1. INTRODUCTION

Physics accretion disks have been one of the most important topics in astrophysics since the pioneering works of Shakura & Sunyaev (1973), Novikov & Thorne (1973), and Lynden-Bell & Pringle (1974, hereafter LBP74). Historically, a lot of attention has been paid to understanding the properties of disks in which the mass accretion rate $\dot{M}$ is constant with radius. But other varieties of disks exist as well, in particular the so-called circumbinary disks—gaseous disks orbiting a central binary, which can be a stellar binary or a pair of supermassive black holes (SMBHs) in the centers of galaxies. The latter type of systems has recently attracted much attention since the tidal interaction of an SMBH binary with the disk removes angular momentum from the former, resulting in its inspiral. This process may help resolve the so-called last parsec problem (Yu 2002; Lodato et al. 2009)—the stalling of SMBH binaries at separations of $10^{-3} - 1$ pc caused by the inefficiency of stellar dynamical processes at shrinking their orbits, even though some recent work (Khan et al. 2011; Preto et al. 2011) has suggested that dynamical processes are effective, even at small radii, as long as the galaxy is not axisymmetric. Such circumbinary disks are the main focus of this work.

Depending on the mass ratio of the binary components, different modes of tidal coupling of the binary with the disk are possible. When the secondary-to-primary mass ratio $q \equiv M_s / M_p$ is very small, the secondary cannot perturb the disk significantly and migrates through it in the so-called Type I migration regime familiar from the studies of protoplanetary disk–planet interactions (Ward 1997). At higher mass ratios, the secondary becomes capable of clearing gas from the annulus around its orbit, switching its orbital evolution into the so-called Type II migration regime. As $q$ approaches unity, the gap around the orbit of the secondary turns into a central cavity that the binary resides within. Numerical simulations (MacFadyen & Milosavljević 2008; Cuadra et al. 2009) typically show that for $q \sim 1$, the size of the cavity is about twice the semi-major axis of the binary orbit.

Deposition of angular momentum of the density waves excited by the binary causes non-trivial evolution of the circumbinary disk. The seminal study of LBP74, followed by works of, e.g., Lightman & Eardley (1974) and Lin & Papaloizou (1996), has shown that the behavior of disks evolving under the action of both internal viscosity and external torque can be described by the simple equation:

$$\frac{\partial \Sigma}{\partial t} = -\frac{1}{r} \frac{\partial}{\partial r} \left[ l \left( \frac{\rho \Sigma \partial \Omega}{r} \right) - 2 \Sigma \Lambda \right], \quad (1)$$

where $\Lambda$ is the external torque per unit mass of the disk. Here, $\Sigma$ and $v$ are the surface density and kinematic viscosity of the disk, $\tau$ and $r$ are time and radius, and $l = \Omega(r) r^2$ is the specific angular momentum for a circular orbit. Equation (1) provides a one-dimensional, azimuthally averaged description of the disk evolution and is computationally very efficient compared to the direct two-dimensional or three-dimensional hydrodynamical simulations. It forms a basis for understanding the long-term evolution of circumbinary disks and SMBH binary inspiral (Lodato et al. 2009; Chang et al. 2010).

In this work, we critically reanalyze the problem of gas-assisted SMBH binary evolution and provide recipes for improved treatment of this problem. Our work is structured as follows. In Section 2, we describe our setup, derive governing equations, discuss some existing results, and obtain steady state equations, discuss some existing results, and obtain steady state
solutions. Following LBP74 and Lyubarskij & Shakura (1987), we reformulate Equation (1) in terms of the viscous angular momentum flux (rather than surface density), which allows a very straightforward interpretation of the steady state disk structure and provides a transparent way of understanding the evolution of circumbinary disks. In Section 4.1, we provide a description of the physical properties of disks around SMBH binaries, parameterizing them via the viscous angular momentum flux rather than the mass accretion rate $M$. In Section 5, we discuss evolution of circumbinary disks, in particular deriving the self-similar solutions to an evolving disk structure. All of these results are then used in Section 6 to describe the coupled evolution of an SMBH binary and a circumbinary disk, including the self-consistent time variation of the disk properties, the orbital inspiral of the binary components, and the electromagnetic manifestations of the system. Our conclusions are summarized in Section 7.

2. PROBLEM SETUP AND EVOLUTION EQUATION

We consider a binary consisting of two point masses $M_1$ and $M_2$ (the total mass is $M = M_1 + M_2$), moving around common barycenter. For simplicity, the binary eccentricity is set to zero, even though simulations show the possibility of eccentricity growth due to tidal binary–disk coupling (Roedig et al. 2011). The binary is surrounded by a prograde, coplanar disk (cf. Nixon et al. 2011), which extends to much larger distances than the binary semi-major axis $r_b$. We will assume that the potential in which the disk orbits is faithfully represented by a Newtonian potential produced by a combined mass $M_c$.

We assume that the binary clear out a central cavity of radius $r_{in}$ in the disk, which requires its mass ratio $q \equiv M_2/M_1$ to be high enough to prevent viscous revilling of the gap (or cavity). We determine the conditions in which overflow across the orbit of the secondary (Kocsis et al. 2012a, 2012b) is possible in Section 6.2.7. The width of the gas-depleted annulus between the orbit of the secondary and $r_{in}$ is $\Delta \equiv |r_{in} - r_b| \lesssim r_b$ for $q \ll 1$, while $\Delta \sim r_b$ for $q \sim 1$ (MacFadyen & Milosavljević 2008). For simplicity, we assume that neither the primary nor the secondary have their own disks (this simplification can be easily relaxed).

An external torque $\Lambda$ due to tides raised by the central binary is concentrated right at the edge of the gap (or cavity) around the orbit of the secondary. This expectation is borne out in calculations of the torque density distribution both in uniform disks, where it is generally found that $\Lambda \propto |r - r_b|^{-4}$ (Goldreich & Tremaine 1980; Armitage & Natarajan 2002), and in nonuniform disks, where it has been shown by Petrovich & Rafikov (2012) that $\Lambda$ decays exponentially near the disk edge. As a result, already at small distances away from $r_{in}$, the external torque can be neglected: Equation (1) then reduces to its classical form obtained by LBP74. The effect of the binary torque is then incorporated into the solution of this simplified equation via the boundary condition (BC, discussed in Section 2.1) imposed at the inner edge of the disk.

The evolution of an accretion disk is best illustrated if we characterize its properties at each radius not by the surface density $\Sigma$ but by the viscous angular momentum flux $F_j$, defined as

$$F_j \equiv -2\pi \nu \Sigma r^3 \frac{d\Omega}{dr} = 3\pi \nu \Sigma l,$$

where $l = \Omega r^2$ is the specific angular momentum in a Newtonian potential with $\Omega = (GM_c/r^3)^{1/2}$. This quantity represents the viscous torque exerted by the part of the disk interior to a given radius $r$ on the external part of the disk, and is thus equal to the amount of angular momentum crossing the disk circumference $2\pi r$ per unit time due to the action of viscosity.

In the absence of external torques, the mass accretion rate through the disk can be directly expressed (LBP74) through the divergence of the viscous angular momentum flux as (note that we take $M > 0$ for mass inflow toward the center of the system):

$$\dot{M}(r) = (\frac{dl}{dr})^{-1} \frac{dF_j}{dr} = \frac{\dot{F}_j}{\dot{\ell}},$$

motivating us to change the independent variable from $r$ to $l$. In this case, the evolution Equation (1) with $\Lambda = 0$ takes on a particularly simple form (LBP74; Filipov 1984; Lyubarskij & Shakura 1987):

$$\frac{\partial}{\partial t} \left( F_j \right) = \frac{\partial^2 F_j}{\partial l^2},$$

where the function similar to the diffusion coefficient

$$D_f \equiv -\nu r^2 \frac{d\Omega}{dl} \frac{dl}{dr}$$

is in general a function of both $l$ and $F_j$ because of the possible dependence of $\nu$ on $F_j$.

Despite the mathematical simplicity of the evolution of Equation (4), which was first recognized by LBP74, over the years it has become conventional to study disk evolution using the more complicated Equation (1), see, e.g., Lodato et al. (2009) and Chang et al. (2010). We will show that the use of Equation (4) offers significant advantages over the standard approach, in particular for obtaining the steady state solutions for the disk structure.

2.1. Boundary Conditions

Using Equation (1) and the continuity equation, one finds

$$-\dot{M} = \left( \frac{\partial \dot{F}_j}{\partial r} \right) - \frac{\dot{F}_j}{\dot{\ell}} + 4\pi \frac{\Sigma \Lambda}{\Omega}.$$

Since $\Lambda$ is significant only in a narrow annulus at the inner edge of the disk, we can assume that $\Lambda = 0$ outside of some radius $r_\Lambda$, which is not too different from $r_{in}$ (to be specific, one can, for example, take $r_\Lambda$ to be the radius interior to which the binary exerts 90% of its torque on the disk). Multiplying Equation (6) by $dl/dr$ and integrating between $r_{in}$ and $r_\Lambda$, one obtains:

$$F_j(r_\Lambda) = 2\pi \int_{r_{in}}^{r_\Lambda} r\Lambda \Sigma dr + \frac{1}{2} \int_{r_{in}}^{r_\Lambda} \dot{M}\Omega dr,$$

where we set $F_j(r_{in}) = 0$ because $\Sigma(r_{in}) = 0$ based on our assumption of a clean gap.

The first integral on the right-hand side of Equation (7) is the total torque that the binary exerts on the disk. As long as the orbital evolution of the binary is driven predominantly by tidal coupling to the disk (and not by the gravitational wave emission), angular momentum conservation ensures that this term is equal to:

$$\frac{dL_b}{dt} = -\frac{L_b}{2r_b} v_b,$$
where $L_b = M_b (G M_r r_b)^{3/2} q/(1 + q)^2$ is the orbital angular momentum of the binary and $v_b \equiv d r_b / d t$ is its inspiral speed.

In the second integral in Equation (7), one can write $M = 2 \pi \Sigma v_r$, where $v_r$ is the radial velocity of the gas, and approximate $v_r \sim v_b$ in the annulus between $r_b$ and $r_h$, thus assuming that gas closely follows the shrinkage of the binary orbit. Then, as long as the mass of the secondary $M_s$ satisfies the condition

$$M_s \gtrsim \frac{1}{\Omega(r_b) r_b} \int_{r_b}^{r_h} \Sigma \Omega r^2 dr,$$

(9)

the second term in the right-hand side of Equation (7) is smaller than the first term. This condition is often replaced by demanding that the “local disk mass” $M_d = \Sigma r^2$ in the vicinity of the binary (i.e., at $r \sim r_b$) be less than the mass of the secondary, i.e. (Syer & Clarke 1995, hereafter SC95; Haiman et al. 2009)

$$\frac{M_d}{M_s} = \frac{\Sigma(r_b) r_b^2}{M_s} \ll 1.$$  

(10)

Whenever the disk is in the “secondary-dominated” regime defined by the conditions (9) or (10), the viscous angular momentum flux in its inner region can be directly related to the orbital evolution of the binary:

$$F_J(r_m) = -\frac{d L_b}{dt}.$$  

(11)

Here, we replaced $r_b$ with $r_h$ since the two radii are very similar (and also very close to $r_h$). On the contrary, if the condition (9) is not fulfilled the inner part of the disk absorbs most of the angular momentum brought in by viscous torques and Equation (11) becomes invalid. In this case, a more general BC in the form (7) must be employed.

Using definition (2) and the expression for $L_b$, one can rewrite Equation (11) as a formula for the orbital evolution time of the binary $t_J = |d \ln r_b / dt|^{-1}$ when disk torques dominate:

$$t_J = \frac{L_b}{2 F_J(r_m)} \approx \frac{t_v}{6 \pi(1 + q)} \frac{M_s}{M_b},$$  

(12)

where we defined the viscous timescale $t_v \equiv r^2/v$, evaluated at $r_h$. Similar expressions (up to a constant factor) have previously been quoted in the literature (Lodato et al. 2009; Baruteau & Masset 2013) on Type II migration in the secondary-dominated limit (Equation (10)). A subtle point in this expression is that both $t_v$ and $M_s$ vary in time as the disk evolves even if the orbit of the secondary does not change appreciably. A self-consistent calculation of this evolution is one of the goals of our present work.

In our work, we normally take the disk at large separations to be a standard constant $M$ disk with the mass supply rate $M_\infty$, so that

$$F_J(t, l \to \infty) \to M_\infty l.$$  

(13)

The inner BC is specified in the form of a constraint on the mass accretion rate $\dot{M}$, which in general is not equal to $M_\infty$ close to the binary because of the tidal barrier. Binary torques do not have to stop the mass inflow completely: if the gap cleared out by the secondary is not deep/broad enough to present a serious obstacle to the gas inflow, a fraction of mass arriving at the inner edge of the disk crosses the orbit of the secondary and is accreted by one of the binary components (Kocsis et al. 2012a, 2012b).

Alternatively, gas can be removed from the circumbinary disk in the form of a wind; see Sections 4.1.4 and 6.2.6. To account for this possibility, we generally use an inner BC in the form

$$\frac{\partial F_J}{\partial l} \mid_{l=l_m} = \dot{M}(l_m) = \chi M_\infty.$$  

(14)

We refer the reader to Equation (3), where $l_m$ is the value of $l$ at the inner disk radius $r_m$, which we will associate with the binary semi-major axis $r_b$ for simplicity. Here, $\chi \ll 1$ is assumed to be constant, allowing $\dot{M}(l_m)$ to be less than $M_\infty$. In practice, the values of $\dot{M}(l_m)$ and $\chi$ are set by the strength of the tidal barrier (Liu & Shapiro 2010) and may vary in time as the system evolves.

Clearly, $\chi = 0$ implies no gas inflow across the orbit of the secondary, and this situation is what we will often consider in this work. For $\chi = 1$, the secondary does not present any barrier to the mass inflow and the disk structure reduces to that of a constant $M = M_\infty$ disk.

2.2. Initial Setup

At $t = 0$, we take the disk outside the binary orbit to be a standard, constant $M = M_\infty$ disk. This setup is naturally realized if a massive cloud of gas falls into the center of a galaxy and circularizes outside the SMBH orbit. Subsequent viscous evolution forms a constant $M$ disk within the circularization radius. We identify $t = 0$ as the moment when the inner edge of the disk reaches the binary orbit and gas inflow is slowed down by the binary torques. Such a setup allows us to follow the complete evolution of the system until the SMBH binary merger, without making simplifying (quasi-) steady state assumptions.

It is important that we also assume that the system starts in the secondary-dominated regime (see Section 6.2.4), so that the condition (10) is fulfilled at $t = 0$. This procedure is different from the setup of, e.g., Haiman et al. (2009), who start their system in the disk-dominated regime, when condition (10) is violated. These authors then follow the transition to the secondary-dominated limit.

2.3. Some Existing Results

We now summarize a number of existing results relevant for our present work and circumbinary disk evolution in general. Pringle (1991) explored viscous evolution of a ring of gas subject to injection of angular momentum at the center, which can be thought of as a model of a circumbinary disk of a finite radial extent. Later, SC95 looked for a steady state solution to circumbinary disk structure using as an inner BC Equation (11) coupled with the requirement $v_r \to v_b$ as $r \to r_m$. This solution implies that $M = M_\infty$, which in fact is never fulfilled in a secondary-dominated regime, as we demonstrate in Appendix A. This solution also explicitly depends on time via the dependence on $r_b(t)$ (and the torque on the binary is found to depend on the mass of the secondary), so that the disk properties end up varying on the migration timescale of the secondary. However, as the secondary spirals in, the migration timescale is in fact much shorter than the viscous timescales of the disk far from the binary, which we demonstrate in Appendix A. As a result, viscous transport in the disk is unable to communicate information about the changing inner BC outside the immediate vicinity of the inner edge of the disk, implying an internal inconsistency in the quasi-steady solution of SC95.

Ivanov et al. (1999, hereafter IPP) identified these problems with the work of SC95 and instead proposed a time-dependent,
self-similar solution to the circumbinary disk evolution using BCs analogous to our Equation (14) with $\chi = 0$. The work of IPP is laid out in terms of $\Sigma$, using Equation (1), even though previously Filipov (1984) and Lyubarskij & Shakura (1987) already constructed self-similar disk solutions in terms of $F_J$ in applications to compact X-ray sources.

Many theoretical studies (Haiman et al. 2009; Kocsis et al. 2011; Yunes et al. 2011) use both the IPP and SC95 solutions to explore gas-assisted SMBH binary evolution, despite the aforementioned issues with the SC95 work; see Appendix A. This methodology has certainly affected the conclusions of these works regarding the evolutionary paths and detectability of SMBH binaries in both the electromagnetic and gravitational wave (GW) domains.

The specific goals of our present study are: (1) to understand the general features of circumbinary disk evolution by providing improved and more general solutions for the disk properties and (2) to apply these results for a better understanding of the gas-driven SMBH binary evolution. Following Pringle (1991) and IPP, we gain significant insight by constructing self-similar solutions for disk evolution in Section 5.1 but do not limit ourselves to only the self-similar evolution; see Section 5.3. Analogous to Filipov (1984), Lyubarskij & Shakura (1987), and Lipunova & Shakura (2000), we describe the disk evolution in terms of $F_J$ rather than $\Sigma$, but our solutions are different because of the different BCs. In Section 4, we compute disk properties in different physical regimes as a function of $F_J$, generalizing similar calculations in Haiman et al. (2009), which are applicable only to constant $M$ disks and are not appropriate for the circumbinary disks because of the ambiguity in relating $F_J$ and $M$; see Section 3. In the no-overflow case ($\chi = 0$), we confirm the self-similar results of IPP (obtained in terms of $\Sigma$ rather than $F_J$), but also extend these results to the case with overflow and directly apply them to gas-assisted SMBH binary evolution in Section 6.

3. STEADY STATE SOLUTION

Equation (4) clearly admits a simple steady state solution (LBP74):

$$F_J(l) = F_{J,0} + F_{J,1}l,$$

where $F_{J,0}$ and $F_{J,1}$ are constants and $l$ is the specific angular momentum. This solution is completely independent of the detailed physics that determines the disk properties, since for any (even highly nonlinear) dependence of $D_J$ on $F_J$ and $l$, the solution (15) still satisfies Equation (4) in the steady state. According to Equation (3), this solution implies that $M = F_{J,1}$ is also constant, which seems to suggest that Equation (15) corresponds to the conventional accretion disk with a constant $M$. This suggestion is, however, not true in general.

Shakura & Sunyaev (1973) explored the physics of constant $M$ accretion disks affected only by internal viscous stresses all the way to the central object. They showed in particular that in such disks the local surface density $\Sigma(r)$ at each radius is related to the mass accretion rate via

$$\dot{M} = 3\pi v \Sigma,$$

while the energy loss per unit surface area of the disk scales as

$$\sigma T_{\text{eff}}^4(r) = \frac{3}{8\pi} \dot{M} \Omega^2.$$

Equation (16) is compatible with the definition (2) and solution (15) only if $F_J = Ml$. However, over the years, the concept of constant $\dot{M}$ disks has evolved to essentially imply disk properties given by Equations (16) and (17). In the rest of this work, we will call such disks standard constant $\dot{M}$ disks.

Let us now consider a steady state solution of the form given by Equation (15) with $F_{J,0} \neq 0$ and $F_{J,1} = M = \text{const}$. Combining Equations (2) and (15), one finds:

$$M = 3\pi v \Sigma - F_{J,0}l^{-1},$$

which reduces to Equation (16) only if $F_{J,0} = 0$. In particular, it follows directly from Equation (3) that it is possible to have a steady disk with $M = 0$ as long as $F_J$ is independent of radius. In such a steady state disk with $F_J = \text{const}$, the disk surface density is related to $F_J$ via

$$F_J = F_{J,0} = 3\pi v \Sigma(r),$$

which replaces Equation (16).

Similarly, the viscous energy dissipation rate in the disk per unit radius (and per unit time), $d\dot{E}/dr$, is given by:

$$\frac{d\dot{E}}{dr} = -F_J \frac{d\Omega}{dr} = \frac{3}{2} \frac{F_J \Omega}{r}.$$

Since $d\dot{E}/dr = 4\pi \sigma T_{\text{eff}}^4(r)$, we can write

$$\sigma T_{\text{eff}}^4(r) = \frac{3}{8\pi} \frac{F_J \Omega}{r^2},$$

which in the steady state described by the solution (15) yields

$$\sigma T_{\text{eff}}^4(r) = \frac{3}{8\pi} \left[ \dot{M} \Omega^2(r) + \frac{F_J \Omega}{r^2} \right].$$

Again, this expression reduces to the conventional result (17) only if $F_{J,0} = 0$. In the case of a $M = 0$ disk with $F_J = F_{J,0}$, one finds

$$T_{\text{eff}}(r) = \left( \frac{3}{8\pi} \frac{F_J \sqrt{GMc}}{\rho} \right)^{1/4} r^{-7/8},$$

so that $T_{\text{eff}}(r)$ increases toward small radii more steeply than in a standard constant $M$ disk (for which $T_{\text{eff}}(r) \propto r^{-3/4}$). This result was first obtained by SC95.

Equation (21) predicts a non-zero $T_{\text{eff}}$ for a disk with $F_J = \text{const}$, even though $M = 0$ in such a disk. Since there is no inward mass flow in this disk and the corresponding release of potential energy is absent, one may naturally wonder where the energy emitted from the disk surface comes from. Integrating Equation (20) between arbitrarily chosen inner and outer radii $r_i$ and $r_o$, one finds the global energy release between these radii:

$$\dot{E}(r_i < r < r_o) = F_J [\Omega(r_i) - \Omega(r_o)],$$

where we have used the fact that $F_J$ is independent of $r$. Thus, the global rate of energy generation by viscous dissipation is equal to the work done on the disk by the viscous stress at its inner and outer edges. The latter ultimately provides the energy source for the radiation from the disk surface.

To summarize, the assumption of a steady state does not necessarily require $M$ and $T_{\text{eff}}$ to be given by Equations (16) and (17)---these statements are sufficient but not necessary characteristics of steady disks. The most general steady state is in fact described by Equation (15); it then follows that $M$ and $T_{\text{eff}}$ must be given by Equations (18) and (22). We explore the properties of such “non-standard” steady disks as a part of this work.
4. DISK PROPERTIES AS A FUNCTION OF $F_J$

It is conventional to describe the structure of steady accretion disks in different physical regimes using $M_R$ assumed to be constant with radius, as a free parameter (e.g., Shakura & Sunyaev 1973). Such calculations (Goodman 2003; Haiman et al. 2009) universally assume that the relations (16) and (17) hold true, i.e., that $F_{J,0} = 0$ in Equation (15). Our demonstration in Section 3 that in general, a constant $M_R$ does not unambiguously determine $F_J$ implies that such scalings cannot be used for circumbinary disks in general. We now revise and generalize scalings of disk properties by expressing them in terms of $F_J$ rather than $M_R$. Note that these scalings are completely general and do not assume $F_J$ to be constant in time and/or radius.

We assume the disk to be optically thick, but it is easy to extend our calculations to optically thin disks as well. Physical conditions in disks vary with radius, resulting in transitions in the opacity behavior. To account for this effect, we derive in Appendix B a set of scaling relations applicable to the gas pressure-dominated disks with a power-law opacity behavior and apply these relations in Section 4.1 to disks around SMBH binaries.

4.1. Circumbinary Disks around SMBH Binaries

In the inner parts of a circumbinary disk around a compact SMBH binary, the radiation pressure $p_r = aT^4/3$ can be very important. We parameterize the role of radiation pressure via the dimensionless ratio of the gas pressure $p_g/p$ to the total pressure $p_g/p = p_g/(p_r + p_g)$. In the radiation-pressure-dominated case $\beta \ll 1$, while in the gas pressure-dominated regime $1 - \beta \ll 1$.

There is still an ongoing debate as to whether the viscosity in the radiation pressure-dominated fluid is determined by the full pressure or just the gas pressure. For this reason, we use a prescription motivated by the conventional $\alpha$-parameterization of Shakura & Sunyaev (1973) and account for the two possibilities in a convenient form (Goodman 2003):

$$\nu = \alpha \beta^{3/2} \frac{k^2}{\Omega}, \quad (25)$$

Here, $\alpha = 0$ corresponds to a kinematic viscosity proportional to the full pressure $p$, while $\alpha = 1$ corresponds to the case when only the gas pressure $p_g$ determines the viscosity. Lightman & Eardley (1974) suggested that radiation pressure-dominated disks with $\alpha = 0$ are thermally unstable but recent numerical work by Hirose et al. (2009a, 2009b) does not show this result to be the case.

Disks around SMBH binaries are heated by internal viscous dissipation and also by energy deposition by the density waves launched by the central binary. For simplicity, we do not consider here the latter contribution (its role has been studied by Lodato et al. 2009)—it may be important only near the inner edge of the disk, where waves are dissipated, and should not affect disk thermodynamics far from the binary. Thus, we assume that disk heating is described by Equations (20) and (21). Using Equation (23), we estimate the effective temperature of the disk as

$$T_{\text{eff}}(r) = \left( \frac{3}{8\pi} \frac{F_J \sqrt{GM_R}}{\sigma} \right)^{1/4} r^{-7/8} \approx 1.1 \times 10^3 K \left( \frac{F_{J,50}}{M_{c,7}} \right)^{1/8} r^{-7/8},$$

where $M_{c,7} \equiv M_c/(10^7 M_\odot)$, $r_n \equiv r/(10^n$ pc), and $F_{J,n} \equiv F_J/(10^n$ erg). This estimate assumes a particular value of $F_J$, which we motivate in Section 4.1.4.

In our treatment of the vertical radiation transfer in the disk, we follow Goodman (2003) and relate the midplane disk temperature $T$ to $T_{\text{eff}}$ via

$$T^4 = \frac{\tau}{2} T_{\text{eff}}^4,$$  \hspace{1cm} (27)

typical for optically thick ($\tau \gg 1$) disks, where $\tau = \kappa \Sigma$ is the optical depth. Also,

$$\Sigma \gg 2 \rho h,$$  \hspace{1cm} (28)

where $\rho$ is the characteristic (midplane) density, $h = c_s/\Omega$ is the scale height, and $c_s = (\rho/\rho)^{1/2}$ is the sound speed, determined by the total pressure $p$. In the innermost region of the disk where the transition between the radiation and gas pressure-dominated regimes occurs, we assume that the opacity is dominated by electron scattering, i.e., $\kappa = \kappa_{\text{es}} \approx 0.4 \text{ cm}^2 \text{ g}^{-1}$. Further out, in the gas pressure-dominated part of the disk, $\kappa$ is determined by the free–free opacity. We separately consider all these regimes below.

Combining Equations (2), (21), (25), (27), and (28), one finds the following equation determining the value of $\beta$ for a disk with $\kappa = \kappa_{\text{es}}$:

$$\frac{\beta^{4+b}}{(1 - \beta)^{10}} = 2^6 \pi^8 \frac{k}{\mu} \left( \frac{m_e \sigma}{\kappa_{\text{es}} \sigma F_J^8} \right)^{1/2} (GM_R)^{1/2} r^{20/2}.$$

One can find the distance $r^{\text{rad/gas}}$ of the transition between the radiation and gas pressure-dominated regimes by setting $\beta = 1/2$ in Equation (29):

$$r^{\text{rad/gas}} = \left[ 2^{6} \pi^{8} \left( \frac{\mu}{k} \right)^{4} \left( \frac{F_{J,50}}{M_{c,7}} \right)^{1/2} \right]^{2/29} \approx 3.8 \times 10^{-3} \text{ pc} \left[ 2^{6} \pi^{8} \left( \frac{\mu}{k} \right)^{4} \left( \frac{F_{J,50}}{M_{c,7}} \right)^{1/2} \right]^{2/29},$$

where $\alpha_n \equiv \alpha/10^n$ and $\mu_n \equiv nm_p$, i.e., $\mu$ in this formula is normalized by the molecular weight of the fully ionized H. Note the extremely weak dependence of $r^{\text{rad/gas}}$ on the binary mass $M_c$.

4.1.1. Radiation Pressure Dominated Regime

Interior to $r^{\text{rad/gas}}$, the disk is radiation pressure-dominated and $\kappa = \kappa_{\text{es}}$. In this case, one finds

$$\frac{h(r)}{r} = \frac{1}{2\pi} \frac{\kappa_{\text{es}}}{c(1 - \beta)} \frac{F_J}{\Omega r^3} \approx 1.1 \times 10^{-3} F_{J,50} M_{c,7}^{-1/2} r_{-7/2},$$

irrespective of the value of $b$ in Equation (25). At the same time, the scalings of $\Sigma(r)$ and $T(r)$ explicitly depend on the viscosity behavior. For $b = 0$:

$$\Sigma(r) = \frac{4\pi}{3} \frac{c^2}{\alpha \kappa_{\text{es}} F_J} r^2 \approx 2.1 \times 10^6 \text{ g cm}^{-2} \left( \frac{\alpha}{a} \right)^{1} F_{J,50}^{-1} r_{-2}^{-2},$$

where $M_{c,7} \approx 10^{-7} M_\odot$.
The Astrophysical Journal scales as \( \kappa \) and definition (2), which are identical in the two cases because involves only the equation of vertical radiation transfer (27) and dominated case with \( \alpha \). Using Equation (34) applicable for the situation described in

\[
T(r) = \left[ \frac{(GM)^{1/2}c^2}{4\alpha \sigma \kappa e} \right]^{1/4} e^{-3/8} \approx 2.9 \times 10^4 \text{K} \alpha_e^{-1} M_{c,7}^{-1/8} r^{-3/8}, \tag{33}
\]

while for \( b = 1 \):

\[
\Sigma(r) = \left[ \frac{2^{4}}{3^{3}\pi^{1}} \left( \frac{\mu}{ak} \right)^{4} \sigma F_{J}^{3} \left( GM_{e}^{1/2} \kappa e \right)^{1/5} r^{-9/10} \right] \approx 6 \times 10^4 \text{g cm}^{-2} \mu \mu_{0} F_{J,50}^{3/10} M_{c,7}^{-1/2} r^{-2/10}, \tag{34}
\]

\[
T(r) = \left[ \frac{1}{2^{3} \pi^{1/2}} \left( \frac{\mu}{\kappa e} \right)^{4} \sigma F_{J}^{3} \left( GM_{e}^{1/2} \kappa e \right)^{1/5} \mu \mu_{0} F_{J,50}^{3/10} M_{c,7}^{-1/2} \right]^{1/5} r^{-11/10} \approx 1.2 \times 10^4 \text{K} \mu \mu_{0} F_{J,50}^{3/10} M_{c,7}^{-1/2} r^{-2/10}. \tag{35}
\]

Note that a radiation pressure-dominated disk with \( F_{J} = \text{const} \) (and \( \alpha = 0 \)) has \( h(r) \propto r^{-1/2} \) (as opposed to the case of a conventional \( M = \text{const} \) disk, for which \( h(r) = \text{const} \)), i.e., the disk puffs up as \( r \) decreases. The surface density shows dramatically different behaviors depending on the value of \( b \): it rises with \( r \) when the viscosity is proportional to the total pressure \( (b = 0) \), but drops with \( r \) when \( v \) scales with the gas pressure. Finally, \( T(r) \) is independent of \( F_{J} \) when \( b = 0 \) but scales as \( \propto F_{J}^{-1} \) for \( b = 1 \).

4.1.2. The Gas Pressure-Dominated Regime with \( \kappa = \kappa_e \)

Outside of \( r \text{rad/gas} \), the gas pressure dominates, meaning that \( \beta \to 1 \) and \( \nu \to \alpha c_{s}^{2} / \Omega \). Initially, however, the opacity is still determined by electron scattering, so \( \kappa = \kappa_e \). It is easy to see that in this case, the midplane temperature and surface density runs in the disk should be the same as in the radiation pressure-dominated case with \( b = 1 \): the derivation of \( T(r) \) and \( \Sigma(r) \) involves only the equation of vertical radiation transfer (27) and definition (2), which are identical in the two cases because \( \nu \) and \( \kappa \) are the same when \( b = 1 \). Thus, the behavior of \( \Sigma(r) \) and \( T(r) \) in the gas pressure-dominated case with \( \kappa = \kappa_e \) is given by Equations (34) and (35), respectively.

However, the behavior of \( h/r \) differs from Equation (31) since, unlike the case studied in Section 4.1.1, the vertical support is now provided by the gas pressure. As a result, one finds:

\[
\frac{h(r)}{r} = \left[ \frac{1}{16\pi^{2}} \left( \frac{k}{\mu} \right)^{4} \frac{\kappa e F_{J}^{2}}{\alpha \sigma (GM_{e})^{1/2}} \right]^{1/10} e^{-1/20} \approx 6.6 \times 10^{-3} \left[ \frac{F_{J,50}^{3/10}}{\alpha^{-1} \mu \mu_{0} F_{J,50}^{3/10} M_{c,7}^{-1/2}} \right]^{1/10} e^{-1/20}. \tag{36}
\]

Note that \( h/r \) decreases with \( r \), meaning that the disk is not flared. However, this dependence on \( r \) is so weak that the aspect ratio is essentially constant with radius.

4.1.3. The Gas Pressure-Dominated Regime with \( \kappa = \kappa_{\beta} \)

At even larger distances, \( \kappa \) is dominated by the free-free opacity \( \kappa_{\beta} = \kappa_{\beta,0} T^{-7/2} \), with \( \kappa_{\beta,0} = 8 \times 10^{32} \text{cm}^{-1} \text{g}^{-2} \text{K}^{-7/2} \). Using Equation (34) applicable for the situation described in Section 4.1.2 and the behavior of the midplane temperature inferred from Equation (36), one finds that the transition between the two opacity regimes takes place at:

\[
r_{\text{trans}} = \frac{3}{8\pi} \frac{\kappa e F_{J}^{2}}{(GM_{e})^{1/2}} \left( \frac{k}{\mu} \right)^{1/4} \approx 8.4 \times 10^{-3} \text{pc} F_{J,50}^{1/2} \mu^{-1/4}. \tag{37}
\]

Outside of this radius, the parameters of the disk scale as:

\[
\Sigma(r) = \left[ \frac{2^{10}}{3^{11}\pi^{14}} \left( \frac{\sigma}{\kappa \mu_{0}} \right)^{2} \left( \frac{\mu}{k} \right)^{15} \frac{F_{J}^{14}}{\alpha^{16}(GM_{e})^{2}} \right]^{1/20} \times r^{-11/10} \approx 7.2 \times 10^{4} \text{g cm}^{-2} \mu \mu_{0} F_{J,50}^{3/10} M_{c,7}^{-1/2} r^{-11/10}, \tag{38}
\]

\[
T(r) = \left[ \frac{2^{-5}}{3^{13}\pi^{3}} \left( \frac{\sigma}{\kappa \mu_{0}} \right)^{2} \left( \frac{\mu}{k} \right)^{15} \frac{F_{J}^{14}}{\alpha^{16}(GM_{e})^{2}} \right]^{1/10} r^{-9/10}, \tag{39}
\]

\[
\frac{h(r)}{r} = \left[ \frac{2^{-5}}{3^{13}\pi^{3}} \left( \frac{\sigma}{\kappa \mu_{0}} \right)^{2} \left( \frac{\mu}{k} \right)^{15} \frac{F_{J}^{14}}{\alpha^{16}(GM_{e})^{2}} \right]^{1/20} \mu^{-1/20}, \tag{40}
\]

We refer the reader to Equations (B2)–(B4) in Appendix B. In this opacity regime, the disk is only weakly flared and \( h/r \) is essentially constant with radius.

4.1.4. Characteristic Values of \( F_{J} \)

We now motivate the characteristic value of the viscous angular momentum flux \( F_{J} \) in disks around SMBH binaries using different arguments. In the case of a standard constant \( M \) disk, a convenient reference value of \( M \) is set by the Eddington mass accretion rate:

\[
\dot{M}_{\text{Edd}} = \frac{4\pi GM_{e}}{c_{\kappa e}} \varepsilon^{-1} \approx 0.2 M_{\odot} \text{yr}^{-1} \frac{M_{c,7}}{10^{9}} \frac{1}{\varepsilon_{0.1}}. \tag{41}
\]

(\( \varepsilon = 0.1 \varepsilon_{0.1} \) is the radiative efficiency), which uniquely relates the maximum value of \( M \) to the mass of the central object \( M_{c} \). However, we have seen in Section 3 that \( M \) does not unambiguously fix the value of \( F_{J} \). Moreover, we will see later that \( F_{J} \) in a disk around an SMBH binary varies in time, making a unique relationship between \( F_{J} \) and \( M \) impossible.

One can still formulate the Eddington limit on the value of \( F_{J} \) by assuming that the limit is reached when the radiation pressure pushes up the disk to such an extent that it becomes geometrically thick, \( h/r \sim 1 \), where a further increase in \( F_{J} \) results in mass loss in a radiation pressure-driven wind. Another possibility for \( h/r \sim 1 \) is mass overflow across the orbit of the secondary high above the midplane where the secondary potential is weaker.\(^5\)
Both processes result in mass loss for the circumbinary disk at its inner edge, which can affect the orbital evolution of the binary, as we demonstrate in Section 5.

Unlike \( M_{\text{Edd}} \) in conventional constant \( \dot{M} \) disks, the Eddington limit on \( F_J \) found from the condition \( h/r = 1 \); see Equation (31):

\[
F_{J,\text{Edd}}(r) = 2\pi \frac{(GM_c)^{1/2}c}{\kappa_{\text{es}} r^{3/2}}.
\]  

(42)

This equation depends on the value of \( r \) at which it is evaluated. It is clear that this condition is most constraining at the inner edge of the disk \( r_{\text{in}} \sim r_b \), where \( r_b \) is the semi-major axis of the SMBH binary.

Obviously, one needs to invoke additional considerations to pick a particular value of \( r_{\text{in}} \) (or \( r_b \)). Here, we assume that this critical value of \( r_{\text{in}} \) is such that at \( r_b \approx r_{\text{in}} \), the orbital decay timescale of the binary due to emission of gravitational waves \( t_{\text{GW}} \) is equal to the characteristic timescale on which the orbit of the binary shrinks due to tidal coupling to the circumbinary disk \( t_J \). The logic behind choosing this condition is that then the disk stays sub-Eddington all the way until the point at which the GW emission becomes more important for the orbital evolution of the binary than its tidal coupling to the disk. Beyond this point, the disk is super-Eddington and loses mass in a radiation-pressure-driven wind, but this process does not affect the binary inspiral since the orbital evolution is no longer sensitive to the tidal torque.

Equating

\[
t_{\text{GW}}(r_b) = \frac{5}{2q_S} \frac{R_S}{c} \left( \frac{r_b}{R_S} \right)^4
\]  

(43)

(here \( q_S \equiv q/(1+q)^2 \) and \( R_S \equiv 2GM_c/c^2 \) is the Schwarzschild radius of the black hole with the mass equal to the total mass of the binary \( M_c \)) and

\[
t_J(r_b) = \left| \frac{d \ln r_b}{dt} \right|^{-1} = \frac{L_b}{2F_J} = q_S M_c \sqrt{GM_b r_b} \frac{\sqrt{GM_c r_b}}{F_J}
\]  

(44)

(here \( L_b = (qs/4)M_c(GM_cr_b)^{1/2} \) is the total orbital angular momentum of the binary), one finds:

\[
r_{\text{in}} = R_S \left( \frac{q_S^2}{25/2^5/2} \frac{M_c c^2}{F_J} \right)^{2/7}.
\]  

(45)

Inserting this value of \( r_{\text{in}} \) into Equation (42), one finds that the Eddington value of \( F_J \) based on the condition \( t_{\text{GW}}(r_{\text{in}}) \approx t_J(r_{\text{in}}) \) is given by:

\[
F_{J,\text{Edd}} = \left[ \frac{2^{10/3}3^{7/2}(GM_c)^{14} M_c^{12/5} c^4}{5^3 e^8 \kappa_{\text{es}}^2} \right]^{1/10}
\]  

\approx 10^{51} \text{ erg} \ M_c^{17/10} q_S^{3/5}.
\]  

(46)

(47)

This argument justifies the adoption of a characteristic value of \( F_J = 10^{50} \text{ erg} \) (corresponding to \( M_c = 10^7 M_\odot \) and \( q = 0.005 \)) in our numerical estimates.

One can come up with other ways of choosing the critical \( r_{\text{in}} \). In particular, one may demand the disk to stay sub-Eddington until the point when \( t_{\text{GW}} \) becomes equal to the viscous timescale \( t_v \) at its inner edge, after which the binary orbit shrinks faster than the viscosity can refill the central cavity, so that the former essentially decouples from the disk (Milosavljević & Phinney 2005). This effect happens at a radius considerably smaller than the \( r_{\text{in}} \) given by Equation (45) simply because \( t_J \) at \( r_{\text{in}} \) is greater than \( t_v \) by a factor \( M_c/M_d \) since disk torques still dominate there (assuming that the secondary mass \( M_d \) dominates over the local disk mass \( M_d \)); see Equation (12). Evaluating \( F_{J,\text{Edd}} \) at the moment of decoupling would result in \( F_J \) being considerably smaller than the estimate given by Equation (46).

However, (1) at this point, the tidal torque is already completely negligible compared to the angular momentum loss due to the GW radiation and (2) such a condition would depend on the poorly understood value of \( b \) in Equation (25). Thus, we avoid this way of constructing Eddington limit-based estimates of \( F_J \).

One can also derive characteristic values of \( F_J \) based on arguments unrelated to the Eddington limit. For example, one may demand \( t_{\text{GW}} \) to be equal to some characteristic time \( t \) \( r_{\text{in}} \) given by Equation (45). This procedure would imply that after the GW emission becomes the dominant cause of the binary inspiral, the lifetime of the binary until the merger is equal to \( t \). Such an estimate gives:

\[
F_{J,t} = M_c c^2 \frac{q_S^{9/8}}{2^{27/8} 7^{1/8}} \frac{R_S}{c t} \frac{1}{7/8}
\]

\approx 4 \times 10^{46} \text{ erg} \ q_S^{9/8} M_c^{15/8} \left( \frac{t}{10^{10} \text{ yr}} \right)^{-7/8}.
\]  

(48)

and sets a lower limit on the value of \( F_J \) necessary for an equal-mass \( (q_S = 1) \), \( M_c = 10^7 M_\odot \) SMBH binary to merge within a Hubble time. For \( F_J = F_{J,\text{Edd}} \), such a binary merges within \( 4 \times 10^5 \text{ yr} \) after the GW emission starts to dominate its orbital evolution.

Finally, using Equations (2) and (25), one can also express \( F_J \) in a constant-\( F_J \) disk through the mass \( M_{\text{disk}} \) enclosed between its inner edge and some outer radius \( r_o \) via

\[
M_{\text{disk}} = \frac{2F_J}{3\alpha} \int_{r_{\text{in}}}^{r_o} \frac{dr}{r c_s^2} \sim \frac{F_J}{\alpha c_s^2 (r_o)}.
\]  

(49)

where it is assumed that gas pressure dominates (or \( b = 1 \)) and \( c_s = (kT/\mu)^{1/2} \). The approximate relation in this formula is valid for \( r_o \gg r_{\text{in}} \), provided that most of the disk mass is in its outer regions—a rather natural assumption as long as the midplane temperature falls with increasing radius, as this equation shows. Note that \( M_{\text{disk}} \) depends on \( r_o \) only via \( T(r_o) \).

Despite the estimates (48) and (49), which have a clear physical meaning, we still advocate the use of \( F_{J,\text{Edd}} \) as it represents an important upper limit on \( F_J \); for \( F_J \lesssim F_{J,\text{Edd}} \) the orbital evolution of the binary is essentially unaffected by the Eddington limit (even though at late stages of inspiral the inner part of the disk may become super-Eddington and be depleted by the radiation pressure-driven wind; see Section 6.2.6).

4.1.5. Global Properties of Steady State Disks

The results of Sections 4.1.1–4.1.3 allow us to understand the global characteristics of steady circumbinary disks described by the solution (15). In Figure 1, we illustrate the variation of disk properties as a function of \( r \) across regions with different opacities and pressure behaviors. We consider an \( M_c = 10^7 M_\odot \) binary orbited by a disk described by the relation (15) with \( F_{J,0} = 5 \times 10^{46} \text{ erg} \) and \( F_{J,1} = M_{\infty} = M_{\text{Edd}} \). Far from the binary (at \( r \rightarrow \infty \)), the disk transitions to a standard
constant $\dot{M} = \dot{M}_\infty$ accretion disk, while at small separations corresponding to $l \lesssim F_{J,0}/\dot{M}_\infty$, the disk turns into a constant-$F_J$ disk. Note that $\dot{M} = \dot{M}_\infty$ at all radii, meaning that mass has to be removed at $r_\infty$ at a rate $\dot{M}_\infty$.

We compare the properties of this disk to a standard constant $\dot{M} = \dot{M}_\infty$ disk in which $F_{J,0} = 0$ (dashed curves in Figure 1). We find that the midplane temperature $T$ is higher in a disk with non-zero $F_{J,0}$ in all regimes. As a result, this disk is more extended vertically and becomes geometrically thick $(h/r \sim 1)$ at $r = 3 \times 10^{-6}$ pc ($\approx 300 \, R_\odot$) in the radiation pressure-dominated regime. Unless it is truncated by the binary torque outside this radius, the disk should be losing mass at this point. Another consequence of higher $T$ is that all transitions between different regimes occur at larger $r$ (typically by a factor of several) in the disk with non-zero $F_{J,0}$. In particular, radiation pressure starts to dominate at $\approx 5 \times 10^{-5}$ pc as opposed to $\approx 10^{-3}$ pc in the $F_{J,0} = 0$ disk.

The behavior of $\Sigma$ depends on whether $b = 1$ or 0 in the radiation pressure-dominated regime. In the former case, $\Sigma$ keeps increasing toward small $r$ in both types of disks, which reflects the inefficiency of viscosity proportional to the (relatively small) gas pressure in the $b = 1$ case. We find $\Sigma$ to be higher in the $F_{J,0} \neq 0$ disk by an order of magnitude in the radiation pressure-dominated regime.

In the case of $v$ scaling with the radiation pressure ($b = 0$), the behavior of $\Sigma$ is completely different as it falls toward small $r$ because of the rapid increase of $T$ and $v$ in the radiation pressure-dominated regime. As a result, close to the binary $\Sigma$ is much smaller (by more than an order of magnitude) in the disk with non-zero $F_{J,0}$. The relatively small amount of mass residing in the vicinity of the SMBH binary in this case may have implications for the properties of the afterglow following the binary merger.

It is well known (Goodman 2003) that the outer parts of disks around SMBHs may be subject to gravitational instabilities (Safonov 1960; Toomre 1964; Goldreich & Lynden-Bell 1965) when the Toomre $Q$ parameter defined as $Q = \Omega c_s / (\pi G \Sigma)$ drops below unity. Using our results for $T(r)$ and $\Sigma(r)$ as a function of $F_J$ derived in Sections 4.1.1–4.1.3, one can easily determine the radius $r^\ast$ at which the disk becomes self-gravitating in different physical regimes. In the interest of brevity we do not perform this exercise here, but for constant $\dot{M}$ disks described by Equation (15) with $F_{J,0} = 0$ the expressions for $r^\ast$ can be found in Goodman (2003) and Haiman et al. (2009). Other effects that may invalidate our treatment at large separations—low optical depth at small $\Sigma$, neutrality of the disk at low $T$, etc.—have been previously discussed in Haiman et al. (2009).

5. DISK EVOLUTION

The steady state disk solutions discussed in Sections 3 and 4.1.5 require $\dot{M}$ to be independent of $r$. This situation is difficult to realize in real circumbinary disks, simply because the mass supply rate at large separations $\dot{M}_\infty$ is determined by processes that have nothing to do with the central binary. At the same time, $\dot{M}$ in the inner disk is limited by the binary torque, and is in general different from $\dot{M}_\infty$.

Because of this mismatch, gas has to accumulate somewhere. This naturally leads to the evolution of the disk properties, which is described by the time-dependent solutions of Equation (4). Deriving these solutions necessarily requires specifying the dependence of the diffusion coefficient $D_J$ (Equation (5)) on $l$ and $F_J$ and a set of BCs, for which we use Equations (13) and (14). Initial conditions in the form of a conventional constant $\dot{M}$ disk with

$$F_J(t = 0, l) = \dot{M}_\infty l$$

have been discussed in Section 2.2.

In some cases, Equation (4) can be solved analytically, which is useful for a qualitative understanding of more complicated situations. In particular, for $D_J = \text{const}$, one finds the following solution (with $l$ playing the role of the spatial coordinate):

$$F_J(t, l) = \chi \dot{M}_\infty l + (1 - \chi) \dot{M}_\infty \times \left\{ l_{\text{in}} + (l - l_{\text{in}}) \text{erf} \left( \frac{l - l_{\text{in}}}{\sqrt{2} D_J t} \right) + \left( \frac{4 D_J t}{\pi} \right)^{1/2} \exp \left[ -\left( \frac{(l - l_{\text{in}})^2}{4 D_J t} \right) \right] \right\}. \quad (51)$$

This solution is shown in Figure 2 for two different values of $M(l_{\text{in}})$ at different times. As time passes, the part of the disk affected by the central binary—the inner region with $F_J \to \text{const}$ in the $M(l_{\text{in}}) = 0$ ($\chi = 0$) case—extends to larger and larger values of $l$ (and $r$). The transition from the initial distribution (50) to $F_J$ strongly affected by the binary torque occurs at the radius of influence $r_{\text{inf}}$ where $l = l_{\text{inf}} \sim (D_J t)^{1/2}$.

4 Here, we follow the notation of IPP, who called the radius at which this transition occurs the "radius of influence" $r_{\text{inf}}$. 

---

**Figure 1.** Properties of a steady disk around an SMBH binary with mass $M_\infty = 10^5 M_\odot$ described by a solution (15) with $F_{J,0} = 5 \times 10^{56}$ erg and a mass accretion rate $\dot{M}_\infty$ (a radiative efficiency of 0.1 is assumed). Solid (black) curves describe the run of (a) aspect ratio $h/r$, (b) temperature $T$, (c) surface density $\Sigma$, and (d) angular momentum flux $\dot{F}_J$ with $r$ for a disk with $b = 0$ in a radiation pressure-dominated part of the disk. Dashed (blue) curves describe the same but for a standard constant $\dot{M}$ disk with the same mass accretion rate $\dot{M}_\infty$. Dotted extensions of these curves at small $r$ correspond to $b = 1$ in the radiation pressure-dominated part of the disk. Open dots on each curve correspond to the transition between the radiation and gas pressure-dominated regimes, while filled dots describe the transition between the electron scattering and free-free opacity.

(A color version of this figure is available in the online journal.)
Equation (4) with $D_J$ in the form (52) admits a self-similar solution provided that the problem at hand has no intrinsic scale. In circumbinary disks, the inner edge of the disk (comparable to the binary semi-major axis $r_{in}$) sets a natural scale. However, after the angular momentum injected by the binary has been transmitted by viscosity out to $r \gg r_b$, this scale becomes irrelevant and the evolution becomes self-similar.

We illustrate this point using solution (51), which can be written in the limit $r \gg l_{in}^2/D_J$ (so that $l_{in} \gg l_\infty$ or $r_{in} \gg r_b$) as

$$
F_J(t, l) = \frac{\dot{M}_\infty \sqrt{D_J t}}{2} \left[ \chi \xi + (1 - \chi) \xi \text{erf} \left( \frac{\xi}{2} \right) \right] \times (1 - \chi) \frac{2}{\sqrt{\pi}} \exp \left( -\frac{\xi^2}{4} \right),
$$

where $\xi = l/\sqrt{D_J t}$ is the dimensionless coordinate, which plays the role of an independent self-similar variable. It is clear from this result that at late times the solution of the evolutionary Equation (4) for $D_J = \text{const}$ is indeed independent of the $l_{in}$ at which the inner BC is imposed; as a result, one can set $l_{in}$ to 0 and eliminate intrinsic scale from the problem.

In the general case of arbitrary $d$ and $p$ in Equation (52), we first define new variables

$$
f_J = F_J/M_\infty, \quad \tau = D_{J,0} M_\infty^d t,
$$

transforming Equation (4) into

$$
\frac{\partial}{\partial \tau} f_J^{1-d} = l_p \frac{\partial^2 f_J}{\partial l^2},
$$

with the BC

$$
\frac{\partial f_J}{\partial l} \bigg|_{l=0} = \frac{\dot{M}(l=0)}{M_\infty} = \chi
$$

(instead of Equation (14)), where $\chi \lesssim 1$ and may be equal to 0. One can easily see that Equation (55) admits a self-similar solution in the form

$$
f_J = \tau^n f(\xi), \quad \xi = \frac{l}{l_p}, \quad n = -\frac{1}{d + p - 2},
$$

where the function $f$ satisfies the ordinary differential equation

$$
f'' + d f^d = n(1 - d)(f - \xi f')
$$

with the BCs

$$
f'(\xi \to 0) = \chi, \quad f'(\xi \to \infty) = 1.
$$

The second BC follows directly from the initial condition (50). We have verified that the self-similar scalings of $\Sigma$ derived in IPP agree with Equations (57) and (58) for the $\chi = 0$ case.

In Figure 3, we show the solutions of this equation for different values of $\chi$ (generalizing the results of IPP to the case with overflow). We have chosen two sets of $d$ and $p$ corresponding to astrophysically relevant situations: a gas pressure-dominated regime with $\kappa = \kappa_0$ ($d = 2/5, p = -6/5, n = 5/14$) and $\kappa = \kappa_{in}$ ($d = 3/10, p = -4/5, n = 2/5$); see

---

5 Previously, Filipov (1984), Lyubarskij & Shakura (1987), and Lipunova & Shakura (2000) derived such solutions for $F_J$ assuming BCs different from what we use here.
analogy with the analytical solution (51) that allowing for some binary. Note, however, that a significant reduction of the binary non-zero mass flow across the orbit of the secondary (i.e., assuming a time $t_{\nu}$ unaffected by the binary torque to the inner solution always occurs at $\xi \approx 1$–2. This transition occurs where the viscous time $t_{\nu}$ is about the time $t$ that has passed since the central binary started tidally interacting with the disk at its inner edge: $t_{\nu}(\xi \sim 1) \sim t$.

We take the radius of influence $r_{\text{infl}}$ to correspond to $\xi_{\text{infl}} = 1$. Then, according to Equations (54) and (57) the value of the specific angular momentum $l_{\text{infl}}$ at $r_{\text{infl}}$ is given by

$$l_{\text{infl}}(t) = (D_{J,0}M_{\infty}^d)^{1/n}t^n, r_{\text{infl}}(t) = \frac{l_{\text{infl}}^2(t)}{GM_c},$$

while the viscous angular momentum flux in the disk is

$$F_J(l, t) = M_{\infty}l_{\text{infl}}(t) \times f (l/l_{\text{infl}}(t), \chi),$$

where the value of $f(\xi, \chi)$ can be found from Figure 3 for a given $\chi$.

### 5.2. Binary Evolution

We now determine the response of the binary to the self-similar disk evolution. The torque acting on the binary in the limit $l_{\text{infl}} \ll l_{\text{infl}}(t)$ (or $r_{\text{infl}} \ll r_{\text{infl}}(t)$) is $F_J(0, t) = M_{\infty}l_{\text{infl}}(t)f(0, \chi)$, where the dependence of $f(0, \chi)$ on the strength of mass inflow through the secondary orbit $\chi$ is shown in Figure 4 for different types of disks. One can see in complete analogy with the analytical solution (51) that allowing for some mass flow across the orbit of the secondary (i.e., assuming a non-zero $\chi$) leads to a reduction of the torque acting on the binary. Note, however, that a significant reduction of the binary torque (e.g., by a factor of two) requires quite substantial overflow, $\chi \gtrsim 0.5$. In the weak overflow case, when $\chi \lesssim 0.1$, the disk structure is very close to the no-overflow case and $F_J$ is roughly constant in the inner disk.

It might seem surprising that the torque experienced by the binary does not depend on the mass of the secondary $F_J(0, t)$ is set only by $M_{\infty}$ and $l_{\text{infl}}(t)$, since ultimately it is the potential of the secondary that gives rise to the tidal coupling with the disk. We comment on this point in Section 6.

Combining Equations (8) and (11) and integrating them using expressions (60) and (61), we obtain the orbital evolution of the secondary in terms of $l_b = \sqrt{GM_c}$:

$$l_b(t) \approx l_b, 0 \left[ \frac{1 - 1 + q}{1 + n} B f(0, \chi) \left( \frac{l_{\text{infl}}(t)}{l_{b,0}} \right)^{(1+n)/n} - 1 \right],$$

starting at an initial value of the specific angular momentum $l_{b,0}$. In this expression, we define, following SC95, a dimensionless parameter

$$B = \frac{M_{d, 0}}{M_*}, \quad M_{d, 0} = \frac{\Sigma(r_{b,0})r_{b,0}^2}{(1 + q)}$$

($S$ in the notation of IPP), which is a ratio of the “local disk mass” $M_{d, 0}$ at $r_{b,0}$ in a reference constant $M = M_{\infty}$ disk (note that $M_{d, 0}$ is not the total mass of the disk), computed using definitions (2), (5), and (52) and assuming $F_J = M_{\infty}l_{b,0}$ to the secondary mass $M_s = qM_c/(1 + q)$. In combination with Equation (60), this result agrees with the $r_{b}(t)$ dependence found in IPP, if we additionally assume the no-overflow case ($\chi \ll 1$) so that $f(0, \chi) \approx 1$. 

---

**Figure 3.** Behavior of the self-similar function $f$ vs. the variable $\xi$ defined in Equation (57). Different panels correspond to different values of the power-law parameters $d$ and $p$: (a) gas pressure-dominated case with $\kappa = \kappa_d$, (b) gas pressure-dominated case with $\kappa = \kappa_{\infty}$. Various curves represent the run of $f(\xi)$ for different boundary conditions imposed at the inner edge, allowing for the possibility of mass inflow past the secondary orbit: $M(0) = \chi M_{\infty}$.

**Figure 4.** Value of $f(\xi = 0, \chi)$ as a function of $\chi$—the fraction of $M_{\infty}$ passing through the orbit of the secondary. Different curves correspond to different values of the power-law parameters $d$ and $p$ indicated in the figure: dotted lines correspond to the gas pressure-dominated case with $\kappa = \kappa_d$, and solid lines correspond to the gas pressure-dominated case with $\kappa = \kappa_{\infty}$. Knowledge of $f(\xi = 0)$ permits a computation of the torque acting on the central binary for an arbitrary value of the “accretion fraction” $\chi$. 

---

Appendix C or Lyubarskij & Shakura (1987). We will see later in Section 6 that the radius of influence often extends into the parts of the disk where one of these regimes is valid.

As expected, the transition from the outer solution $f(\xi) = \xi$ unaffected by the binary torque to the inner solution always occurs at $\xi \approx 1$–2. This transition occurs where the viscous time $t_{\nu}$ is about the time $t$ that has passed since the central binary started tidally interacting with the disk at its inner edge: $t_{\nu}(\xi \sim 1) \sim t$.
In the secondary-dominated regime described by Equation (10), \( B \ll 1 \). This result means that the binary merges (\( l_b \to 0 \)) when the radius of influence reaches a critical value

\[
\dot{m}_\text{infl,m} \sim [f(0, \chi)B]^{-n/(n+1)} l_b,0 \gg l_b,0, \tag{64}
\]

at a time (see IPP)

\[
\dot{m}_\text{infl} \sim [f(0, \chi)B]^{-1/(n+1)} t_{\nu,0} \tag{65}
\]

(here \( t_{\nu,0} = l_b,0/(D_J m_{\text{infl},0}) \)) is the viscous timescale at \( r_b,0 \) in a reference constant \( M = M_{\infty, \text{disk}} \), i.e., only after the binary torque has affected the disk out to radii much larger than the initial binary semi-major axis \( r_b,0 \). During this time, the mass

\[
\Delta M = \dot{M}_{\infty, \text{infl}} \sim \left[ \frac{B^n}{f(0, \chi)} \right]^{1/(n+1)} M_t \ll M_t \tag{66}
\]

is going to be processed through the circumbinary disk. The total mass of the disk can easily be higher than \( \Delta M \) if it extends far beyond \( r_{\text{infl,m}} = l_b,0/(D_J m_{\text{infl},0}) \), but having a disk extending out to \( r_{\text{infl,m}} \) containing only \( \Delta M \) of gas is enough for the merger to occur. Thus, a binary starting in the secondary-dominated regime \( B \ll 1 \) with weak or no overflow \( (f(0, \chi) \sim 1) \) can merge even if its disk is less massive than the secondary.

This statement may seem paradoxical at first but is in fact a consequence of angular momentum conservation in the binary–disk system: because the specific angular momentum goes as \( r^{1/2} \), the binary can lose all of its angular momentum and merge by sharing its angular momentum with even a relatively small mass of gas far away. The necessary minimum gas mass is smaller for higher \( r_{\text{infl,m}} \), since \( \Delta M/M_t \sim l_b,0/l_{\text{infl,m}} \sim (r_{b,0}/r_{\text{infl,m}})^{1/2} \) for \( f(0, \chi) \sim 1 \), as trivially follows from Equations (64) and (66).

5.3. General Description of the Disk Evolution

In Section 5.1, we outlined the main features of the self-similar evolution of a circumbinary disk that arises when three essential conditions are met: (1) the behavior of the diffusion coefficient \( D_J \) is given by a simple power-law form (52), (2) the outer parts of the disk are well approximated by a standard constant \( M \) disk with \( F_J = M_{\infty, J} \), and (3) the radius of influence \( r_{\text{infl}} \) far exceeds the semi-major axis of the central binary \( r_b \) and the radius of the inner disk edge \( r_{in} \). We now describe how the picture of the disk evolution changes when these assumptions are relaxed by concentrating on the situation when there is no mass inflow across the orbit of the secondary, i.e., \( \chi = 0 \) or \( M(r_{\text{in}}) = 0 \). A more complicated setup allowing for some mass inflow across the orbit of the secondary can be understood by generalizing the picture that emerges from the \( M(r_{\text{in}}) = 0 \) case.

Different parts of circumbinary disks can feature different physical environments, as illustrated in Section 4. For example, the inner disk can be in the radiation pressure–dominated regime, while further out gas pressure dominates with the opacity initially given by \( K_{\text{ps}} \) and then by \( K_{\text{fr}} \). In all these regimes, one can still use the power-law description of \( D_J \) locally as described in Appendix C, with smooth transitions between the different scalings at the boundaries of different regimes but the global self-similarity will be destroyed. Nevertheless, we can still understand the disk behavior in this more complicated situation as follows.

Interior to \( r_{\text{infl}} \), the viscous time in the disk decreases with decreasing \( r \), meaning that for \( r \lesssim r_{\text{infl}} \) the disk tends to approach a steady state solution. Then, Equation (4) implies that \( F_J \) in this part of the disk is given by a simple solution (15) independent of the complicated behavior of \( D_J \) caused by the transitions between different physical regimes. The disk properties such as \( \Sigma(r) \), \( T(r) \), etc. can be computed as functions of this radially constant \( F_J \) (for \( M(r_{\text{in}}) = 0 \)) using the formulae derived in Section 4; these properties will show different dependences on \( r \), \( F_J \), and other system parameters in different regimes.

At every moment of time, the value of the radially constant \( F_J \) (for \( r \lesssim r_{\text{infl}} \)) angular momentum flux \( F_J \) is obviously set by the disk properties in a particular physical regime corresponding to \( r \sim r_{\text{infl}} \). This regime can change in time since \( F_J \) steadily increases and both the transition radii of different regimes and \( r_{\text{infl}} \) vary. Nevertheless, it is clear that our results for the self-similar disk evolution obtained previously should allow one to easily understand this more complicated situation; see Section 6.

A second complication arises if the circumbinary disk does not start out as a standard constant \( M \) disk with \( F_J \) given by Equation (50) but is instead characterized by some more complicated initial distribution of \( F_J(l, t = 0) \). Again, our understanding of the self-similar disks allows us to qualitatively characterize disk evolution in this case. The region influenced by the binary torque still expands in time with the dependence \( l_{\text{infl}}(t) \) given by an implicit relation

\[
l^2_{\text{infl}} \sim t \times D_J (F_J(l_{\text{infl}}, t = 0), l_{\text{infl}}). \tag{67}
\]

We refer the reader to Equation (4), in which we explicitly indicated the dependence of \( D_J(F_J, l) \) on \( F_J \) and \( l \). Interior to \( r_{\text{infl}} \), the angular momentum flux is roughly constant with radius and is equal to \( F_J(l_{\text{infl}}, t = 0) \).

Outside \( r_{\text{infl}}(t) \), the disk still maintains the distribution of \( F_J \) close to the initial distribution \( F_J(l, t = 0) \) since the viscous time there is long compared to the system lifetime (which is also equal to \( t_{\text{visc}}(r_{\text{infl}}) \)). If the initial distribution of \( F_J \) exhibits a maximum at some radius \( r_{\text{max}} \), then past the moment when \( r_{\text{infl}} \sim r_{\text{max}} \) the circumbinary disk will turn into a decretion disk (Pringle 1991) and the mass accumulated in the central part of the disk will flow out, driven by the continuing injection of angular momentum by the binary (assuming that the binary does not merge by that time).

Finally, initially \( r_{\text{infl}} \) may be close to the inner edge of the disk. This assumption is true if the evolutionary lifetime of the system has not yet exceeded the viscous time at the inner edge of the disk. However, at later times, the condition \( r_{\text{infl}} \gtrsim r_{\text{in}} \) is guaranteed to be fulfilled since \( r_{\text{infl}} \) steadily grows while both the binary semi-major axis and \( r_{\text{in}} \) can only decrease. As a result, at late times the system inevitably converges to the self-similar track (see, e.g., the evolution shown in Figure 2) or its generalizations described above for the more complicated situations.

6. IMPLICATIONS FOR SMBH BINARY EVOLUTION

We now apply the results obtained in previous sections to the coupled evolution of SMBH binaries and the disks around them. The orbital evolution of the binary and the evolution of the disk properties must be considered simultaneously because of their mutual influence on each other. In exploring this evolution, we pay special attention to the non-local nature of the binary–disk coupling, namely, the fact that the angular momentum flux \( F_J \) carried through the disk near the binary (which determines its orbital evolution) is set by the disk properties at the radius \( r_{\text{infl}} \), which can far exceed the semi-major axis of the binary.
This fact, overlooked in previous studies of SMBH binary evolution, is very important, as we show below.

In our subsequent calculations, we are not addressing the “final parsec” problem (Lodato et al. 2009) directly as we typically follow SMBH binaries starting at rather small separations, $10^{-2} - 10^{-4}$ pc. Such binaries may be created by previous (possibly multiple) episodes of gas infall into the center of the galaxy in which the binary resides, each of which tightens the binary orbit. At the same time, some of our findings (e.g., the significant reduction of the binary inspiral timescale when the disk evolution is self-consistently included) are likely to be relevant for attempting to resolve the “final parsec” problem.

We take the viscosity in the radiation pressure-dominated regime to scale with the total rather than the gas pressure, i.e., $\dot{b} = 0$. We also assume that the tidal torque of the binary presents a sufficiently strong barrier to inflowing gas to completely suppress gas overflow across the orbit of the secondary, i.e., $\chi = M(r_{\text{in}})/M_\infty = 0$. In principle, one can easily extend our results to the case of $\chi \neq 0$. Finally, even though the binary itself is not accreting when $M(r_{\text{in}}) = 0$, the inner parts of the disk are still gaining mass, which changes the potential in which gas orbits further out. In this work, we are mainly concerned with the disk-related effects on the binary evolution and for that reason we neglect the increase of the binary+disk mass throughout the calculation.

### 6.1. Binary Inspiral: Basic Features

In Figure 5, we show a representative case of the binary+disk evolution. The binary orbit is evolved according to

$$\frac{dr_b}{dt} = -\frac{r_b}{t_G} - \frac{r_b}{t_J},$$

where $t_G$ and $t_J$ are given by Equations (43) and (44). The disk properties are described via the following simple piecewise dependence of $F_J$ on $r$:

$$F_J(r, t) = \begin{cases} M_\infty (GM_b r)^{1/2}, & r > r_{\text{inf}}(t), \\ M_\infty (GM_b r_{\text{inf}}(t))^{1/2}, & r \leq r_{\text{inf}}(t), \end{cases}$$

where $r_{\text{inf}}(t)$ is self-consistently calculated in each physical regime using Equation (60). This prescription means that outside the radius of influence, $F_J(r, t)$ behaves as appropriate for a constant $M$ disk, while inside $r_{\text{inf}}$ the disk attains a constant-$F_J$ structure, in qualitative agreement with the picture outlined in Section 5. Unlike the calculations of the self-similar disk behavior in Section 5.1, we do not consider here the details of the smooth transitions between the parts of the disk inside and outside of $r_{\text{inf}}$.

We start with an equal-mass ($q = 1$) binary with a total mass $M_c = 10^7 M_\odot$ with a semi-major axis of $10^{-3}$ pc, surrounded by a standard constant $M = M_{\text{Edd}}$ disk extending from very large distances down to the binary semi-major axis (for simplicity, we disregard the difference between the binary semi-major axis $r_b$ and the inner radius of the disk, which is a factor of two uncertainly at most; see MacFadyen & Milosavljevic 2008).

The radial dependence of the disk properties at time $t = 0$ can be found in Haiman et al. (2009) or by setting $F_J = M_\infty (GM_b r)^{1/2}$ in the formulae derived in Section 4.1. As Figure 5 shows, the binary starts in the radiation pressure-dominated part of the disk but the transition to the gas pressure-dominated regime occurs not too far outside of $r_b$, at $r_{\text{gas}}/r_{\text{rad}} \approx 2.5 \times 10^{-3}$ pc. The opacity switches from being dominated by electron scattering to being dominated by free–free opacity at $r_{\text{es}}/r_{\text{ff}} \approx 0.03$ pc. These regimes are clearly labeled in Figure 5.

We also show the run of the radius of influence $r_{\text{inf}}(t)$ with time by the dashed curve. At $t = r_{\text{inf}}(t)$, the local viscous time equals the time since the start of the evolution $t$ but this number is meaningful only if at time $t$ the disk extends to $r < r_{\text{inf}}(t)$, which is not always the case. Nevertheless, this dependence is still a useful concept as it allows us to see important transitions in the disk properties if we were to adopt an initial SMBH semi-major axis $r_0(0)$ different from the value shown in Figure 5. The actual dependence $r_{\text{inf}}(t)$ is calculated using definition (60) and the expressions for $D_{\text{rad}}, d$, and $n$ that can be found in Appendix C. Since these expressions are different in various physical regimes, the behavior of $r_{\text{inf}}(t)$ exhibits distinct transitions as it crosses the boundaries of different regimes, clearly visible in Figure 5.

Figure 5 also shows the behavior of the torque $F_J(r_{\text{in}}, t)$ at the inner edge of the disk (thick solid curve calibrated on the right axis), which is absorbed by the binary and causes its orbital evolution. The $F_J(r_{\text{in}}, t)$ curve initially closely follows $r_b(t)$ because the torque exerted on the binary by a constant $M$ disk is well approximated by $F_J(r_{\text{in}}, t) = M_\infty (GM_b r_b)^{1/2}$ as long as the inner edge of the disk tracks the binary orbit. This torque is

---

6 For clarity, we have slightly shifted the curve $F_J(r_{\text{in}}, t)$ upward in Figure 5 to avoid overlap with other curves, e.g., $r_b(t)$. 
small enough for $r_b$ not to change appreciably for a rather long time. Approximately at $t = 2 \times 10^4$ yr, when $r_{\text{visc}}(r_b(0)) \sim t$, the radius of influence of the binary torque $r_{\text{infl}}$ grows beyond the initial binary semi-major axis $r_b(0)$. At this point, $F_J(r_{\text{infl}}, t)$ starts tracking the run of $r_{\text{infl}}(t)$ in Figure 5. The angular momentum flux at the inner edge of the disk $F_J(r_{\text{infl}}, t)$ increases as $F_J(r_{\text{infl}}, t) \propto l_{\text{infl}}(t) \propto t^{1/2}$, thus accelerating the orbital evolution of the binary. Initially, $r_{\text{infl}}$ stays in the radiation pressure-dominated regime and varies as $r_{\text{infl}} \propto t^{2/7}$; see Equations (60) and (C1). At $t \approx 2 \times 10^3$ yr, the radius of influence reaches out into the gas pressure-dominated regime with $\kappa = \kappa_{\text{cs}}$. There, $r_{\text{infl}}$ grows as $r_{\text{infl}} \propto t^{2/7}$ (see the break in the slope of the $r_{\text{infl}}$ curve), and the increase of $F_J(r_{\text{infl}}, t)$ acting on the binary accelerates, along with the torque.

As the value of $F_J$ in the inner disk ($r \lesssim r_{\text{infl}}$) grows, the disk properties keep changing as well. In particular, the boundaries of the different regimes, i.e., $r_{\text{rad/gas}}$ and $r_{\text{visc/ff}}$, expand as time goes by. Thus, the state of the disk at the location of the binary may change not only because of the variation of the binary orbit, but also due to the disk evolution; see Figures 8 and 6.

Note that even though the binary is always in contact with the radiation pressure-dominated region of the disk, past $t \approx 2 \times 10^5$ yr the torque on the binary is determined by the gas pressure-dominated part of the disk with $\kappa = \kappa_{\text{cs}}$, since this location is where $r_{\text{infl}}$ is. This result demonstrates the non-locality of the disk–binary coupling: the state of the disk near the binary is irrelevant for its orbital evolution at late times because the torque on the binary is set at $r \sim r_{\text{infl}}$.

To better illustrate the role of mass pileup and disk evolution for the orbital evolution of the binary, we compare our results with calculations in which disk properties remain well represented by the properties of a constant $M = M_{\infty}$ disk at all times. Then, the torque on the binary is always given by $F_J(t) = M_{\infty}[G M r_b(t)/r_{\text{infl}})^{1/2}$; see the dotted curve in Figure 5. Clearly, at late times, this torque is much smaller than the real $F_J(t)$, meaning that calculations assuming a constant $M$ disk overestimate the inspiral time of the binary.

6.2. Binary Inspiral: Parameter Exploration

We now provide a more detailed and systematic view of the SMBH binary evolution under different conditions. We explore two representative values of $M_c$: $10^5 M_\odot$ (Schwarzschild radius $R_S = 10^{-3}$ pc) and $10^7 M_\odot$ ($R_S = 10^{-6}$ pc), but our results, as shown in Figures 6 and 7, can be trivially extended to other values of $M_c$. The accretion rate in the disk is taken to be either $\dot{M}_{\infty} = 10^{-2} \dot{M}_{\text{Edd}}$ or $\dot{M}_{\text{Edd}}$ and the binary mass ratio is varied between $q = 10^{-2}$ and 1. We also consider two different values of the starting semi-major axis of the binary $r_b(0): 10^{-4}$ pc and $10^{-3}$ pc for $M_c = 10^5 M_\odot$ and $10^{-3}$ pc and $10^{-2}$ pc for $M_c = 10^7 M_\odot$. These values are close to the “bottleneck” semi-major axes at which the stellar dynamical orbital evolution of SMBH pairs decelerates dramatically; see Yu (2002).

One might worry that the outer parts of the disk are prone to gravitational instabilities. We determined that in the initial
constant \( M \) disk, the Toomre Q equals unity at \( r^*_{\text{Edd}} = 9 \times 10^{-3} \) pc, 0.3 pc, 0.3 pc, and 0.5 pc for the systems shown in panels (a)-(d), respectively, in Figure 6 (for \( \alpha = 0.1, \varepsilon = 0.1, \) and \( \mu = 0.5m_\odot \)). Thus, at the start of our calculations, one needs to worry about the importance of the disk self-gravity only for the system with \( M = 10^7 M_\odot \) and \( M_{\text{in}} = M_{\text{Edd}} \) and the binary starting at \( r_0(0) = 10^{-2} \) pc, where the disk can be marginally gravitationally unstable. One has to keep in mind though that the torque on the binary is determined at \( r_{\text{infl}} \), which later expands beyond \( r_b(0) \); see Section 6.1. If \( r_{\text{infl}} \) exceeds \( r^*_{\text{Edd}} \) at some point, the calculations of the disk evolution would need to be refined. In our present study, we neglect this complication; it may only be an issue for the high-mass systems shown in Figure 6(a).

Figure 6 shows the evolutionary tracks of the binary orbit mapped onto the disk state in a format analogous to that used in Figure 5. In addition, in Figures 7(a) and (b) we show the dependence of the binary orbital evolution timescale \( t_{\text{ev}} \equiv -\frac{d}{dt} \log \left( \frac{r}{r_b} \right) \) versus the binary orbital period \( P_{\text{orb}} \) or semi-major axis \( r_b \) (upper axis) for the evolutionary tracks displayed in Figure 6. This plot allows us to easily see the transition from the disk-driven evolution at longer periods, where \( t_{\text{ev}} = t_J \), to the GW-dominated phase, which is clearly described by straight line tracks at small values of \( P_{\text{orb}} \). Figures 7(a) and (b) can be directly compared to the analogous \( t_{\text{ev}}(P_{\text{orb}}) \) plots in Haiman et al. (2009).

The results presented in Figures 6 and 7 can be summarized in the following set of conclusions.

### 6.2.1. Circumbinary Disks can be Efficient in Driving the Orbital Evolution of the Binary

Figures 6 and 7 clearly show that coupling to circumbinary disks can appreciably accelerate the orbital evolution of SMBH binaries, in agreement with the results of existing studies (e.g., IPP; Haiman et al. 2009). For most of the systems, the inspiral timescale due to GW emission alone is very long: for example, Equation (43) predicts \( t_{\text{GW}} = 8.3 \times 10^{10} \) yr for an equal-mass \( 10^7 M_\odot \) SMBH binary starting at \( r_0(0) = 10^{-2} \) pc, and systems with small mass ratios evolve even slower. The only exception in our sample is the \( q = 1, 10^7 M_\odot \) binary starting at \( 10^{-2} \) pc and surrounded by a disk with \( M_{\text{in}} = 10^{-2} M_{\text{Edd}} \), for which the GW emission dominates from the start; see Figure 7(a). This system merges faster than its \( q = 10^{-2} \) counterpart, even though the latter is more affected by disk torques.

On the other hand, the same \( q = 1, 10^7 M_\odot \) binary starting at \( r_0(0) = 10^{-2} \) pc and surrounded by the disk accreting at \( M_{\text{in}} = M_{\text{Edd}} \) merges within 7 Myr—more than four orders of magnitude faster than without the disk! Lower mass disks are of course less efficient at driving the orbital evolution of SMBH binaries—the same binary surrounded by a disk with \( M_{\text{in}} = 10^{-2} M_{\text{Edd}} \) merges within \( 3 \times 10^8 \) yr, but this time is still much shorter than the
corresponding $t_{GW}$. Lower $q$ binaries are affected by the disk even more, provided that they present an efficient barrier to the mass inflow at the orbit of the secondary: a $10^7 M_\odot$ binary with $q = 10^{-2}$ surrounded by a disk with $M_\infty = 10^{-2} M_{\text{Edd}}$ merges within $3 \times 10^7$ yr, about an order of magnitude faster than the $q = 1$ binary with the same parameters.

In Figure 6, black square dots mark the location where the disk-driven evolution switches to the GW-driven orbital decay. In systems with massive, high-$M_\infty$ disks, this transition typically occurs when the binary is in contact with the radiation pressure-dominated part of the disk.

Another way to state the importance of the disk-driven evolution is to note that in most cases, the transition to the GW-dominated regime occurs at $r_b \ll r_b(0)$, i.e., after the binary semi-major axis has been significantly reduced by the disk torques. For example, the evolution of an $M_c = 10^3 M_\odot$ binary with $q = 10^{-2}$ starting at $10^{-3}$ pc is dominated by torques produced by a $M_\infty = M_{\text{Edd}}$ disk down to $r_b \approx 2 \times 10^{-6}$ pc $\approx 200 R_\odot$ (Figure 6(c)), which is almost three orders of magnitude smaller than $r_b(0)$. This result highlights the importance of disk torques for shrinking SMBH binary orbits and their possible relevance for solving the last parsec problem.

6.2.2. Non-local Character of the Disk–Binary Coupling

As described in Section 6.1, the torque exerted on the binary is set by the disk properties at $r_{\text{infl}}$. Initially, a binary can affect only its immediate surroundings as it takes a certain time for the disk to absorb enough angular momentum injected by the binary to affect the $\Sigma$ distribution further out in the disk. For that reason, initially, $F_J(r_{in})$ is essentially the same as in the case of a constant $M = M_\infty$ disk, and is set by the disk locally, at $r \sim r_b$. This torque is usually rather small, implying a slow initial orbital evolution and long values of $t_{GW}$.

However, after the system has evolved for a time comparable to the viscous timescale at the inner disk edge $t_{\text{visc}}(r_{in})$, $r_{\text{infl}}$ starts exceeding $r_b$. Past that point, the torque on the binary $F_J(r_{in})$ is being set globally, at distances far exceeding $r_b$, which can be clearly seen in several evolutionary tracks shown in Figure 6. The increase of $F_J(r_{in})$ often occurs initially at almost constant $r_b$, which is reflected in the almost vertical initial evolutionary tracks in the $I_{\nu} - P_{\text{orb}}$ plane shown in Figures 7(a) and (b), clearly visible for $M_c = 10^3 M_\odot$ and low $M_\infty/M_{\text{Edd}} = 10^{-2}$.

It might seem strange that the torque acting on the binary is independent of the mass of the secondary $M_c$ and $q$ (see also IPP). Also, the dependence of $F_J(r_{in})$ on the total mass of the binary $M_t$ arises only because $M_t$ determines the angular frequency in the disk. One has to keep in mind that the full torque exerted by the binary on the disk (equal and opposite to $F_J(r_{in})$) is generally found to scale as (Goldreich & Tremaine 1980; Papaloizou & Lin 1984; Petrovich & Rafikov 2012):

$$F_J(r_{in}) \propto \frac{M_t^2 \Sigma_0}{\Delta^3},$$  \hspace{1cm} (70)

irrespective of the precise form of the torque density distribution. Here, $\Sigma_0$ is the disk surface density just outside the region where the binary torques are important and $\Delta = |r_{in} - r_b|$ is the width of the gap—the separation between the secondary orbit and the inner edge of the disk.

Equation (70) shows that a particular value of $F_J(r_{in})$ can be obtained not only by changing $M_t$ but also by varying the width of the gap $\Delta$. This explains how the disk–binary tidal interaction self-regulates itself to provide the necessary torque on the disk. As $F_J(r_{in})$ varies in time for a fixed $M_t$, the width of the gap changes as well. The same is true if one varies the mass of the secondary while keeping $F_J(r_{in})$ constant—the width of the gap simply scales as $\Delta \propto q^{2/3}$. Of course, $q$ cannot be arbitrarily small; see Section 6.2.7 for details. But as long as the gap opening conditions are satisfied for a given $q$, $\Delta$ should always be such as to provide just the right amount of torque on the disk.

The global nature of the disk–binary coupling is more pronounced for lower mass binaries and for lower $M_\infty$, because a higher $M_t$ implies an earlier transition to the GW-dominated orbital decay (see Figures 7(a) and (b)), shortening the binary lifetime and preventing $r_{\text{infl}}$ from extending as far as in the lower $M_c$ case. Higher $M_\infty/M_{\text{Edd}}$ values play a similar role, shortening the binary lifetime and reducing the radius of influence at the end of inspiral compared to the lower $M_\infty$ case.

For almost all the tracks shown in Figure 6, the inner edge of the disk (assumed to be equal to $r_b$) and $r_{\text{infl}}$ reside at some point in parts of the disk corresponding to different physical regimes. Under these circumstances, calculations assuming the behavior of $r_{\text{infl}}(t)$ corresponding to the physical state of the inner edge of disk (Haiman et al. 2009; Kocsis et al. 2011) should not produce accurate results. Figure 6 demonstrates that using this simple-minded procedure for a binary in the radiation pressure-dominated regime can easily underestimate the torque $F_J(r_{in})$ driving its orbital evolution, thus overestimating the binary lifetime (see, e.g., Figures 6(a) and (c)).

Another consequence of the non-locality of the disk–binary coupling is the clear hysteresis in the evolution of the system—the dependence of the current rate of the orbital decay of the binary on the previous history of the disk evolution. This property is most readily seen in Figure 7, in which the evolutionary tracks computed for the same $M_c$, $q$, and $M_\infty$ but starting at different $r_b(0)$ result in different (easily by a factor of several) orbital decay timescales $t_{GW}$ at a given orbital period in the disk-driven regime. This result is in contrast with local calculations presented in Haiman et al. (2009), in which the orbital evolution depends on the physical conditions at the current value of $r_b$.

6.2.3. Evolution of the Disk Accelerates the Orbital Evolution of the Binary

We find the difference of the circumbinary disk structure from that of a standard constant $M$ disk to be very important for binary evolution. In Figure 8, we display $r_b(0)$ calculated using constant $M$ disk properties (dotted curves) and fully accounting for the binary-driven disk evolution (solid curves) for two equal-mass ($q = 1$) SMBH binaries starting at different radii $r_b(0)$. One clearly sees that in most cases the evolution of $r_b$ computed using the standard constant $M$ disk properties considerably overestimates the binary inspiral time (by almost an order of magnitude in some cases). Whenever disk torques dominate (not the case for the $M_c = 10^3 M_\odot$, $M_\infty = 10^{-2} M_{\text{Edd}}$, and $r_b(0) = 10^{-3}$ pc system for which GW emission dominates at all times), binaries shrink faster when the self-consistent disk evolution is properly taken into account. For this reason, the standard constant $M$ disk model should not be used for exploring SMBH binary evolution.

Because the radius $r_{\text{infl}}$ setting the value of the inner torque $F_J(r_{in})$ steadily grows (see Equation (69)), we necessarily find that disk evolution speeds up binary decay. Interestingly, previously Haiman et al. (2009) claimed that disk evolution computed according to IPP, who accounted for the mass pileup near the orbit of the secondary, slows down the inspiral (see their Figures 6 and 7). We believe that this conclusion resulted
from their choice of the reference state—they compared IPP-like tracks with the evolutionary paths calculated using the quasi-steady SC95 solution. As we demonstrate in Section 2.3 and Appendix A, the assumptions used in deriving the latter solution are never realized in practice so the evolutionary tracks that use these assumptions do not represent the correct behavior of the system. As pointed out in Ju et al. (2013), the use of the SC95 solution in population synthesis models predicting the distribution of SMBH binary semi-major axes may result in considerable underestimate of the abundance of binaries with periods ~0.1–1 yr for $M_\star \sim 10^5$–10$^7$ $M_\odot$.

6.2.4. Validity of the Secondary-dominated Regime

Our present study always assumes evolution to be secondary-dominated (SC95; IPP; Haiman et al. 2009), in the spirit of Equation (10). In Figures 7(c) and (d), we display the ratio $\Sigma(\mathbf{r}_0) r_0^2 / M_\star$ and one can see that almost all evolutionary tracks shown there indeed satisfy $M_d / M_\star \ll 1$ at all times. This result is partly due to the relatively small values of $r_0(0)$ that we adopt. Indeed, concentrating on the gas pressure-dominated regime with free–free opacity, we can compute the initial value of $M_d / M_\star = B$ using Equation (38) with $F_J = M_\infty l(r_0(0))$:

$$B \approx 3 \frac{1 + q}{q} \left[ \frac{r_0(0)}{\text{pc}} \right]^{5/4} \left[ \frac{M_\infty}{M_\odot} \right]^{3/10} \left[ \frac{M_c}{10^7 M_\odot} \right]^{-1/5} \left( l \right)^{3/10}.$$

(71)

According to this formula, low $q$ and high $M_\infty$ can make the evolution disk-dominated even if $r_0(0)$ is small. For our parameters, only the evolutionary track for a $q = 10^{-2}$, $M_\star = 10^7 M_\odot$ SMBH binary starting at 10$^{-2}$ pc and surrounded by a disk with $M_\infty = M_{\text{Edd}}$ is in danger of having $B \sim 1$, which is indeed the case according to Figure 7(c) ($M_d / M_\star \sim 0.2$ initially). As a result, this track exhibits noticeable evolution even prior to crossing the $r_{\text{inf}}(t)$ curve, an effect that shows up in Figures 7(a) and (b) as flat initial portions of the evolutionary tracks for $q = 0.01$ and large $r_0(0)$. Treating such marginal cases with $B \sim 1$ is rather problematic as the details of the disk-to-secondary-dominated transition have not been explored. As a result, at large separations, our results for $r_{\text{ev}}$ may not be very accurate. Despite this complication, we agree with Haiman et al. (2009) that the evolution always becomes secondary-dominated at some stage in all the systems that we look at.

6.2.5. Low-mass Disks can Accelerate a Binary Merger

Circumbinary disks are efficient drivers of binary inspiral not only based on the timescale argument presented in Section 6.2.1. Under the assumption of the secondary-dominated evolution, $B \ll 1$, the disk can be significantly lighter than the secondary and still cause the binary to merge. Indeed, Equation (66) states that the minimum amount of mass in a disk capable of driving a binary to merge is $\sim B^{(q+1)/2} M_\star \ll M_\star$. In the gas pressure-dominated regime, the minimum disk-to-secondary mass ratio is $\sim B^{5/19}$ for $\kappa = \kappa_{\text{es}}$ and $\sim B^{2/7}$ for $\kappa = \kappa_{\text{ff}}$.

Evolutionary tracks in Figure 6 confirm this fact: a $q = 1$, $M_\star = 10^7 M_\odot$ system starting at 0.01 pc in a disk accreting at $M_\infty = M_{\text{Edd}} \approx 0.2 M_\odot$ yr$^{-1}$ merges in $t_m \approx 7$ Myr ($B \approx 7 \times 10^{-7}$; see Figure 7(c)). During this time, the minimum mass $\Delta M \sim M_{\infty} t_m \approx 1.4 \times 10^6 M_\odot = 14\% M_\star$ passes through the disk, i.e., less than $M_\star$. Systems accreting at lower rates require even lower disk masses to merge; for instance, lowering $M_{\infty}$ to 0.01$M_{\text{Edd}}$ for the same system ($B \approx 2 \times 10^{-4}$) implies a minimum $\Delta M \approx 5\% M_\star$ (although the merger time becomes $\sim 2.5 \times 10^5$ yr).

The efficiency of gas-assisted evolution in terms of disk mass is ultimately related to the non-local nature of the binary–disk coupling: a binary shares its angular momentum with the distant parts of the disk located at $r_{\text{inf}}$, where the specific angular momentum is higher than at $r_0$. As a result, less mass is required to absorb the angular momentum lost by the binary and $\Delta M$ can be less than $M_\star$.

6.2.6. Importance of the Eddington Limit

We next address the importance of the Eddington limit (42) for SMBH binary evolution. In Figure 6, we plot the radius $r_{\text{Edd}}$ at which the Eddington limit becomes important as a dotted curve. This dependence can be easily derived from Equation (42):

$$r_{\text{Edd}}(t) = \left[ \frac{\kappa_{\text{es}} F_J(t)}{2 \pi c (GM_\star)^{1/2}} \right]^{2/3} = \left( \frac{\kappa_{\text{es}} M_\infty}{2 \pi c} \right)^{2/3} r_{\text{inf}}^{1/3}(t),$$

(72)
where we take the value of $F_j$ in the inner part of the disk from Equation (69) with $r_{\text{in}}(t)$ plotted in the same figure. The Eddington limit is important whenever $r_0 < r_{\text{Edd}}(t)$.

Figure 6 demonstrates that for our choices of $M_c$, $q$, $M_\infty$, and $r_0(0)$, the Eddington limit is essentially irrelevant during the disk-driven phase of the orbital evolution of the binary: $r_{\text{Edd}}(t)$ always passes below or close to the square dots marking the transition to the GW-dominated regime. This fact provides justification for our calculations since the details of the disk physics, including the advent of the Eddington limit, do not affect the binary evolution after it switches to the GW-dominated regime. On the other hand, for all evolutionary tracks depicted in Figure 6, the Eddington limit does become important at some point and affects the disk properties right before the binary merger. This result may have important implications for the electromagnetic precursor of the merger.

It is worth stressing that in a circumbinary disk the Eddington limit can be important even if $M_\infty \ll M_{\text{Edd}}$; see Figures 6(b) and (d). This fact is another consequence of mass accumulation at the inner edge of the disk resulting in a higher temperature and a larger radiation pressure than in a constant $M$ disk. At the same time, the Eddington limit is still definitely less important for low $M_\infty / M_{\text{Edd}}$: for $M_\infty = 10^{-2} M_{\text{Edd}}$, it kicks in only at distances of several tens of $R_0$, while for $M_\infty = M_{\text{Edd}}$ it becomes important as soon as the binary enters the GW-dominated regime.

6.2.7. Gas Overflow across the Orbit of the Secondary

The assumption $M(r_m) = 0$ adopted throughout most of this work, i.e., no gas overflow across the orbit of the secondary, is equivalent to demanding that the width of the orbit of the secondary and the inner edge of the disk $\Delta$ be larger than the disk scale height $h$. Indeed, the torque density produced by the secondary drops for $|r - r_0| \lesssim h$ due to the phenomenon of “torque cutoff” (Goldreich & Tremaine 1980), implying that the secondary can effectively repel the disk fluid only if the gap width satisfies $\Delta \gtrsim h$. When this condition is not fulfilled, gas enters the torque cutoff zone where the tidal repulsion is no longer effective and starts overflowing the orbit of the secondary (Kocsis et al. 2012a). As a result, the inner BC in the form $M(r_0) \ll M_\infty$ may be violated.

Assuming axisymmetry for now and following SC95, we can estimate the gap width as

$$\frac{\Delta}{h} \sim \left[ \frac{q^2}{\alpha \beta^5} \left( \frac{r}{h} \right)^5 \right]^{1/3}, \tag{73}$$

where we assume the viscosity to be given by Equation (25). In both the gas pressure and the radiation pressure-dominated regimes with $b = 0$ (which we adopt in our calculations), one obtains the same value of $\Delta$ as SC95. However, in the radiation pressure-dominated regime with $b = 1$ (as adopted, e.g., by Kocsis et al. 2012a, 2012b), the gap is wider by a factor of $\beta^{-1/3}$, making overflow less likely.

Liu & Shapiro (2010) calculated the mass accretion rate across the orbit of the secondary using a local steady state model for the disk structure. They found that $M(r_m)$ is exponentially small when the factor7 in square brackets in Equation (73) is large ($\gtrsim 5$–10). On the contrary, when this factor is $\lesssim 5$, the mass accretion rate across the orbit of the secondary is found to be close to $M_\infty$. This result provides justification for our use of the

BC $M(r_0) = 0$ ($\chi = 0$) whenever $\Delta/h \gtrsim 1$, and for assuming that overflow occurs when $\Delta/h \lesssim 1$.

Condition (73) allows us to find the value of $r$ at which $\Delta/h = 1$, which we call the “overflow” radius $r_{\text{of}}$. Assuming the disk around the binary to be in the radiation pressure-dominated regime (as is always the case in our calculations when overflow starts; see Figure 6) with $b = 0$ and $h/r$ given by Equation (31), we find that $\Delta/h = 1$ at

$$r_{\text{of}} \approx r_{\text{Edd}} \left( \frac{\alpha}{q} \right)^{2/15},$$

where $r_{\text{Edd}}$ is given by Equation (72).

It is clear from this expression that for equal-mass binaries overflow occurs only after the Eddington limit becomes important since then $r_{\text{of}} \lesssim r_{\text{Edd}}$. On the other hand, as long as $q \lesssim \alpha^{1/2}$, one finds $r_{\text{of}} \gtrsim r_{\text{Edd}}$ but the actual value of $r_{\text{of}}$ never deviates too much (i.e., not by orders of magnitude) from $r_{\text{Edd}}$ because of the weak dependence of $r_{\text{of}}$ on $q$ and $\alpha$. Indeed, for $\alpha = 0.1$ and $q = 10^{-2}$, one finds $r_{\text{of}} \approx 2.5 r_{\text{Edd}}$.

In Figure 6, we show the position of $r_{\text{of}}$ for each evolutionary track with open pentagons. The positions always lie close to the dot-dashed curve showing the Eddington limit (72). Another important observation that can be made by inspecting this figure is that, at least for $q > 10^{-2}$, overflow always occurs close to the start of the GW-dominated phase of the orbital evolution of the binary. Indeed, only for $q = 10^{-2}$ and only in massive disks with $M_\infty = M_{\text{Edd}}$ do we find that overflow precedes (by only a factor of $\sim 2$ in terms of $r_0$) the stage of the GW-driven evolution; see Figures 6(a) and (c). Both for $q \sim 1$ and for $M_\infty \ll M_{\text{Edd}}$ (essentially irrespective of $q$), overflow occurs when the orbital evolution of the binary is already fully determined by the GW emission.

Kocsis et al. (2012a, 2012b) have derived quasi-steady solutions for the disk structure in the presence of the overflow, reminiscent of the SC95 results. For these solutions to become valid after the overflow begins, information on the changing BC at the inner edge of the disk must propagate to the current radius of influence $r_{\text{infl}}$ where the unperturbed, standard constant $M$ disk starts. Otherwise, the solution does not converge to a standard constant $M = M_\infty$ solution at large radii and $M$ at the inner edge of the disk cannot be assumed to be equal to $M_\infty$.

Establishing a connection to the outer disk takes of the order of the viscous time at $r_{\text{infl}}$, which is about the age of the system in our calculations. Using Figures 6 and 7, one can easily see that unless $q \lesssim 10^{-2}$ and $M_\infty = M_{\text{Edd}}$, the orbit of the binary evolves on a much shorter timescale than the viscous time at $r_{\text{infl}}$.

Indeed, a disk with $M_\infty = M_{\text{Edd}}$ around a binary with $M_c = 10^7 M_\odot$, $q = 10^{-2}$ starting at 0.01 pc begins to overflow when its period is $\approx 0.1$ yr (see Figure 6(a)) and the inspiral timescale is $\approx 7 \times 10^7$ yr (Figure 7(a)). This time is almost the same as the viscous time at $r_{\text{infl}}$ for the corresponding evolutionary track. Thus, a quasi-steady solution can be marginally valid in this case. But, if we now look at a $q = 1$ binary keeping everything else the same, we find overflow to occur at $P_{\text{orb}} = 0.03$ yr, when $t_{\text{ev}} \approx 10^3$ yr, which is much shorter than the viscous time at $r_{\text{infl}}$ ($\sim 10^7$ yr). As a result, a global quasi-steady solution does not get established in this case. A similar situation occurs for $M_\infty \ll M_{\text{Edd}}$ (and arbitrary $q$).

To summarize, in the axisymmetric approximation, overflow can be important for low $q$, high-$M_\infty$ systems but only close to the point when the binary inspiral is dominated by the GW emission. These conclusions are reached in a setup most favorable for

---

7 Liu & Shapiro (2010) call this factor $\varrho$. 
the emergence of the overflow—a radiation pressure-dominated disk with $\nu$ proportional to the radiation pressure $p_r$. If instead $\nu$ scales with the gas pressure, overflow is going to be even less important for the gas-assisted SMBH binary evolution.

The assumption of axisymmetry invoked in deriving Equation (73) and used by Liu & Shapiro (2010) and Kocsis et al. (2012a) may be violated in real systems, allowing gas flow into the central cavity even when conditions for axisymmetric overflow are not fulfilled, as suggested by simulations (MacFadyen & Milosavljević 2008; Cuadra et al. 2009; Roedig et al. 2011, 2012; Shi et al. 2012). In particular, MacFadyen & Milosavljević (2008) find that binary torques do not completely prevent accretion across the inner cavity but only reduce it to $\lesssim 0.2 M_\infty$. Such modest overflow will only weakly affect the orbital evolution of the binary: for $\chi = M(0)/M_\infty = 0.2$ the limiting value of $F_J(0)$, which sets the torque on the binary, is different from the $\chi = 0$ (no overflow) case by $\lesssim 25\%$; see Figure 3. Thus, all our results for the binary evolution and disk properties remain largely unchanged even when non-axisymmetric overflow at the level of 0.2 $M_\infty$ is allowed. What does change is the observational appearance of the system since the amount of overflow controls the accretion luminosity of the SMBHs.

6.2.8. Implications for the Gravitational Wave Signatures of SMBH Binary Mergers

Our results have interesting implications for future space-based gravitational wave antennae such as eLISA/NGO (Amaro-Seoane et al. 2012, 2013). If we adopt 0.03 mHz ($\approx 10^{-3}$ yr$^{-1}$) as the characteristic lowest frequency probed by such experiments, then, according to Figures 7(a) and (b), equal-mass binaries with $M_c \gtrsim 10^4 M_\odot$ are detectable only when their orbital decay is already fully dominated by GW emission, even for massive disks with $M_\infty = M_{\text{Edd}}$.

However, low-$q$ systems including the so-called intermediate-mass ratio inspirals (IMRIs) can enter the detection band of the GW experiments during the stage when their orbital evolution is still dominated by the tidal coupling to the disk. An example of this result can be seen in Figure 7(b), where the SMBH binaries with $q = 10^{-2}$ are pushed by the disk with $M_\infty = M_{\text{Edd}}$ all the way until the transition to the GW-dominated regime occurs at $P_{\text{orb}} \approx 3$ yr. The decay of systems with even lower $q$ or $M_c$ can be dominated by their disks down to even shorter orbital periods, making detection of disk-driven migration quite plausible for low-$M_c$ IMRIs (provided that the strain they produce is above the signal-noise of the GW antenna).

Disk effects on the GW signal manifest themselves via the orbital phase shift of the binary caused by the variation of its semi-major axis due to the disk-driven migration; see Kocsis et al. (2011) and Yunes et al. (2011). Even if signatures of the disk-driven migration were found in the GW signal of coalescing binaries, it is unlikely that one will be able to use these measurements to probe the properties of the radiation pressure-dominated part of the disk in the immediate vicinity of the binary. The reason for that again lies in the non-local nature of the torque acting on the binary: close to $P_{\text{orb}} \approx 3$ yr, the evolution of all low-$M_c$ binaries shown in Figures 6(c) and (d) is typically dominated by torques set at $r_{\text{inf}}$ located in the gas pressure-dominated regime. Thus, the GW phase signal will inform us only of the properties of the gas pressure-dominated part of the disk.

We also note in this regard that calculations of the GW shifts presented in Kocsis et al. (2011) and Yunes et al. (2011) should be revised to account for this non-locality of the disk torques on the binary.

6.3. Spectra of Disks around SMBH Binaries

Circumbinary disks can be observationally distinguished from their regular constant $M$ counterparts (SC95). We illustrate the difference in Figure 9 by showing the spectral energy distribution (SED) of a circumbinary disk at different evolutionary stages of the disk+binary system with $q = 1$, $M_c = 10^5 M_\odot$, $r_{\text{inf}}(0) = 10^{-3}$ pc, and $M_\infty = 10^{-2} M_{\text{Edd}}$; see Figure 6(d) for the corresponding evolutionary track. One set of spectra shown by thick curves in Figure 9 assumes $F_J(r, t)$ in the form (69) with $r_{\text{inf}}(t)$ taken from the self-consistent calculations presented in Figure 6(d). Another set of SEDs (thin curves) computed at the same moments of time (and for the same values of $r_{\text{inf}}(t)$) is for a constant standard $M = M_\infty$ disk with $F_J(r, t) = M_\infty (GM_r t)^{1/2}$, which is truncated at $r_b$. Both kinds of calculations assume the outer edge of the disk lies at 0.1 pc (in this calculation we disregard complications arising at large separations, which were mentioned in Section 4.1.5; the choice of the outer radius is not important to us). We now go over the details of these calculations.

After $t = 10^3$ yr of evolution, the binary semi-major axis has essentially not changed and the radius of influence has extended only out to $3 \times 10^{-4}$ pc. Since this value of $r_{\text{inf}}$ is close to $r_b$, binary torques affect only the very innermost part of the disk and the spectra computed in two ways (thick and thin long-dashed curves in Figure 9) do not show a significant difference. At
$\lambda = 10-500 \mu m$, both are well fit by a power-law $F_\nu \propto \nu^{4/3}$ typical for a constant $M$ disk.

At $t = 10^6$ yr, $r_b$ is still very close to $r_b(0)$, but the effects of the binary torque have been viscously transmitted through the disk out to $r_{infl} \approx 1.7 \times 10^{-3}$ pc. This fact results in a factor of $\approx 4$ difference in the torque acting on the binary in two cases, and noticeably changes the spectrum of the disk in a self-consistent calculation: the peak wavelength of the spectrum shifts to a slightly shorter wavelength and the peak amplitude of $F_\nu$ increases by a factor of two compared to $t = 10^5$ yr. Note that the latter is solely due to the rearrangement in the disk structure due to binary torques—the inner disk radius stays essentially the same in these two epochs. For the same reason, there is no difference in the spectra computed assuming a constant $M$ disk (the thin curves for $t = 10^5$ and $10^6$ essentially overlap in Figure 9).

At $t = 10^7$ yr, $r_b$ has shrunk to $8.7 \times 10^{-5}$ pc, while $r_{infl} \approx 0.03$ pc. The constant $M$ disk SED has changed—its peak is now around $1 \mu m$ (thin solid line) and the maximum of $F_\nu$ is about an order of magnitude higher than before. But from now on, the disk around the binary is in gas pressure-dominated regime with $\kappa = \kappa_e$, see Figure 6(d). While the spectrum of a constant $M$ disk is almost the same at this epoch, the SED of a self-consistently evolved disk (thick dotted curve) exhibits not only an increase in amplitude and a shift toward shorter wavelengths, but also a steepening at $\lambda = 10-500 \mu m$. All of these results are still predominantly due to the evolution of the radial structure of the disk under the action of the binary torque.

Finally, at $t \approx 4 \times 10^7$ pc, the binary orbit shrinks to $10^{-3}$ pc (orbital evolution is now dominated by the GW emission), while $r_{infl} \approx 0.03$ pc. The constant $M$ disk SED has changed—its peak is now around $1 \mu m$ (thin solid line) and the maximum of $F_\nu$ is about an order of magnitude higher than before. But the SED of a self-consistently evolved disk changes far more dramatically—it now peaks in the optical at $0.3 \mu m$ and the peak value of $F_\nu$ is $\approx 40$ times higher than for a constant $M$ disk. Between $1 \mu m$ and $500 \mu m$, the shape of the SED is well fit by $F_\nu \propto \nu^{12/7}$, as expected for a constant $F_J$ disk (SC95), now occupying the inner third of the radial extent of the disk and accounting for most of its luminosity.

To summarize, the SED of a disk affected by the torque of a central binary is steeper, brighter, and extends to shorter wavelengths than the SED of its constant $M$ counterpart having the same inner radius and mass accretion rate $M_\infty$ at large distances. These features (especially the steepness of the spectrum) provide a way of inferring the existence of an SMBH binary surrounded by the disk based on broadband spectroscopy alone. This method is distinct from other indicators of the SMBH binary presence, both spectroscopic—the double-peak lines due to binary components and/or disk (Bogdanović et al. 2009) and the time-variable velocity shifts of single quasar lines (Tsamantza et al. 2011; Eracleous et al. 2012; Ju et al. 2013)—and photometric (Sesana et al. 2012). This method may be the only way of inferring the presence of an SMBH binary with face-on orientation or a wide binary not showing evolution of spectral lines. The relative brightness of circumbinary disks should facilitate the detection of such systems out to large distances.

6.3.1. Sensitivity of the SED to Mass Inflow across the Secondary Orbit

We now relax our assumption of $M(r_{infl}) = 0$ and see how this procedure affects the SED of the circumbinary disk. If there is overflow, then one can still construct both the steady state and the evolving self-similar solutions as demonstrated in Sections 3 and 5.1 and extend the results for the orbital evolution of the SMBH binary presented in Sections 6.1 and 6.2. Instead of Equation (69), we now approximate the spatial distribution of $F_J$ by the following simple formula:

$$F_J(r) = \begin{cases} M_\infty/r_{infl} & r > r_{infl}, \\ M_\infty[l_{infl} - \chi l_{infl} - l(r)]/r \leq r_{infl}, \end{cases}$$

where $l(r) = (G M/r)^{1/2}$ and $l_{infl} = (G M r_{infl})^{1/2}$. This prescription consists of two steady state solutions continuously matched at $r = r_{infl}$ but with different $M$ inside and outside of this point. For $\chi = 0$, this formula naturally reduces to Equation (69).

In Figure 10, we show disk spectra computed assuming different values of $\chi = M(r_{infl})/M_\infty \leq 1$. Our calculations assume a system for which a spectrum is shown in Figure 9 at $t = 4 \times 10^7$ yr for $M(r_{infl}) = 0$, i.e., an equal-mass, $M_e = 10^5 M_\odot$ SMBH binary at $r_b = 10^{-3}$ pc surrounded by a disk with $M_\infty = 10^{-2} M_{edd}$ and $r_{infl} \approx 0.03$ pc. We show the SED of only the circumbinary disk, i.e., in this calculation we again do not account for the emission produced by accretion disk(s) around the primary and/or secondary, which should form when mass flows across the orbit of the secondary.

One can see that as the transparency of the tidal barrier $\chi$ increases toward unity the SED approaches that of a constant $M = M_\infty$ disk. This result is not at all surprising because in the limit of $\chi \rightarrow 1$ the disk structure reduces to that of a constant $M$ disk; see Figure 3. However, it is also clear from Figure 10 that the SED is strongly affected compared to the case of a constant $M$ ($\chi = 1$) disk even if only a small amount of inflowing mass is accumulated at the inner edge of the disk. For example, the peak amplitude of the disk spectrum for $\chi = 0.98$ (implying that only 2% of the accreting mass gets stopped by the binary
torques) is a factor of two higher than in the \( \chi = 1 \) case. If the tidal barrier allows penetration of only 30\% of the gas across the gap, the SED of the circumbinary disk is hardly distinguishable from that of a disk with no gas inflow at \( r_{in} \).

These results imply that the broadband SED of the disk is a rather sensitive measure of even a small amount of matter penetrating into the cavity cleared by the SMBH binary. Coupled with the measurements of the SED of the accretion disk(s), which may form around each of the binary components if \( \chi \neq 0 \), these observations can inform us of the efficiency of the binary torques at clearing a clean cavity at the center of the system.

7. SUMMARY

In this work, we explored the coupled evolution of an SMBH binary and its gaseous disk. Disk properties (surface density, temperature, etc.) evolve under the action of binary torque, which constrains the flow in the inner part of the disk. To study this problem, we have reformulated evolution equations in terms of the angular momentum flux \( F_J \), which significantly simplifies the mathematical description of the problem. We derived the disk properties as a function of \( F_J \) in different physical regimes that are important in circumbinary disks around SMBH binaries.

When the external mass supply to the disk at large distances is not matched at the inner edge of the disk because of the barrier presented by the binary torques, the disk evolves toward establishing a quasi-steady state in the inner region, where the local viscous timescale is shorter than the evolution time of the system. The viscous angular momentum flux in the inner disk steadily grows with time, accelerating the orbital evolution of the binary.

We explored the dependence of this general picture on the system parameters (mass of the binary, mass accretion rate through the disk, etc.) and found the following results, in agreement with previous studies:

1. Tidal coupling to a circumbinary disk can substantially shorten (by orders of magnitude) the lifetime of the binary (IPP; Lodato et al. 2009; Haiman et al. 2009).
2. For a long time before the GW emission dominates, the mass of the secondary is larger than the local disk mass (Haiman et al. 2009).
3. Disk-driven evolution of the binary can be measured by space-based gravitational wave antennae for low-\( q \) systems with relatively low \( M_\star \) (Kocsis et al. 2011; Yunes et al. 2011).
4. The spectrum of the disk affected by the binary torques is different from that of a conventional constant \( M \) disk: it extends to shorter wavelengths, is steeper at intermediate wavelengths (SC95), and more power is emitted.

We also obtained a number of new results, summarized below.

1. The self-consistent evolution of the disk resulting in a pileup of mass at its inner edge accelerates the orbital evolution of the binary compared to some previous calculations.
2. Even low-mass disks (less massive than \( M_\star \)) can effectively drive binary mergers, as long as the system starts in the secondary-dominated regime \( B \ll 1 \).
3. Disk–binary coupling has a non-local character: torques acting on the binary are determined by the state of the disk far from the binary, at the radius of influence \( r_{inf} \), which steadily increases in time, rather than by the disk properties in the immediate vicinity of the binary.
4. The evolution of the binary orbits exhibits a phenomenon of hysteresis—it depends on the past history of the system, which is caused by the non-locality of the disk–binary coupling.
5. The Eddington limit can strongly affect the disk structure even when the mass accretion rate at large distances is considerably sub-Eddington, but it is unlikely to affect the binary inspiral.
6. Overflow across the orbit of the secondary becomes possible mainly (or only) during the GW-dominated phase of inspiral and is most important for low-\( q \), high-\( M_\infty \) systems.
7. Spectra of circumbinary disks strongly depend on the degree of mass overflow across the orbit of the secondary.

This list clearly implies that properly accounting for the fully self-consistent, time-dependent evolution of circumbinary disks is crucial for understanding gas-assisted SMBH mergers. This general conclusion will hopefully inspire re-evaluation of some of the existing results for the orbital evolution of SMBH binaries and their observational manifestations, both in the electromagnetic and the GW domains. Results of this work can also be extended to studying circumbinary disks around stellar mass binaries.

We are indebted to Pavel Ivanov, Bence Kocsis, and Zoltan Haiman for a careful reading of the manuscript, an open exchange of opinions, and a number of useful suggestions. We are grateful to Zoltan Haiman for suggesting the overflow calculations presented in Section 6.2.7. The financial support for this work is provided by the Sloan Foundation, NASA grant NNX08AH87G, and NSF grant AST-0908269.

APPENDIX A

EVOLUTION WITH AN ALTERNATIVE INNER BOUNDARY CONDITION

We consider the evolution of the disk+binary system with an inner BC different from Equation (14). Following SC95, we assume that the inward motion of the secondary gives rise to a continuous density redistribution in the inner disk resulting in the mass flux \( 2\pi r_b v_b \Sigma(r_b) \) at \( r = r_b \), assuming that the inner edge of the disk closely follows the binary orbit. Expressing \( v_b \) and \( \Sigma(r_b) \) via \( F_J \) and \( D_J \) with the aid of Equations (2), (5), (8), and (11), and allowing an additional mass flux \( \chi M_\infty \) across the orbit of the secondary, we come up with a new BC:

\[
\frac{\partial F_J}{\partial t} \bigg|_{t=\tau_b} = \dot{\chi}_{\tau_b} = \chi \dot{M}_\infty = \frac{(1 + q)F_J}{M_\star D_J}, \tag{A1}
\]

which replaces Equation (14). Since in general \( D_J \) is a function of \( F_J \) and \( r_b \), \( \tau_b \) varies as a result of the orbital evolution of the secondary. Equation (A1) represents a nonlinear and time-dependent inner BC for Equation (4).

We now show that in astrophysically relevant setups the secondary-dominated evolution with a BC in the form of (A1) reduces to that using the BC (14). For simplicity, we restrict ourselves to the no-overflow case \( \chi = 0 \) and demonstrate that \( \dot{M}(r_b) \ll \dot{M}_\infty \), in which case the disk–binary system should behave as if it were governed by the BC (14) with \( \chi = 0 \); see Section 5.1.

As always, we start with a constant \( M = M_\infty \) disk extending all the way to the initial semi-major axis of the binary \( r_{b,0} \) (corresponding value of the initial specific angular momentum
We assume that \( M(r_b) \ll M_\infty \), and then demonstrate this assumption to be the case. With this assumption, and for \( D_J \) given by Equation (52), the disk evolution is described by solutions (60) and (61) for \( \tau_{\text{infl}} \) and \( F_J \). In the latter, we take \( f(l/l_{\text{infl}} \to 0) \approx 1 \) (justified when \( M(r_b) \ll M_\infty \), i.e., \( F_J \approx M_\infty / l_{\text{infl}} \)). Then, it is easy to show that Equation (A1) can be written, after some manipulations, as

\[
\dot{M}(r_b) \approx (1 + q)\dot{M}_\infty B \left( l_{\text{infl}} / b_0, \right)^{2-d} \left( l_b / b_0 \right)^{p} , \tag{A2}
\]

where \( B \ll 1 \) is defined by Equation (63). We also calculate the ratio of the binary evolution timescale \( t_J \) (Equation (12)) to the viscous timescale \( \tau_{v,\text{infl}} = l_{\text{infl}} / D_J \) at the radius of influence, where the torque acting on the secondary is determined:

\[
t_J / \tau_{v,\text{infl}} \sim B^{-1} l_b / l_{\text{infl}} \left( l_{\text{infl}} / b_0 \right)^{-(n+1)/n} . \tag{A3}
\]

At the beginning of the secondary-dominated (\( B \ll 1 \)) evolution, \( \tau_{\text{infl}} = l_b = l_{b,0} \) and \( M(r_b) / M_\infty = B \ll 1 \), while \( t_J / \tau_{v,\text{infl}} \sim B^{-1} \ll 1 \). As the system evolves, \( l_{\text{infl}} \) increases but even for the highest \( l_{\text{infl}} = l_{\text{infl},0} \) corresponding to the binary merger (Equation (64)), one finds \( B(l_{\text{infl},0} / b_0)^{2-d} \sim B^{n-p(n+1)/n} \ll 1 \), since \( p < 0 \) and \( n > 1 \); see Appendix C.

Thus, starting in the disk-driven, secondary-dominated regime, \( M(r_b) \) is guaranteed to remain below \( M_\infty \) (contrary to the assumption by SC95). At \( l \sim l_{\text{infl},0} \), one also finds \( t_J / \tau_{v,\text{infl}} \sim l_b / l_{\text{infl},0} \).

As the binary orbit begins shrinking and \( l_b / l_{\text{infl},0} \to 0 \), the ratio \( M(r_b) / M_\infty \) rapidly decreases (because \( p < 0 \)); see Section 6.2.4 and Figures 7(c) and (d). Also, \( t_J / \tau_{v,\text{infl}} \to 0 \), meaning that the disk at \( r_{\text{infl}} \) is no longer causally connected to the binary by the viscous stresses—the disk at \( r_{\text{infl}} \) does not respond to the binary inspiral. As a result, the torque acting on the binary during this phase freezes at \( F_J \sim M_\infty / l_{\text{infl},0} \), as can be seen in Figures 5 and 6.

Two important conclusions can be drawn from this exercise for systems starting in the secondary-dominated regime (\( B \ll 1 \)):

1. They always have mass accretion rates in the inner portion of the disk \( M(r_b) \ll M_\infty \).
2. After the binary significantly evolves away from its original orbit, the inspiral timescale \( t_J \) is much shorter than the viscous evolution timescale \( \tau_{v,\text{infl}} \) at the radius of influence.

APPENDIX B

SCALING RELATIONS FOR AN ARBITRARY POWER-LAW OPACITY

Here, we summarize the scaling relations for the disk properties that result when the opacity is a power-law function of the gas temperature \( T \) and the density \( \rho \), represented by their midplane values (see Lyubarskij & Shakura 1987 for similar treatment):

\[
\kappa = \kappa_0 \rho^{\mu_1} T^{\mu_2} . \tag{B1}
\]

We make two additional assumptions regarding disk properties: (1) the disk is optically thick and (2) radiation pressure is negligible compared to the gas pressure (the radiation pressure-dominated case is described by Equations 31–35)). Combining Equations (2), (21), (27), and (28), and \( \nu = \alpha c_s^2 / \Omega \) valid in the gas pressure-dominated regime, one finds

\[
\Sigma(r) = \left[ \frac{2^{(4+\mu_1)} \kappa_0}{3^{10+\mu_1-2\mu_2} \sigma \kappa_0 / \sigma} \left( \frac{\kappa}{\kappa_0} \right)^{2(4-\mu_2)} \right. \tag{B2}
\]

\[
\times \left. \frac{\kappa}{\sigma} \left( \frac{\kappa}{\sigma} \right)^{2(4-\mu_2)} \right] r^{-(\eta-\mu_1-4\mu_2)/\epsilon} , \tag{B3}
\]

\[
\frac{\dot{h}(r)}{r} = \frac{2^{-5+\mu_1}}{3^{10+\mu_1-2\mu_2}} \left( \frac{\kappa}{\sigma} \right)^{4-\mu_2} \tag{B4}
\]

\[
\times \frac{\kappa}{\sigma} \left( \frac{\kappa}{\sigma} \right)^{4-\mu_2} \frac{\dot{F}_J}{\dot{F}_{J,0}} \left[ \sigma / \kappa_0 \right]^{2(10+3\mu_1-2\mu_2)} , \tag{B5}
\]

\[
d = \frac{2(2 + \mu_1)}{10 + 3\mu_1 - 2\mu_2} , \quad p = -\frac{12 + 11\mu_1 + 2\mu_2}{10 + 3\mu_1 - 2\mu_2} . \tag{B6}
\]

APPENDIX C

SUMMARY OF THE DIFFUSION COEFFICIENT BEHAVIOR IN DIFFERENT REGIMES

Here, we summarize the behavior of the diffusion coefficient \( D_J \) in the power-law form (52) and the behavior of the self-similar exponent \( n \) defined by Equation (57) for different objects and in different regimes explored in this work (see also Lyubarskij & Shakura 1987).

In the radiation pressure-dominated case (Section 4.1.1), one finds, for \( b = 0 \):

\[
D_{J,0} = \frac{3}{4} \frac{\kappa_0^2 (GM_\infty)^4}{\alpha^2 c_s^2} , \quad d = 2 , \quad p = -7 , \quad n = \frac{1}{7} . \tag{C1}
\]

For \( b = 1 \), one finds:

\[
D_{J,0} = \frac{3}{4} \left[ \frac{1}{24\pi^2} \frac{\kappa}{\sigma} \left( \frac{\kappa}{\sigma} \right)^{4} \left( \frac{GM_\infty^4 \sigma}{\kappa} \right)^{1/5} \right]^{1/5} \tag{C2}
\]

\[
d = \frac{2}{5} , \quad p = -\frac{6}{5} , \quad n = \frac{5}{14} . \tag{C2}
\]
Expression (C2) also holds true for the gas pressure-dominated case with $\kappa = \kappa_{\text{es}}$ (Section 4.1.2).

In the gas pressure-dominated case with $\kappa = \kappa_{\text{ff}}$ (Section 4.1.3):

$$D_{J,0} = \frac{3}{4} \left( GM_c \right) \left[ \frac{2^{-5} \kappa_0}{3 \pi^3 \sigma} \left( \frac{k}{\mu} \right)^{15/2} \alpha^8 \right]^{1/10},$$

$$d = \frac{3}{10}, \quad p = -\frac{4}{5}, \quad n = \frac{2}{5}.$$  \hspace{1cm} (C3)

**REFERENCES**

Amaro-Seoane, P., Aoudia, S., Babak, S., et al. 2012, CQGra, 29, 124016

Amaro-Seoane, P., Aoudia, S., Babak, S., et al. 2013, GW Notes, 6, 4

Armitage, P. J., & Natarajan, P. 2002, ApJL, 567, L9

Baruteau, C., & Masset, F. 2013, Tides in Astronomy and Astrophysics (Lecture Notes in Physics, Vol. 861; Berlin: Springer), 201

Bogdanović, T., Eracleous, M., & Sigurdsson, S. 2009, ApJ, 697, 288

Chang, P., Strubbe, L. E., Menou, K., & Quataert, E. 2010, MNRAS, 407, 2007

Cuadra, J., Armitage, P. J., Alexander, R. D., & Bergman, M. C. 2009, MNRAS, 393, 1423

Duffell, P. C., & MacFadyen, A. I. 2013, ApJ, 769, 41

Eracleous, M., Boroson, T. A., Halpern, J. P., & Liu, J. 2012, ApJS, 201, 23

Filipov, L. G. 1984, AdSpR, 3, 305

Goldreich, P., & Lynden-Bell, D. 1965, MNRAS, 130, 97

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

Goodman, J. 2003, MNRAS, 339, 937

Haiman, Z., Kocsis, B., & Menou, K. 2009, ApJ, 700, 1952

Hirotæ, S., Blaes, O., & Krolik, J. H. 2009a, ApJ, 704, 781

Hirotæ, S., Krolik, J. H., & Blaes, O. 2009b, ApJ, 691, 16

Ivanov, P. B., Papaloizou, J. C. B., & Polnarev, A. G. 1999, MNRAS, 307, 79 (IPP)

Ju, W., Greene, J. E., Rafikov, R. R., Bickerton, S. J., & Badenes, C. 2013, arXiv:1306.4987

Khan, F. M., Just, A., & Merritt, D. 2011, ApJ, 732, 89

Kocsis, B., Haiman, Z., & Loeb, A. 2012a, MNRAS, 427, 2660

Kocsis, B., Haiman, Z., & Loeb, A. 2012b, MNRAS, 427, 2680

Kocsis, B., Yunes, N., & Loeb, A. 2011, PhysD, 84, 024032

Lightman, A. P., & Eardley, D. M. 1974, ApJL, 187, L1

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

Lin, D. N. C., & Papaloizou, J. 1996, ARA&A, 34, 703

Lipunova, G. V., & Shakura, N. I. 2000, A&A, 356, 363

Liu, Y. T., & Shapiro, S. L. 2010, PhysD, 82, 123011

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

Lynden-Bell, D., & Pringle, J. E. 1974, MNRAS, 168, 603 (LB74)

Lyubarskij, Y. E., & Shakura, N. I. 1987, SvAL, 13, 386

MacFadyen, A. I., & Milosavljević, M. 2008, ApJ, 672, 83

Milosavljević, M., & Phinney, E. S. 2005, ApJL, 622, L93

Nixon, C. J., Cossins, P. J., King, A. R., & Pringle, J. E. 2011, MNRAS, 412, 1591

Novikov, I. D., & Thorne, K. S. 1973, in Black Holes, ed. C. DeWitt & B. S. DeWitt (New York: Gordon & Breach), 343

Papaloizou, J., & Lin, D. N. C. 1984, ApJ, 285, 818

Petrovich, C., & Rafikov, R. R. 2012, ApJ, 758, 33

Preto, M., Benz, T., & Huetz, F. 2011, ApJ, 732, L26

Pringle, J. E. 1991, MNRAS, 248, 754

Rafikov, R. R. 2002, ApJL, 572, 566

Roedig, C., Dotti, M., Sesana, A., & Colpi, M. 2011, MNRAS, 415, 3033

Roedig, C., Sesana, A., & Dotti, M. 2012, A&A, 545, A127

Safonov, V. S. 1960, AnAp, 23, 979

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

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

Shi, J.-M., Krolik, J. H., Lubow, S. H., & Hawley, J. F. 2012, ApJ, 749, 118

Syer, D., & Clarke, C. J. 1995, MNRAS, 277, 758 (SC95)

Toomre, A. 1964, ApJ, 139, 1217

Tsaltmanza, P., Decarli, R., Dotti, M., & Hogg, D. W. 2011, ApJ, 738, 20

Warb, W. R. 1997, Icar, 126, 261

Yu, Q. 2002, MNRAS, 331, 935

Yunes, N., Kocsis, B., Loeb, A., & Haiman, Z. 2011, PhysL, 107, 171103

Rafikov