, $\omega_{\rm lump}$ of RunSE is used instead. (Left to right) RunSE, $\rm Run_{q=1/2}$, $\rm Run_{q=1/5}$, $\rm Run_{q=1/10}$.

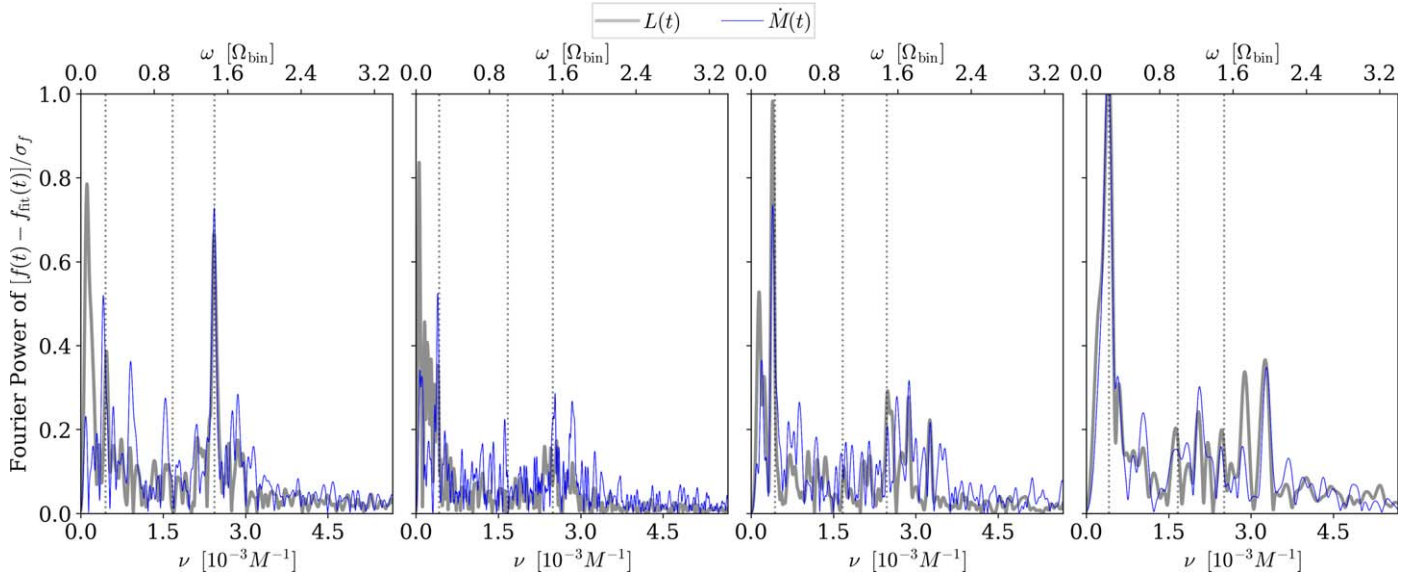

Figure 26. Fourier power spectra of the light curves and accretion rates, $\dot{M}(r = a, t)$, including only times $t > T_{\rm lump}$. For those runs with no observed lump, we use the simulation's last $2.5 \times 10^4 M$ of time. Before performing the Fourier power spectrum, the function is conditioned by subtracting a 5th-order polynomial fit and then applying a normalization factor equal to the curve's standard deviation. Vertical dotted lines in each plot lie, from left to right, at $\omega = \omega_{\rm lump}$, $\Omega_{\rm bin}$, and $2(\Omega_{\rm bin} - \omega_{\rm lump})$; for those runs without a lump, $\omega_{\rm lump}$ of RunSE is used instead. (Left to right) RunSE, $\rm Run_{med}$, $\rm Run_{lrg}$, $\rm Run_{inj}$.

locating the maximum of the $m = 1$ mode amplitude integrated over the radial extent of the lump region.

The lump's orbital frequency can then be defined by $\omega_{\rm lump} \equiv \dot{\phi}_{\rm lump}$. The average $\omega_{\rm lump}$ over this period (for those runs with a lump) is reported in Table 2; it is consistently $0.26\Omega_{\rm bin}$ or $\simeq \Omega_{\rm bin}/4$ for all runs, suggesting a 4:1 resonance between the lump and binary. As expected, in the runs exhibiting a lump, the phase difference is modulated at the beat frequency $(\Omega_{\rm bin} - \omega_{\rm lump})$ and its higher harmonics.

Several aspects of this timing analysis confirm the coherence of lump motion. Plots of $\omega_{\rm lump}(t)$ are continuous. Most clearly seen in RunSE and—to a lesser extent—in the other lump-forming runs, $\omega_{\rm lump}$ varies at twice the beat frequency, suggesting that the lump is accelerated by each passing BH. The eccentricity of lump orbits fluctuates primarily at the lump's average Keplerian rate, $\omega_{\rm lump}$ as well as at twice the beat frequency; this, too, indicates that the lump is a distinct physical element, not a pattern.

In order to connect the accretion rate and light-curve variability to the lump, we compare their PSDs (in Figure 25) to those associated with the density structure of the lump-forming region (in Figure 27). Strikingly, in both sets of PSDs the most prominent peaks for RunSE and $\rm Run_{q=1/2}$, the runs with the clearest lumps occur at the same frequencies, twice the beat frequency and $\omega_{\rm lump}$, for $q = 0.5$, 1.

For the magnetic flux series we compare Figure 26 to Figure 28. Relative to the other runs, those exhibiting larger or

Table 3
Standard Deviations $\sigma_{\dot{M}}$ (σ_L) of Accretion Rate (Luminosity) for Each Run

Run Name	$\sigma_{\dot{M}}/\overline{\dot{M}}$	$\overline{\dot{M}}\,[10^{-3}]$	$\sigma_L/\overline{L}$	$\overline{L}\,[10^{-4}]$
RunSE	0.29	5.6	0.027	3.3
$\mathrm{Run}_{q=1/2}$	0.54	3.3	0.048	1.9
$\mathrm{Run}_{q=1/5}$	0.33	2.2	0.025	1.4
$\mathrm{Run}_{q=1/10}$	0.20	4.8	0.022	1.8
$\mathrm{Run}_{\mathrm{med}}$	0.30	11.	0.043	5.4
$\mathrm{Run}_{\mathrm{lrg}}$	0.38	10.	0.033	5.1
$\mathrm{Run}_{\mathrm{inj}}$	0.56	4.1	0.054	1.4

Note. Each standard deviation is normalized by the mean of the quantity in question over this period. These averages are also displayed, though in code units. Runs were taken over the same period in which the PSDs were calculated in Figures 25 and 26.

smaller fluctuation power in the lump's properties at ω_{lump} and twice the beat frequency also exhibit larger or smaller signals in $L(t)$ and $\dot{M}(t)$ at these frequencies. $\mathrm{Run}_{\mathrm{inj}}$, which is the only run of this series that does not show significant variability in $|\varphi_{\mathrm{lump}} - \varphi_{\mathrm{bin}}|$ at $2(\Omega_{\mathrm{bin}} - \omega_{\mathrm{lump}})$, does show significant variability in this quantity at twice the other beat frequency, $2(\Omega_{\mathrm{bin}} + \omega_{\mathrm{lump}})$ as does its accretion rate.

5. Discussion

5.1. Origins and Conditions of Lump Formation

A primary goal of our paper is to investigate how and why the lump forms, and what conditions are amenable to the lump's growth. We used multiple diagnostics to verify the presence of a coherent, orbiting, overdense region of gas. In particular, the surface density, cavity wall eccentricity, lump phase, and orbital velocity, and the PSD spectra of several quantities all consistently show signs of the lump when one is present.

The mass ratio series demonstrated that sufficiently large mass ratios are required to manifest a significant lump. From the magnetic flux series, we learned that the amount of mass in the disk had little effect on the lump, but the amount of magnetic flux did: adding a relatively modest, ordered, poloidal magnetic field distribution beyond the lump region was enough to perturb it to the point the nascent lump was disrupted and formed approximately $\simeq 100 t_{\mathrm{bin}}$ or $25 t_{\mathrm{lump}}$ later. Also, the time, T_{lump}, at which a run passed our lump-formation criterion, increased with decreasing mass ratio (gravitational torque) and with increasing available mass/magnetic flux. These results suggest that a circumbinary disk's ability to form a lump is robust to minor deviations in conditions, though requires a sufficiently strong gravitational torque.

Although MHD turbulence usually has the most power on the longest spatial wavelength modes, implying that small m azimuthal modes have the most power, a distinct mechanism for sustaining a *coherent* $m = 1$ mode is required, as a $m = 1$ turbulent mode would be *incoherent*. This coherence is supplied by the binary's gravitational torque in two ways. One has previously been cited: the lump is reinforced by those portions of the accretion streams thrown back to the circumbinary disk by the gravitational torque (Shi et al. 2012; D'Orazio et al. 2013). Phase coherence is further maintained because the lump's orbit is resonant with the binary's orbit: $t_{\mathrm{lump}}: t_{\mathrm{bin}} = 4:1$. The importance of this resonance is evident in these two timescales' prominence in the PSDs of the accretion rate, luminosity, eccentricity, and $m = 1$ density mode.

As we saw in $\mathrm{Run}_{\mathrm{inj}}$ a perturbation to the magnetic field was sufficient to disturb the lump, so why is the inherent magnetic field insufficient to shear apart a growing $m = 1$ fluctuation? After all, one typically finds, in disks about single black holes, the magnetic stress per unit mass is nearly uniform in azimuth, with incoherent fluctuations having a fairly smooth inverse polynomial power spectrum w.r.t. wavelength. Obviously there is a competition between the forcing and the local shear stress.

From measuring the MRI quality factors (Appendix C), we know that the magnetic field per unit enthalpy degrades within the overdensity region. So how does the resonant interaction encourage mass growth over magnetic field growth in the circumbinary disk? For there to be a physical origin for the depletion of specific magnetic field strength in the lump, we need to understand how the magnetic field is preferentially destroyed there. The mechanism also needs to depend on the mass ratio since we find that a significant lump forms for only sufficiently large q. The answer comes from animations of magnetic field structure in the torqued streams striking the circumbinary disk, which show that the magnetic field in these streams is directed opposite to the field in the disk where the stream arrives. The collision of oppositely oriented magnetic field distribution with the inner cavity wall material leads to large-scale reconnection and dissipation of the field into heat. This process can therefore explain how the magnetic field in the lump region decreases.

The local magnetic field may grow through local MHD instabilities like the MRI and be replenished by the field carried into the region by inward fluid motion. Our interest in exploring these processes was the reason for plotting the magnetic stress per unit mass, $W^r{}_\phi$ in Figures 11 and 12. Lumps form only when $W^r{}_\phi$ falls below $\simeq 10^{-4}$ in the region near the circumbinary disk's inner edge. While this is just a correlation, it is one that works for runs with different T_{lump}, suggesting it is not a simple function of the mass ratio or initial conditions. In order to explore why this value is important, let us compare the timescales for magnetic field advection across the lump, $\Delta\, t_{\mathrm{lump}}$, and the timescale over which the magnetic field is dissipated, t_{diss}, by compression of expelled streams with an oppositely oriented magnetic field.

Assuming time steadiness of the accretion flow and that Maxwell stress accounts for the majority of the total stress, one can show that far from the edge of the disk:

$$W^r{}_\phi \simeq r\Omega_K(r)\langle u^r\rangle_\rho, \tag{35}$$

where Ω_K is the local Keplerian orbital rate, and $\langle u^r\rangle_\rho$ is the accretion inflow speed which can be used to estimate the timescale for advection of plasma across the lump, Δt_{adv}:

$$\Delta t_{\mathrm{adv}} = \frac{\Delta r_{\mathrm{lump}}}{\langle u^r\rangle_\rho} = \left.\frac{\Delta r_{\mathrm{lump}}\, r\, \Omega_K(r)}{W^r{}_\phi}\right|_{r=r_{\mathrm{lump}}} \tag{36}$$

$$= \left(\frac{\Delta r_{\mathrm{lump}}}{a}\right)\left(\frac{r_{\mathrm{lump}}}{a}\right)\left(\frac{a}{M}\right)^2 \omega_{\mathrm{lump}} M^2 (W^r{}_\phi)^{-1}$$
$$= \left(\frac{\Delta r_{\mathrm{lump}}}{a}\right)\left(\frac{r_{\mathrm{lump}}}{a}\right)^{-\frac{1}{2}}\left(2\pi W^r{}_\phi \frac{a}{M}\right)^{-1} t_{\mathrm{bin}} \tag{37}$$

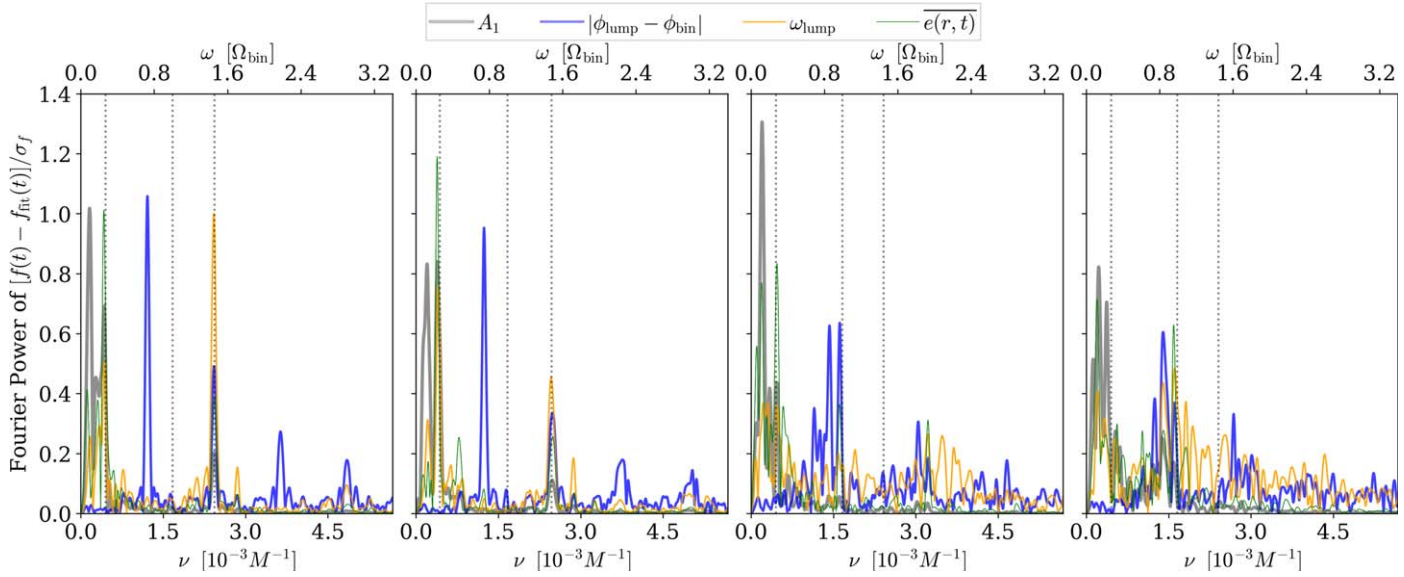

Figure 27. Fourier power spectra of quantities related to the lump's amplitude, motion, and position including only times $t > T_{\rm lump}$. For those runs with no observed lump, we use the simulation's last $2.5 \times 10^4 M$ of time. Before performing the Fourier power spectrum, the function is conditioned by subtracting a 5th-order polynomial fit and then applying a normalization factor equal to the curve's standard deviation. Vertical dotted lines in each plot lie, from left to right, at $\omega = \omega_{\rm lump}$, $\Omega_{\rm bin}$, and $2(\Omega_{\rm bin} - \omega_{\rm lump})$; for those runs without a lump, $\omega_{\rm lump}$ of RunSE is used instead. (Left to right) RunSE, Run$_{q=1/2}$, Run$_{q=1/5}$, Run$_{q=1/10}$.

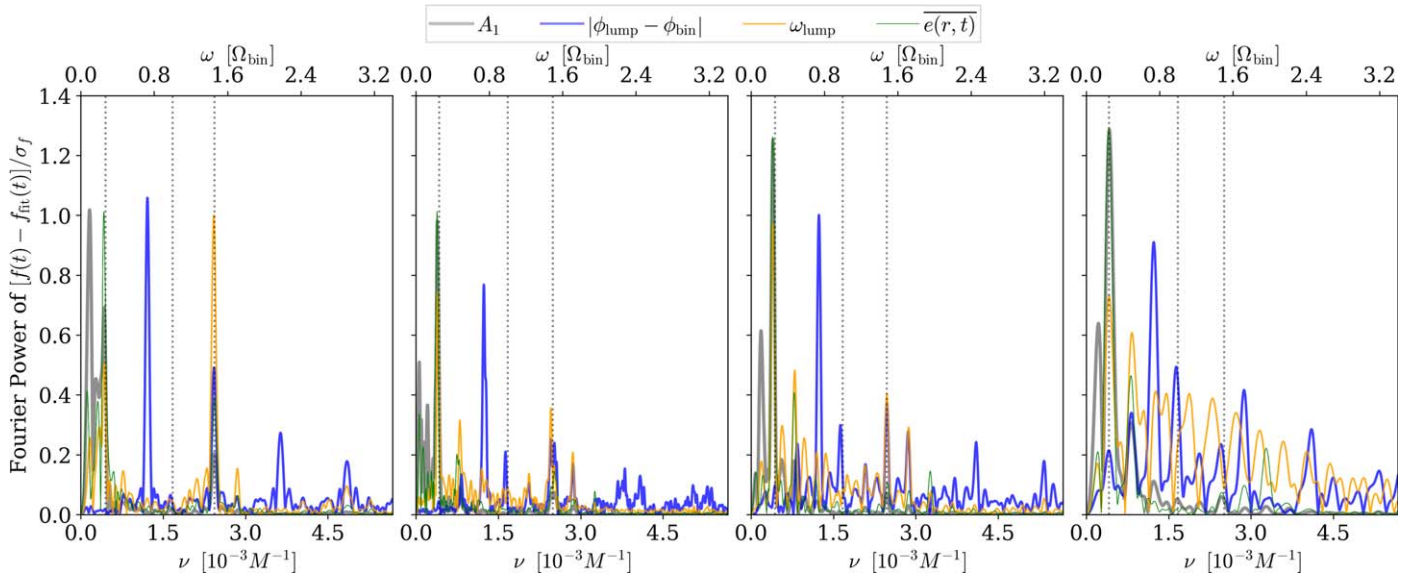

Figure 28. Fourier power spectra of quantities related to the lump's amplitude, motion, and position including only times $t > T_{\rm lump}$. For those runs with no observed lump, we use the simulation's last $2.5 \times 10^4 M$ of time. Before performing the Fourier power spectrum, the function is conditioned by subtracting a 5th-order polynomial fit and then applying a normalization factor equal to the curve's standard deviation. Vertical dotted lines in each plot lie, from left to right, at $\omega = \omega_{\rm lump}$, $\Omega_{\rm bin}$, and $2(\Omega_{\rm bin} - \omega_{\rm lump})$; for those runs without a lump, $\omega_{\rm lump}$ of RunSE is used instead. (Left to right) RunSE, Run$_{\rm med}$, Run$_{\rm lrg}$, Run$_{\rm inj}$.

$$\simeq 5t_{\rm bin}\left(\frac{\Delta r_{\rm lump}}{0.1a}\right)\left(\frac{r_{\rm lump}}{2.5a}\right)^{\frac{1}{2}} \times \left(\frac{a}{20M}\right)^{-1}\left(\frac{W^{r}{}_{\phi}}{10^{-4}}\right)^{-1}, \quad (38)$$

where we have used the Newtonian rotation rates, $\omega_{\rm lump} = M^{-1}(r_{\rm lump}/M)^{-3/2} = M^{-1}(r_{\rm lump}/a)^{-3/2}(a/M)^{-3/2}$, and $t_{\rm bin} = 2\pi M(a/M)^{3/2}$. The average radial extent of the growing lump, $\Delta\ r_{\rm lump}$, is often found to be a fixed fraction of the binary separation, a. This fraction is generically small because the $m = 1$ overdensity originates from an expelled accretion stream compressed by its shock against the cavity wall. We estimate $\Delta r_{\rm lump} \sim 0.1a$ at the time the lump begins to form, which has been observed in a number of simulations (MacFadyen & Milosavljević 2008; Noble et al. 2012; Zilhão et al. 2015; Farris et al. 2014; Miranda et al. 2017).

The dissipation timescale of magnetic field loss in the lump is the period between successive BH-overdensity interactions, which occur at twice the beat frequency $\Omega_{\rm diss} = 2(\Omega_{\rm bin} - \Omega_K(r_{\rm lump})) \simeq \frac{3}{2}\Omega_{\rm bin}$:

$$t_{\rm diss} = \frac{2\pi}{\Omega_{\rm diss}} \simeq \frac{2}{3}t_{\rm bin}. \quad (39)$$

If one process occurs at a faster rate, it will eventually win out. The ratio of the two timescales, Y, is therefore useful:

$$Y \equiv \frac{t_{\rm diss}}{\Delta t_{\rm adv}} \tag{40}$$

$$\simeq 0.13\left(\frac{\Delta r_{\rm lump}}{0.1a}\right)^{-1}\left(\frac{r_{\rm lump}}{2.5a}\right)^{-\frac{1}{2}} \times \left(\frac{a}{20M}\right)\left(\frac{W^r{}_\phi}{10^{-4}}\right). \tag{41}$$

We would expect a lump to develop once $Y < 1$, and may not otherwise because the MRI operates on a $t_{\rm lump} \gg t_{\rm bin}$ timescale at the location of the lump. For our parameters, we find this ratio implies a lump will grow. We found that the lump does not occur earlier in the lump-forming evolutions because $W^r{}_\phi$ is an order of magnitude larger, pushing $Y > 1$. When we inject magnitude field in $\rm Run_{inj}$, $W^r{}_\phi$ grows by an order of magnitude resulting in $Y > 1$ until the specific magnetic stress returns to the 10^{-4} level and the lump returns. We also note that once the lump begins to form, its radial extent grows, which makes it more difficult to rejuvenate its magnetic field through advective mixing because $Y \propto \Delta r_{\rm lump}^{-1}$.

Although this model does not explicitly depend on the mass ratio, the qualitative picture does help us understand why it is more difficult for binaries with smaller q to form a lump. First, the magnitude of the binary's time-dependent quadrupole moment decreases for smaller q. As a result, the binary's gravitational torque weakens, diminishing how much of the matter entering the gap is thrown back at the disk to feed the lump. This effect may also decrease the effectiveness of magnetic field dissipation for smaller q because a smaller amount of mass returned to the disk carries a smaller amount of (oppositely directed) magnetic field, thereby increasing W^r_ϕ. This trend is bolstered by the fact that Y increases by a further factor of 2 as the BH-lump interaction frequency falls from twice the beat frequency (for $q = 1$) to exactly the beat frequency (for $q \ll 1$).

Simulations employing viscous hydrodynamics have also demonstrated lumps (MacFadyen & Milosavljević 2008; D'Orazio et al. 2013, 2016; Farris et al. 2014, 2015; Muñoz & Lai 2016; Miranda et al. 2017; Moody et al. 2019; Mösta et al. 2019; Muñoz et al. 2019, 2020; Duffell et al. 2020; Muñoz & Lithwick 2020; Tiede et al. 2020; Zrake et al. 2021), even though magnetic reconnection cannot limit their internal stresses. However, the mathematical form of α-viscosity leads to the same effect. If a coherently orbiting structure (a nascent lump) happens to form, it has, by definition, little internal shear. Consequently, the stress given by the α-model is also small, so there is no tendency for the structure to be pulled apart by a viscous stress coupling it to surrounding matter following different orbits. On the other hand, it may grow in mass as binary torques fling outward material that has begun to move inward from the cavity wall's edge (Shi et al. 2012; Shi & Krolik 2015). The only analysis of velocity gradients in a VH simulation (Ragusa et al. 2017) found evidence supporting this picture: the vorticity component orthogonal to the disk plane has a significant local minimum in the lump. Thus, by a lucky coincidence, α-viscosity mimics the action of physical MHD in fostering the growth and longevity of lumps in circumbinary disks.

5.2. Comparison to Prior Work

In the following, we will summarize how our results compare and contrast with prior work, limiting our comparison to those studies using binaries on fixed circular orbits with prograde disks.

By all accounts, the lump is a significant density enhancement that forms near or along the cavity wall, spans a relatively narrow radial extent, and orbits as a coherent structure at the local Keplerian rate (Section 4 and MacFadyen & Milosavljević 2008; Farris et al. 2012; Noble et al. 2012; Shi et al. 2012; D'Orazio et al. 2013, 2016; Gold et al. 2014a; Zilhão et al. 2015; Bowen et al. 2018, 2019; Lopez Armengol et al. 2021; Paschalidis et al. 2021). In its asymptotic state, $\Delta r_{\rm lump} \simeq a$, while spanning a significant fraction of the possible azimuthal extent $\pi/3 \lesssim \delta\phi_{\rm lump} \lesssim \pi$. VH simulations deviate a little from this picture, at least for the larger q cases that exhibit significant $m = 1$ structure, in that they also develop a density enhancement at the apoapse (MacFadyen & Milosavljević 2008; D'Orazio et al. 2013, 2016; Farris et al. 2015; Miranda et al. 2017; Ragusa et al. 2020) that precesses at a much slower rate than the local Keplerian velocity of the disk. Some report that the lump travels through this "traffic jam" of gas at the local Keplerian rate (Ragusa et al. 2020), while others show the lump vanish and return cyclically Miranda et al. (2017).

Regardless of differences in the azimuthal distribution of the lump, all[10] simulations find that it leads to accretion rate modulation at a timescale approximately equal to the Keplerian period at the cavity's edge. This modulation timescale usually is approximately $t_{\rm lump} \simeq 4 - 5t_{\rm bin}$ for cases with $h/r \sim 0.1$ (MacFadyen & Milosavljević 2008; Noble et al. 2012; Shi et al. 2012; D'Orazio et al. 2013, 2016; Farris et al. 2015; Zilhão et al. 2015; Muñoz & Lai 2016; Miranda et al. 2017; Moody et al. 2019; Muñoz et al. 2020; Duffell et al. 2020; Dittmann & Ryan 2021; Lopez Armengol et al. 2021). All our simulations with a lump show $t_{\rm lump} \simeq 4t_{\rm bin}$, which were the cases when $q > 0.2$. Interestingly, Duffell et al. (2020) and Dittmann & Ryan (2021) also report the modulation changes rather sharply at $q = 0.2$ for circular orbits and $h/r = 0.1$ disks, and they both found that $t_{\rm lump}$ gradually decreased from $5t_{\rm bin}$ at $q = 1$ to $4t_{\rm bin}$ at $q \to 0.2$. Others show transitions in the same ballpark: $q > 0.25$–0.5 (D'Orazio et al. 2013), $q > 0.25 - 0.43$ (Farris et al. 2015), and $q \geqslant 0.4$ (Muñoz et al. 2020). For cases with larger cavities (e.g., with colder disks), longer periods are found: Ragusa et al. (2016) find $t_{\rm lump} \simeq 5t_{\rm bin}$ ($t_{\rm lump} \simeq 4t_{\rm bin}$) for $h/r = 0.04$ ($h/r = 0.1$), and Ragusa et al. (2020) report modulations at $7-8t_{\rm bin}$ from a cavity of radius $r \simeq 3.5a$. A dissenting case is the $q = 0.1$ simulation of Shi & Krolik (2015) which shows evidence of the lump in the surface density and the accretion rate, with the lump even forming later in their $q = 1$ case; the difference may be due to their use of an isothermal equation of state. As our MHD simulations, like those of others, include turbulent circumbinary disks, our PSDs of $\dot{M}(t)$ are more complex than those of VH simulations and include significant power over a range of frequencies, like those observed in real AGN disks. Demonstrating that variability associated with the lump is evident above these broadband fluctuations for several different cases is a key finding of this paper.

[10] Few numerical relativity simulations have the temporal range to calculate a PSD of $\dot{M}(t)$, and those that have PSDs have relatively poor frequency resolution and dynamic range to yield significant signal-to-noise ratios (Gold et al. 2014a).

Few papers, however, investigate how the lump forms and endures. Shi et al. (2012), D'Orazio et al. (2013), and Shi & Krolik (2015) went to significant lengths to show how returning streams reinforce the lump. In Noble et al. (2012) we showed how the MRI quality of our simulation degraded in the lump; this analysis is repeated for our simulations here in Appendix C. Evidence of degradation of magnetic and hydrodynamic turbulence in the lump is shown in Shi & Krolik (2016) by spatially associating the lump with the relative contributions of the vertical components of the velocity and magnetic field. And, as we mentioned earlier, Ragusa et al. (2017) explore whether the lump is stabilized by vorticity, which they found to not be the case as vorticity was shown to be smaller in the lump than in the rest of the circumbinary disk. All evidence presented so far is consistent with the notion that the lump is a coherently orbiting structure with a deficit or absence of internal differential rotation or local vorticity; as we have mentioned, the lack of differential rotation helps to stabilize it by hindering the MRI or viscous stresses.

6. Summary

We have explored two series of simulations of circumbinary accretion disks about binary black holes using GRMHD and 2.5PN approximate spacetimes: the mass ratio series, in which the relative masses of the black holes were $q = 0.1$, 0.2, 0.5, or 1, while all other aspects remained the same, and the magnetic flux series, which used an equal-mass binary but different distributions of mass and magnetic field. Axisymmetric and nonaxisymmetric properties of the circumbinary disks were explored in each series. Special emphasis was given to aspects of the circumbinary disk that may affect binary signatures in EM emission, such as the lump.

In terms of axisymmetric aspects, we found that the circumbinary disk in most cases approached a steady state of mass inflow by the end of each run. Aspects such as the surface density peak (gradient) were found to broaden (move inward) as the mass ratio decreased. Runs with larger distributions of mass/magnetic flux took longer to asymptote to a steady state, but otherwise resembled the other equal-mass runs in axisymmetric aspects.

Regarding the nonaxisymmetric properties, we explored the disk's azimuthal structure, eccentricity, and variability. We found a new diagnostic for the relative strength of the $m = 1$ azimuthal mode of density such that lump formation was always sustained if this quantity exceeded a certain value. The growth in time of this diagnostic was found to be coincident with the growth rate of the eccentricity of material near the cavity. When arising, the lump is found to be a coherent structure in which magnetic stress is kept comparatively weak by repeated collisions with expelled material carrying oppositely oriented magnetic field. The lump is also associated with variability in the light-curve and accretion rate on two timescales related, but not identical, to the binary period: $4t_{\rm bin}$ and $2/3t_{\rm bin}$, the former associated with the lump's local Keplerian period, the latter with the beat mode between the binary and the lump.

The lump was observed in $q = 0.5$, 1 and not found in $q = 0.2$, 0.1, suggesting the no-lump/lump transition lies between $0.2 < q < 0.5$ for our simulation conditions. We also demonstrated that persistent reservoirs of mass and magnetic flux available to be accreted into the lump region do not hinder its growth. However, a perturbation to the accretion flow in the form of a modest additional ordered magnetic field was sufficient to disrupt a nascent lump and delay its development for $\sim 10t_{\rm bin}$.

Previous work on the origin and sustenance of lumps stressed the coordinated delivery of matter to the lump region by "torqued-up" streams (Shi et al. 2012; D'Orazio et al. 2013; Shi & Krolik 2015). We have found that the matter returned to the disk in this way carries with it a magnetic field that tends to be directed opposite to the already-existing field in the lump. Reconnective dissipation then suppresses the magnetic stress per unit mass in the lump, helping the lump region retain its mass. As a result, lumps grow once the magnetic stress per unit mass falls below a critical level, a threshold that is essentially independent of simulation parameters. Our model may also explain how lumps are stabilized in viscous hydrodynamics simulations since these calculations include shear stress proportional to the velocity shear of the flow and the lump region has insignificant shear; without shear stress, there is no local mechanism for angular momentum transport across the lump and no means to disrupt it.

Because we have not explored the entire parameter space of conditions, our conclusions are limited to the choices made herein (e.g., the thermodynamic model, disk aspect ratio). We will reserve exploring other parts of parameter space for future work. Further, these new, relaxed circumbinary disk simulations provide starting conditions with which we may pursue simulations with resolved black holes (Bowen et al. 2018, 2019; d'Ascoli et al. 2018), as we have done using circumbinary disk data of Noble et al. (2012).

We thank Carlos Lousto (RIT), Federico Lopez Armengol (RIT), Luciano Combi (Instituto Argentino de Radioastronomía), Mark Avara (Cambridge), and Dan D'Orazio (NBI, Copenhagen) for valuable discussions on work described in this manuscript. We also thank Alexandra A. Gladkova (U. Colorado Boulder, NASA Internship Program) for work that aided the analysis presented in Appendix B.

This work was supported by several National Science Foundation (NSF) grants. S.C.N., M.C., M.Z., and Y.Z. received support from awards AST-1028087 and J.H.K. from AST-1028111, PHY-1707826, and AST-2009260. M.C., B.C. M, H.N., Y.Z. also acknowledge partial support from grants PHY-0929114, PHY-0903782, PHY-0969855, and PHY-1707946. Additionally, M.C. was supported by awards AST-1516150 and AST-2009330, and S.C.N. was supported by awards AST-1515982 and OAC-1515969.

H.N. acknowledges support from JSPS KAKENHI grant Nos. JP16K05347 and JP17H06358.

S.C.N. was supported by an appointment to the NASA Postdoctoral Program at the Goddard Space Flight Center administrated by USRA through a contract with NASA.

This research is part of the Blue Waters sustained-petascale computing project, which is supported by the NSF (awards OCI-0725070 and ACI-1238993) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications. This work is also part of the "Computational Relativity and Gravitation at Petascale: Simulating and Visualizing Astrophysically Realistic Compact Binaries" PRAC allocation support by the NSF (award OCI 0832606). All but one simulation ($q = 1/2$) was performed on Blue Waters. The remaining simulation was performed on Stampede

at the Texas Advanced Computing Center through the XSEDE allocation TG-PHY060027N. Analysis was performed on the NewHorizons and BlueSky Clusters at Rochester Institute of Technology, which were supported by NSF grant No. PHY-0722703, DMS-0820923, AST-1028087, and PHY-1229173.

This work was performed in part at Aspen Center for Physics, which is supported by NSF grant PHY-1607611.

This research was supported by the Munich Institute for Astro- and Particle Physics (MIAPP) which is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC-2094—390783311.

Software: astropy (Astropy Collaboration et al. 2013), HARM3D (Gammie et al. 2003; Noble et al. 2006, 2009, 2012).

Appendix A
NZ Metric Range of Validity

In order to discuss the range of validity for the NZ metric, we must first discuss the two types of PN approximation. The first is the PN approximation to the metric. For simplicity, we consider $g_{rr} = 1/(1-2M/r)$ of the Schwarzschild metric in the Schwarzschild coordinates, and the series expansion with respect to M/r up to 2PN order, $g_{rr}^{(3\mathrm{PN})} = 1 + 2M/r + 4M^2/r^2$. In practice, we do not have $O((M/r)^3)$ terms in the spatial component of the NZ metric. The error due to the PN truncation is evaluated as $|g_{rr}^{(2\mathrm{PN})}/g_{rr} - 1| = 1.6\%$, 0.8% and 0.24% for $r/M = 8$, 10, and 15, respectively.

The second type of PN approximation is in the PN equations of motion. Here, we apply the result shown in Sago et al. (2016) and Fujita et al. (2018) which are extensions of Yunes & Berti (2008) and Zhang et al. (2011) for the PN region of validity for quasi-circular orbits of a point particle orbiting around a massive BH. According to the appendix of Sago et al. (2016), we have the radius of convergence for the orbital velocity v around $v \approx 0.5$. Considering the orbital velocity as $v = \sqrt{M/r}$, the above fact means that we may use $1/(1-4M/r)$ as a resummed form. The series expansion with respect to M/r up to 3PN order that is used in this paper, becomes $1 + 4M/r + 16M^2/r^2 + 64M^3/r^3$. From a similar analysis to the metric case, we have the PN truncation error, 6.3%, 2.6%, and 0.51% for $r/M = 8$, 10, and 15, respectively. The above two analyses give only the PN truncation errors, and in practice, we need to compare the PN orbital evolutions with those obtained by numerical relativity simulations for BBHs (see, e.g., Szilagyi et al. 2015 for comparisons of the waveforms). According to Ajith et al. (2012) which gives a guideline for a gravitational-wave frequency to hybridize the PN and numerical relativity waveforms, the PN approximation may be good up to r/M ~9. Furthermore, from the right figure of Figure 3 in Mundim et al. (2014), which is shown for the $M_1 = M_2 = M/2$ and $a/M = 20$ case, the accuracy of the NZ and inner zone metrics becomes comparable around $r/M_1 = 5$,

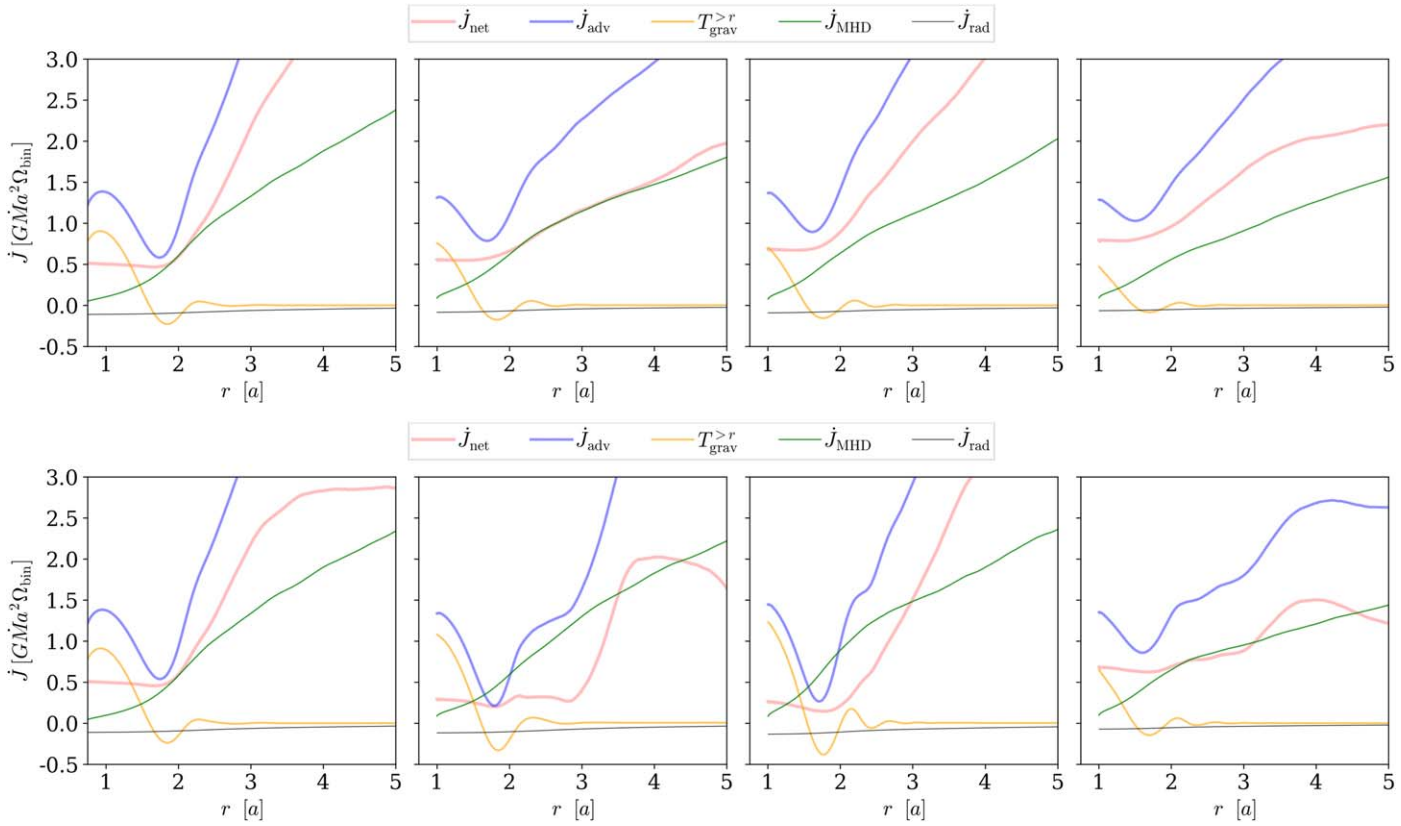

Figure 29. Contributions to the time-averaged radial distribution of $\dot{J}_{\rm net}$ (red) in the mass ratio series. Shown are the contributions from the hydrodynamic stress ($\dot{J}_{\rm adv}$, blue), gravitational torque ($T^{>r}_{\rm grav}$, orange), Maxwell stress ($\dot{J}_{\rm MHD}$), green), and torque from the radiation flux ($\dot{J}_{\rm rad}$, black). All quantities in the top (bottom) row plots are time-averaged over $40{,}000 < t/M < 76{,}000$ (last $30{,}000M$ of evolution per run). (Left to right) RunSE, $\mathrm{Run}_{q=1/2}$, $\mathrm{Run}_{q=1/5}$, $\mathrm{Run}_{q=1/10}$.

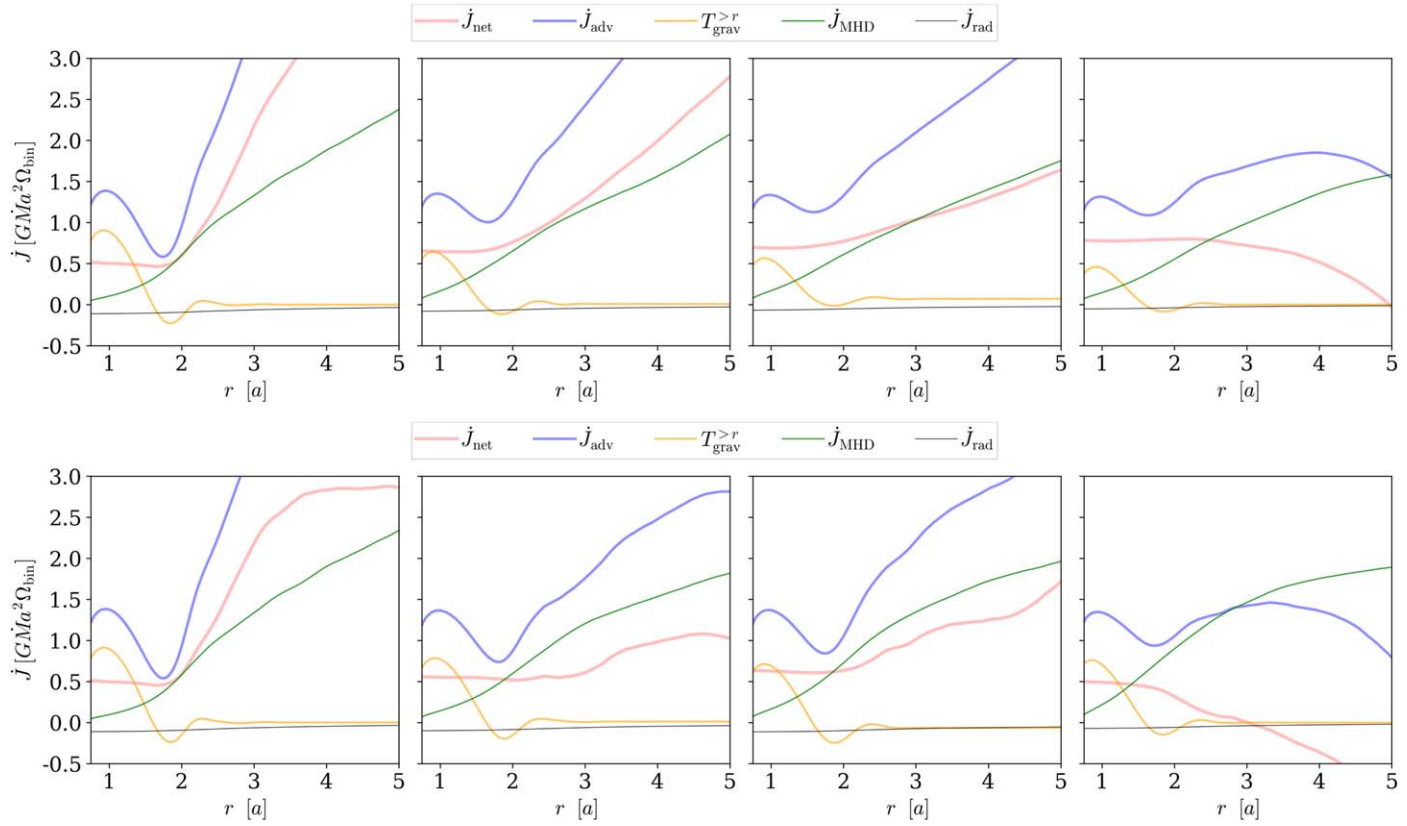

Figure 30. Contributions to the time-averaged radial distribution of $\dot{J}_{\rm net}$ (red) in the magnetic flux series. Shown are the contributions from the hydrodynamic stress ($\dot{J}_{\rm adv}$, blue), gravitational torque ($T^{>r}_{\rm grav}$, orange), Maxwell stress ($\dot{J}_{\rm MHD}$, green), and torque from the radiation flux ($\dot{J}_{\rm rad}$, black). All quantities in the top (bottom) row plots are time-averaged over $40{,}000 < t/M < 76{,}000$ (last $30{,}000M$ of evolution per run).

i.e., $r/M = 10$. From the above observations, the NZ metric is appropriate for the region of both $r_1 \gtrsim 10M_1$ and $r_2 \gtrsim 10M_2$ to describe the spacetime, and this gives $r_{\rm in}/a \gtrsim (1/2+q)/(1+q)$ from the primary BH and $r_{\rm in}/a \gtrsim (1+q/2)/(1+q)$ from the secondary BH. Therefore, the constraint on $r_{\rm in}$ is derived as Equation (3).

Appendix B
Net Angular Momentum Transfer Rate

One way we explored the contributions of the angular momentum transfer budget was through the gradient of the net angular momentum rate, $d^2J/drdt$, as we did in Section 3.4. Any local departures from steady state, during the time periods we averaged over, are reflected in nonzero values of this time-averaged quantity. Another way to evaluate the torque budget and steady state is to compare the contributions to the integrated rate, $\dot{J} = dJ/dt$, as was done in Miranda et al. (2017) and repeated in other Newtonian viscous hydrodynamics studies. These analyses reported the rates of angular momentum change in terms of advected mass, viscous, and gravitational components. Since we use GRMHD, we cannot use the exact expressions employed in prior work, but must generate our own based on the covariant conservation laws, described in Farris et al. (2011) and Appendix C of Noble et al. (2012).

Starting from the gradient expressions, Equation (20), and combining the two hydrodynamics-only terms together, we obtain:

$$\partial_r \dot{J} = dT/dr - \{\mathcal{F}_\phi\} - \partial_r\{M^r{}_\phi\} - \partial_r\{T^r_{H\phi}\}. \tag{B1}$$

Since each term should asymptote to zero as $r \to \infty$, we will assume that this condition holds at our outer boundary and integrate to it from each radius to obtain each rate as a function of r:

$$\dot{J}_{\rm net} \equiv -\dot{J}(r) = \int_r^\infty \partial_{r'}\partial_t\, J\, dr', \tag{B2}$$

$$\dot{J}_{\rm adv} \equiv \int_r^{r_{\rm out}} \partial_{r'}\{T^r_{H\phi}\}\, dr', \tag{B3}$$

$$\dot{J}_{\rm MHD} \equiv -\int_r^{r_{\rm out}} \partial_{r'}\{M^r{}_\phi\}\, dr', \tag{B4}$$

$$\dot{J}_{\rm rad} \equiv -\int_r^{r_{\rm out}} \{\mathcal{F}_\phi\}\, dr', \tag{B5}$$

$$T^{>r}_{\rm grav} \equiv \int_r^{r_{\rm out}} \frac{dT}{dr}\, dr'. \tag{B6}$$

We have defined these angular momentum rate variables with the appropriate sign to be consistent with the convention of Miranda et al. (2017) so that Equation (B1) times a minus sign yields Equation (A15) of Miranda et al. (2017), where our $\dot{J}_{\rm net}$ equals their "$\langle \dot{J} \rangle$":

$$\dot{J}_{\rm net} = \dot{J}_{\rm adv} - \dot{J}_{\rm MHD} - T^{>r}_{\rm grav} - \dot{J}_{\rm rad}. \tag{B7}$$

As far as we know, this is the first time that this expression has been derived for GRMHD systems.

From our $d^2J/drdt$ analysis, each run grows steadier in the later time averages, demonstrated here by the flatter $\dot{J}_{\rm net}$ curves in Figures 29 and 30. The flattest $\dot{J}_{\rm net}$ curves are seen at later times for $\rm Run_{med}$ and $\rm Run_{lrg}$, which both were run longer than any other of the simulations, implying that further evolution is needed for each simulation to achieve a steadier state as evaluated by this criterion. When $\dot{J}_{\rm net}$ is flat, we find that $\dot{J}_{\rm net} \simeq 0.25 - 0.75$, though only $\rm Run_{q=1/2}$ and $\rm Run_{q=1/5}$ yield $\dot{J}_{\rm net} < 0.5$ over the flatter inner portion of $\dot{J}_{\rm net}(r)$.

Appendix C
MRI Resolution

The ability for a finite volume/difference code to adequately resolve MHD turbulence depends largely on whether the fastest growing mode of the MRI is resolved Sano et al. (2004). By performing a series of resolution studies of global MHD disks, Hawley et al. (2011), Sorathia et al. (2012) found that global, extrinsic characteristics of the accretion flow asymptote with resolution. They found that those simulations that met or surpassed a particular set of resolution criteria would reside in the asymptotic regime. We follow these guidelines in constructing the simulations here, just as we did in Noble et al. (2012), and use the same resolution before. Please refer to Appendix B of Noble et al. (2012) for more details.

The MRI quality factors $Q^{(i)}$ are ratios of the local MRI wavelength and the local grid scale in a particular direction Noble et al. (2010):

$$Q^{(i)} = \frac{2\pi |b^i|}{\Delta x^{(i)}\ \Omega_K(r)\ \sqrt{\rho h\ +\ 2p_m}}, \tag{C1}$$

where index "i" denotes the spatial numerical coordinate, with $i = 1,\ 2,\ 3$ representing the radial, poloidal, and azimuthal directions, respectively. Averages in polar angle ($x^{(2)}$) are mass-weighted to bias the integral over the turbulent portion near the disk's midplane rather than the laminar regions of the corona and funnel:

$$\langle Q^{(i)} \rangle_\rho \equiv \frac{\int_0^1 Q^{(i)} \rho\ \sqrt{-g}\ dx^{(2)}}{\int_0^1 \rho\ \sqrt{-g}\ dx^{(2)}}. \tag{C2}$$

The vertically averaged quality factors, $\langle Q^{(1)} \rangle_\rho$, $\langle Q^{(2)} \rangle_\rho$, $\langle Q^{(3)} \rangle_\rho$ are shown in Figure 31 for the mass ratio series and Figure 32 for the magnetic flux series. The quality factors $Q^{(2)}$ and $Q^{(3)}$ are typically the most challenging to achieve, especially for thinner disks. This is why Hawley et al. (2011) recommend targets for $Q^{(2)}$ and $Q^{(3)}$ of 10 and 25, respectively, above which they found simulations to be in the convergent regime.

In the mass ratio series (Figure 31), we find that all runs satisfy the $Q^{(2)} > 10$ and $Q^{(3)} > 25$ MRI quality conditions for all

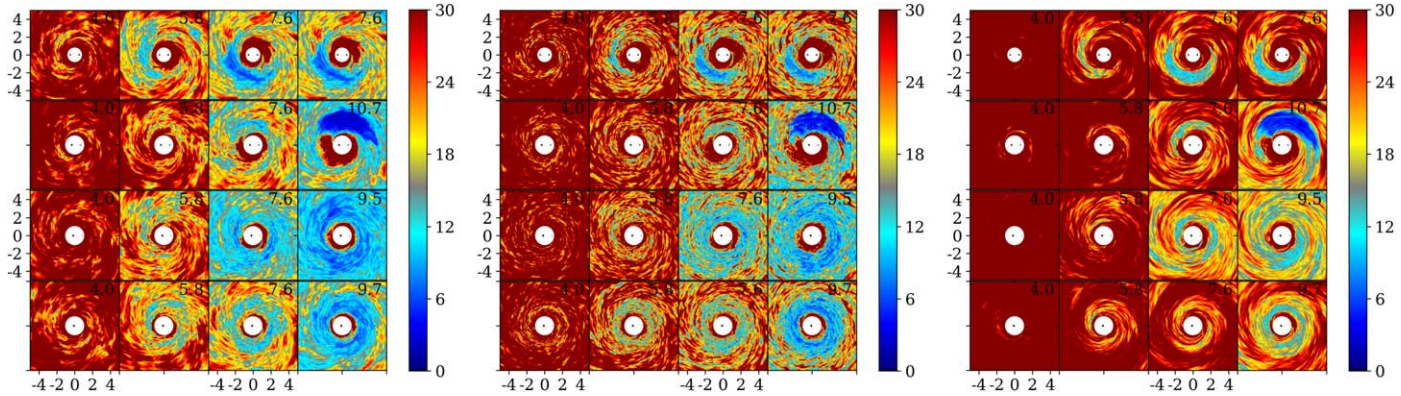

Figure 31. (Left to right) The mass-weighted vertically integrated MRI quality factor, $\langle Q^{(i)} \rangle_\rho$, in the radial, polar, and azimuthal directions, respectively, at three overlapping times and the final time of each simulation. The times of each snapshot are specified in the upper-right corner of each frame in units of 10^4M. The vertical and horizontal axes are in units of $a = 20M$. (Top to bottom) RunSE, $\rm Run_{q=1/2}$, $\rm Run_{q=1/5}$, $\rm Run_{q=1/10}$.

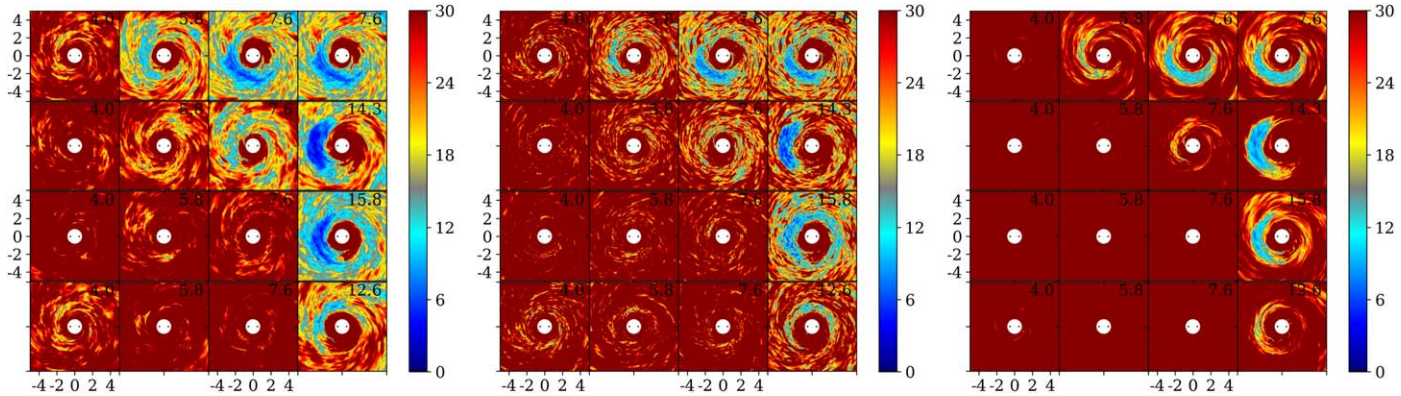

Figure 32. (Left to right) The mass-weighted vertically integrated MRI quality factor, $\langle Q^{(i)} \rangle_\rho$, in the radial, polar, and azimuthal directions, respectively, at three overlapping times and the final time of each simulation. The times of each snapshot are specified in the upper-right corner of each frame in units of 10^4M. The vertical and horizontal axes are in units of $a = 20M$. (Top to bottom) RunSE, $\rm Run_{med}$, $\rm Run_{lrg}$, $\rm Run_{inj}$.

times, with some exceptions. The first exception is that $\mathrm{Run}_{q=1/5}$'s quality factors diminish to just below their threshold values for $t \gtrsim 7.6 \times 10^4 M$; $\mathrm{Run}_{q=1/10}$ also shows a decrease in MRI quality over this period, but not as steep a decline in time, and only in localized regions are sub-threshold levels reached. Also, all quality factors are less than their targets in the lump once it forms, as we found in Noble et al. (2012), too.

The magnetic flux series (Figure 32) on the other hand is much better resolved in terms of its MRI quality factors. The only place where the MRI quality factors drop slightly below their target values is within the lump toward the end of each run, with the quality factors far exceeding the targets everywhere else. Interestingly, the perturbation added in $\mathrm{Run}_{\mathrm{inj}}$ seems to have led to conditions in which the quality factors are sustained at much higher levels than otherwise, i.e., in RunSE; it is somewhat not surprising that ordered poloidal field fosters more active MRI-driven turbulence, which is precisely why we added it in the first place.

In general, we consider our simulations well resolved, with the loss of MRI quality explained as a natural consequence of binary-stream interactions (Section 5.1) or marginal enough to not affect our qualitative conclusions.

Appendix D
Aspect Ratio

The aspect ratio, $H(r)/r$, where $H(r)$ is the vertical scale height of the disk at radius r, is often used in accretion disk theory to understand a disk's steady-state thermodynamics. Retained heat can lift gas from the minimum energy surface at the equator, therefore, the aspect ratio is directly related to the vertical distribution of rest-mass density, ρ, and the local temperature, $T(r)$.

In our simulations, the density and temperature are independent quantities, which means that there are two primary means at arriving at the aspect ratio: 1) from the vertical density-weighted angular moment of the disk (H_ρ/r), and 2) from the available local pressure support of the gas (H_p/r) (e.g., Noble et al. 2010). The aspect ratio from the density distribution is:

$$\frac{H_\rho}{r} \equiv \langle \sqrt{g_{\theta\theta}} |\theta - \pi/2| \rangle_\rho. \tag{D1}$$

If T_* is the temperature of an isothermal disk with aspect ratio H/r, then the condition for vertical gravity equilibrium yields:

$$T_* = \frac{\pi}{2} \frac{R_z(r)}{r} \left[\frac{H(r)}{r} \right]^2, \tag{D2}$$

where $R_z(r)$ is related to the Ricci tensor and includes a relativistic correction to the vertical gravity component

$$R_z(r) = \frac{1}{r}[l_k^2 - a^2(\epsilon_k^2 - 1)], \tag{D3}$$

and l_k and ϵ_k are—respectively—the specific angular moment and energy of circular time-like geodesics in the equator of a Kerr black hole with spin parameter a (Abramowicz et al. 1997; Krolik 1999; note the correction in $R_z(r)$ reported in Footnote 5 of Noble et al. 2010). Since our gas is not isothermal, we must approximate T_* with an average temperature of our gas. Because the gas may be supported by thermal or magnetic pressure, we calculate this average temperature with just the gas pressure, $T_* = \{p\}/\{\rho\}$, or with the sum of the gas and magnetic pressures, $p_{\mathrm{tot}} \equiv p + p_{\mathrm{mag}}$: $T_* = \{p_{\mathrm{tot}}\}/\{\rho\}$. Then $H_p(r)/r$ or $H_{p_{\mathrm{tot}}}(r)/r$ are found by solving Equation (D2) for H/r using the appropriate T_*. We only include bound material in the shell integrals when calculating T_*.

The different aspect ratios for the two series of runs are shown in Figures 33–34. In all runs, we find that H_ρ/r and $H_{p_{\mathrm{tot}}}/r$ deviate by less than 10% for any time interval and for radii $r > 2a$, and H_p/r is typically 10%–20% smaller than H_ρ/r over the same radial range. This demonstrates that total pressure support can explain the vertical density distribution on average, with thermal pressure providing the majority of support.

In general, there is no clear trend in any of the scale height measures with respect to mass ratio or initial data. Each aspect ratio quantity is constant to within 10% over $2.5a < r < 5a$, which means that our cooling function works well at regulating a constant aspect ratio in the disk. However, we see that each H/r function ratio shows significant deviations within $r = 2a$ indicating that the cooling rate is less than the heating rate there, not surprising due to the significant dissipation and ballistic nature of the accretion streams. In runs with measurable lumps we see excess in H_p/r in the region $2a < r < 3a$, but no clear sign in the other measures of aspect ratio. This indicates that extra heating occurs at the site of the lump and this extra heating does not significantly alter the

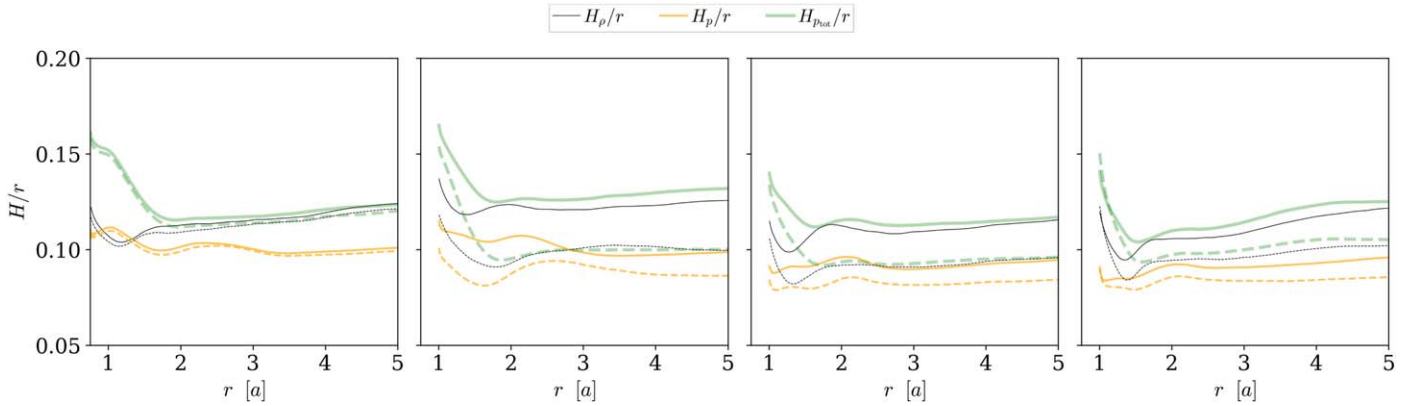

Figure 33. Time average of the aspect ratio as a function of r for the mass ratio series. The time intervals used for the averages are $40{,}000 < t/M < 76{,}000$ (solid curves), the secularly evolving period of RunSE, and the last $3 \times 10^4 M$ of each run (dashed curves).

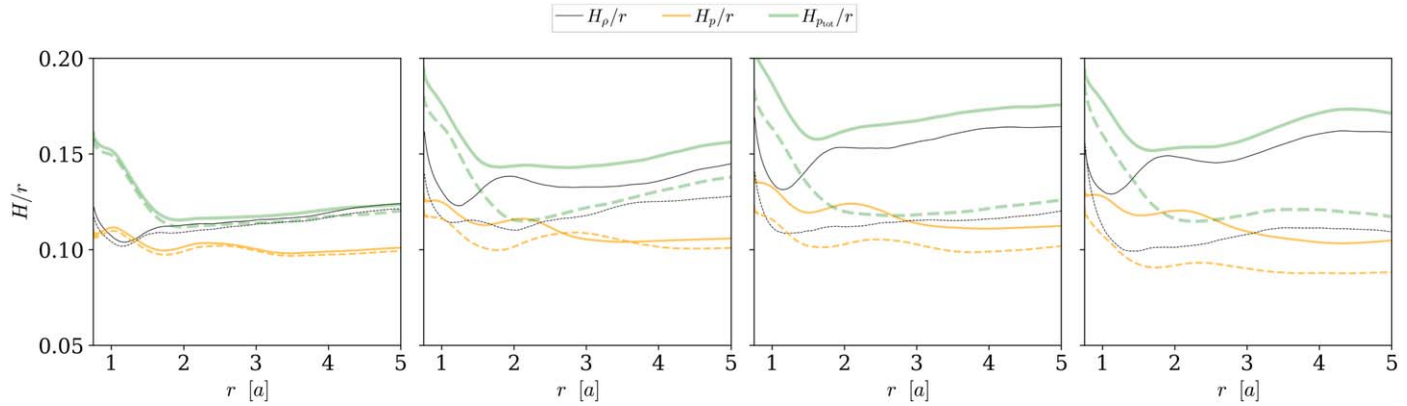

Figure 34. Time average of the aspect ratio as a function of r for the magnetic flux series. The time intervals used for the averages are $40{,}000 < t/M < 76{,}000$ (solid curves), the secularly evolving period of RunSE, and the last $3 \times 10^4 M$ of each run (dashed curves).

density scale height of the disk, likely because the excess heat is removed through the cooling mechanism before it may further lift the gas. We do see small excesses in the two runs without a lump, $\mathrm{Run}_{q=1/5}$ and $\mathrm{Run}_{q=1/10}$, but they reside closer to $r = 2a$, are narrower in extents, and are likely due to excess dissipation occurring at the site of the maximum gradient in the surface density.

ORCID iDs

Scott C. Noble https://orcid.org/0000-0003-3547-8306
Julian H. Krolik https://orcid.org/0000-0002-2995-7717
Manuela Campanelli https://orcid.org/0000-0002-8659-6591
Yosef Zlochower https://orcid.org/0000-0002-7541-6612
Hiroyuki Nakano https://orcid.org/0000-0001-7665-0796
Miguel Zilhão https://orcid.org/0000-0002-7089-5570

References

Abramowicz, M. A., Lanza, A., & Percival, M. J. 1997, ApJ, 479, 179
Ajith, P., Boyle, M., Brown, D., et al. 2012, CQGra, 29, 124001
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Baker, J., Haiman, Z., Rossi, E. M., et al. 2019, BAAS, 51, 123, https://baas.aas.org/pub/2020n3i123
Balbus, S. A., & Hawley, J. F. 1998, RvMP, 70, 1
Bankert, J., Krolik, J. H., & Shi, J. 2015, ApJ, 801, 114
Beckwith, K., Hawley, J. F., & Krolik, J. H. 2008, ApJ, 678, 1180
Blanchet, L., Faye, G., & Ponsot, B. 1998, PhRvD, 58, 124002
Bollimpalli, D. A., Mahmoud, R., Done, C., et al. 2020, MNRAS, 496, 3808
Bowen, D. B., Mewes, V., Campanelli, M., et al. 2018, ApJL, 853, L17
Bowen, D. B., Mewes, V., Noble, S. C., et al. 2019, ApJ, 879, 76
Charisi, M., Bartos, I., Haiman, Z., et al. 2016, MNRAS, 463, 2145
Chen, Y.-C., Liu, X., Liao, W.-T., et al. 2020, MNRAS, 499, 2245
d'Ascoli, S., Noble, S. C., Bowen, D. B., et al. 2018, ApJ, 865, 140
Derdzinski, A., D'Orazio, D., Duffell, P., Haiman, Z., & MacFadyen, A. 2021, MNRAS, 501, 3540
Derdzinski, A. M., D'Orazio, D., Duffell, P., Haiman, Z., & MacFadyen, A. 2019, MNRAS, 486, 2754
Dittmann, A., & Ryan, G. 2021, ApJ, 921, 71
D'Orazio, D. J., Haiman, Z., Duffell, P., MacFadyen, A., & Farris, B. 2016, MNRAS, 459, 2379
D'Orazio, D. J., Haiman, Z., & MacFadyen, A. 2013, MNRAS, 436, 2997
Duffell, P. C., D'Orazio, D., Derdzinski, A., et al. 2020, ApJ, 901, 25
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, 446, L36
Farris, B. D., Gold, R., Paschalidis, V., Etienne, Z. B., & Shapiro, S. L. 2012, PhRvL, 109, 221102
Farris, B. D., Liu, Y. T., & Shapiro, S. L. 2011, PhRvD, 84, 024024
Fontecilla, C., Lodato, G., & Cuadra, J. 2020, MNRAS, 499, 2836
Fujita, R., Sago, N., & Nakano, H. 2018, CQGra, 35, 027001
Gammie, C. F. 1996, ApJ, 457, 355
Gammie, C. F., McKinney, J. C., & Tóth, G. 2003, ApJ, 589, 444
Gold, R. 2019, Galax, 7, 63
Gold, R., Paschalidis, V., Etienne, Z. B., Shapiro, S. L., & Pfeiffer, H. P. 2014a, PhRvD, 89, 064060
Gold, R., Paschalidis, V., Ruiz, M., et al. 2014b, PhRvD, 90, 104030
Gole, D., Simon, J. B., Lubow, S. H., & Armitage, P. J. 2016, ApJ, 826, 18
Graham, M. J., Djorgovski, S. G., Stern, D., et al. 2015, Natur, 518, 74
Hawley, J. F., Guan, X., & Krolik, J. H. 2011, ApJ, 738, 84
Hawley, J. F., & Krolik, J. H. 2002, ApJ, 566, 164
Heath, R. M., & Nixon, C. J. 2020, A&A, 641, A64
Hogg, J. D., & Reynolds, C. S. 2016, ApJ, 826, 40
Johnson-McDaniel, N. K., Yunes, N., Tichy, W., & Owen, B. J. 2009, PhRvD, 80, 124039
Katz, M. L., Kelley, L. Z., Dosopoulou, F., et al. 2020, MNRAS, 491, 2301
Kelley, L., Charisi, M., Burke-Spolaor, S., et al. 2019, BAAS, 51, 490, https://baas.aas.org/pub/2020n3i490
Keppler, M., Penzlin, A., Benisty, M., et al. 2020, A&A, 639, A62
Klein, A., Barausse, E., Sesana, A., et al. 2016, PhRvD, 93, 024003
Krolik, J. H. 1999, Active Galactic Nuclei: From the Central Black Hole to the Galactic Environment (Princeton, NJ: Princeton Univ. Press)
Liao, W.-T., Chen, Y.-C., Liu, X., et al. 2021, MNRAS, 500, 4025
Liu, T., Gezari, S., Ayers, M., et al. 2019, ApJ, 884, 36
Liu, T., Gezari, S., Heinis, S., et al. 2015, ApJL, 803, L16
Liu, T., Gezari, S., & Miller, M. C. 2018, ApJL, 859, L12
Lopez Armengol, F. G., Combi, L., Campanelli, M., et al. 2021, ApJ, 913, 16
MacFadyen, A. I., & Milosavljević, M. 2008, ApJ, 672, 83
Mangiagli, A., Klein, A., Bonetti, M., et al. 2020, PhRvD, 102, 084056
Miranda, R., Muñoz, D. J., & Lai, D. 2017, MNRAS, 466, 1170
Moody, M. S. L., Shi, J.-M., & Stone, J. M. 2019, ApJ, 875, 66
Mösta, P., Taam, R. E., & Duffell, P. C. 2019, ApJL, 875, L21
Muñoz, D. J., & Lai, D. 2016, ApJ, 827, 43
Muñoz, D. J., Lai, D., Kratter, K., & Miranda, R. 2020, ApJ, 889, 114
Muñoz, D. J., & Lithwick, Y. 2020, ApJ, 905, 106
Muñoz, D. J., Miranda, R., & Lai, D. 2019, ApJ, 871, 84
Mundim, B. C., Nakano, H., Yunes, N., et al. 2014, PhRvD, 89, 084008
Noble, S. C., Gammie, C. F., McKinney, J. C., & Del Zanna, L. 2006, ApJ, 641, 626
Noble, S. C., & Krolik, J. H. 2009, ApJ, 703, 964
Noble, S. C., Krolik, J. H., & Hawley, J. F. 2009, ApJ, 692, 411
Noble, S. C., Krolik, J. H., & Hawley, J. F. 2010, ApJ, 711, 959
Noble, S. C., Mundim, B. C., Nakano, H., et al. 2012, ApJ, 755, 51
Paschalidis, V., Bright, J., Ruiz, M., & Gold, R. 2021, ApJL, 910, L29
Ragusa, E., Alexander, R., Calcino, J., Hirsh, K., & Price, D. J. 2020, MNRAS, 499, 3362
Ragusa, E., Dipierro, G., Lodato, G., Laibe, G., & Price, D. J. 2017, MNRAS, 464, 1449
Ragusa, E., Lodato, G., & Price, D. J. 2016, MNRAS, 460, 1243
Reynolds, C. S., & Miller, M. C. 2009, ApJ, 692, 869
Roedig, C., Krolik, J. H., & Miller, M. C. 2014, ApJ, 785, 115
Sago, N., Fujita, R., & Nakano, H. 2016, PhRvD, 93, 104023

Sano, T., Inutsuka, S.-i., Turner, N. J., & Stone, J. M. 2004, ApJ, 605, 321
Sesana, A., Roedig, C., Reynolds, M. T., & Dotti, M. 2012, MNRAS, 420, 860
Shi, J.-M., & Krolik, J. H. 2015, ApJ, 807, 131
Shi, J.-M., & Krolik, J. H. 2016, ApJ, 832, 22
Shi, J.-M., Krolik, J. H., Lubow, S. H., & Hawley, J. F. 2012, ApJ, 749, 118
Shiokawa, H., Gammie, C. F., & Doeleman, S. S. 2017, ApJ, 846, 29
Sorathia, K. A., Reynolds, C. S., Stone, J. M., & Beckwith, K. 2012, ApJ, 749, 189
Szilagyi, B., Blackman, J., Buonanno, A., et al. 2015, PhRvL, 115, 031102
Tiede, C., Zrake, J., MacFadyen, A., & Haiman, Z. 2020, ApJ, 900, 43
Tóth, G. 2000, JCoPh, 161, 605
Yunes, N., & Berti, E. 2008, PhRvD, 77, 124006
Yunes, N., & Tichy, W. 2006, PhRvD, 74, 064013
Yunes, N., Tichy, W., Owen, B. J., & Bruegmann, B. 2006, PhRvD, 74, 104011
Zhang, Z., Yunes, N., & Berti, E. 2011, PhRvD, 84, 024029
Zhu, X.-J., & Thrane, E. 2020, ApJ, 900, 117
Zilhão, M., Noble, S. C., Campanelli, M., & Zlochower, Y. 2015, PhRvD, 91, 024034
Zrake, J., Tiede, C., MacFadyen, A., & Haiman, Z. 2021, ApJL, 909, L13