The Astrophysical Journal, 939:48 (28pp), 2022 November 1 https://doi.org/10.3847/1538-4357/ac93f6

Anatomy of a Slow Merger: Dissecting Secularly Driven Inspirals of LIGO/Virgo Gravitational Wave Sources

Chris Hamilton[1,2] and Roman R. Rafikov[1,2]
[1] Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA; chamilton@ias.edu
[2] Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK
Received 2022 July 5; revised 2022 September 17; accepted 2022 September 20; published 2022 November 2

Abstract

The dozens of compact object mergers detected by LIGO/Virgo raise a key theoretical question: how do initially wide binaries shrink sufficiently quickly that they are able to merge via gravitational wave (GW) radiation within a Hubble time? One promising class of answers involves secular driving of binary eccentricity by some external tidal perturbation. This perturbation can arise due to the presence of a tertiary point mass, in which case the system exhibits Lidov-Kozai (LK) dynamics, or it can stem from the tidal field of the stellar cluster in which the binary orbits. While these secular tide-driven mechanisms have been studied exhaustively in the case of no GW emission, when GWs are included the dynamical behavior is still incompletely understood. In this paper we consider compact object binaries driven to merger via high-eccentricity excitation by (doubly averaged, test-particle quadrupole level) cluster tides—which includes LK-driven mergers as a special case—and include the effects of both general relativistic precession and GW emission. We provide for the first time an analytical understanding of the different evolutionary stages of the binary's semimajor axis, secular oscillation timescale, and phase-space structure all the way to merger. Our results will inform future population synthesis calculations of compact object binary mergers from hierarchical triples and stellar clusters.

Unified Astronomy Thesaurus concepts: Compact binary stars (283); Relativistic binary stars (1386); Gravitational wave sources (677); Celestial mechanics (211); Star clusters (1567); LIGO (920); Astrodynamics (76)

1. Introduction

The LIGO/Virgo Collaboration has now detected around 90 compact object binary mergers (Abbott et al. 2021). However, there is still ambiguity on the theoretical side about which mechanisms drive these mergers.

One much-studied candidate is the Lidov-Kozai (LK) mechanism which operates in hierarchical triple systems (Kozai 1962; Lidov 1962; Naoz 2016), in which a compact object binary is both orbited and tidally torqued by a bound tertiary perturber. In the quadrupole-tide approximation, provided the inclination angle between the two orbital planes of the hierarchical triple system is sufficiently large, the binary's eccentricity e can be driven periodically toward values approaching unity on secular timescales (i.e., much longer than any of the orbital periods in the system).[3] This greatly reduces the pericenter distance $p \equiv a(1-e)$, where a is the binary's semimajor axis. Gravitational wave (GW) radiation is strongly enhanced at small p, i.e., is strongly concentrated around the maximum $e \rightarrow e_{\max} \approx 1$, so there is a "burst" of orbital energy loss at each peak. By energy conservation, this corresponds to a decay in binary semimajor axis by some amount Δa. Such losses accumulate over multiple secular cycles until the binary is so compact that it effectively decouples from the tidal perturbations and undergoes a GW-dominated inspiral. This basic understanding has inspired scores of papers that aim to understand binary black hole mergers and predict their rates (e.g., Blaes et al. 2002; Miller & Hamilton 2002; Wen 2003; Thompson 2011; Antonini & Perets 2012; Antognini et al. 2014; Liu & Lai 2017; Silsbee & Tremaine 2017).

More generally, every binary that resides in a stellar cluster feels the cluster's gravitational potential. As we showed in Hamilton & Rafikov (2019b, 2019c)—hereafter referred to as Papers I and II, respectively—the cluster potential provides a tidal torque on the binary just as a tertiary perturber would. This cluster-tidal torque can drive LK-like eccentricity oscillations in binaries on astrophysically relevant timescales. One can therefore consider the cluster itself to be a ubiquitous "third body" which has the capacity to induce "cluster tide-driven" mergers (Hamilton & Rafikov 2019a; Bub & Petrovich 2020; Arca Sedda 2020). The primary caveat to both the LK-driven and cluster tide-driven merger mechanisms is that 1pN general relativistic (GR) apsidal precession acts to quench the oscillations and hence delay mergers. In Hamilton & Rafikov (2021)—hereafter Paper III—we provided a systematic account of how the LK-like dynamics of binaries are modified as the strength of GR precession is varied.

Guided by the basic understanding outlined above, most of the aforementioned studies either opt for direct numerical integration of the binary equations of motion (e.g., Silsbee & Tremaine 2017), or they aim at a parameterization of the total merger timescale as a function of $e_{\max}$ (e.g., Thompson 2011; Liu & Lai 2018), and very little theoretical understanding has been developed beyond this (but see Randall & Xianyu 2018). On the other hand, a detailed look at numerical integrations suggests that LK-driven and cluster tide-driven mergers are in fact very dynamically rich, even in the doubly averaged, test-particle quadrupole limit[4] that we consider

[3] Octupolar tides can sometimes lead to high-eccentricity excitation even for low initial inclination. We will ignore octupolar effects throughout this paper, but see Section 6 for a brief discussion.

[4] In this limit we average the dynamics over both the binary's inner Keplerian orbit and its barycentric orbit around the cluster/perturber; we assume the outer orbit is unchanging; and we perform a tidal expansion of the perturbation upon the binary only to quadrupole order. For more details, see Antonini et al. (2014), Naoz (2016), and Grishin et al. (2018).

exclusively in this paper. They exhibit nontrivial time evolution of the binary's semimajor axis, maximum and minimum eccentricity, phase-space location, secular period, etc. To demonstrate this, we now provide a numerical example of an LK-driven merger.

1.1. Example of an LK-driven Merger

Figure 1 shows the result of integrating the doubly averaged (hereafter "DA") test-particle quadrupole LK equations of motion for a binary with constituent point masses $m_1 = m_2 = 10M_\odot$ orbiting a supermassive black hole (SMBH) of mass $\mathcal{M} = 4 \times 10^6 M_\odot$. The peri/apocenter of the binary's "outer" orbit around the SMBH, (r_p, r_a), as well as its initial "inner" semimajor axis a, eccentricity e, inclination i, and argument of pericenter ω are given in the text at the top of the figure.[5] In panels (a)–(d) we plot the time evolution of a, e, i, and pericenter distance $p = a(1-e)$ with black lines.

We note here that we have chosen these rather special "initial" conditions by hand, to create an example that has a clean separation of dynamical behaviors so we can most easily elucidate the underlying physics. Of course they are by no means the most likely initial conditions for a black hole binary. For instance, we have purposely chosen an extremely high inclination $i_0 = 89.75°$, which if one assumes statistical isotropy will only be achieved by a fraction $\cos i_0 \sim 4 \times 10^{-3}$ of triple systems.

As expected, the binary undergoes large-amplitude eccentricity oscillations, which drive a decay in semimajor axis and ultimately lead to merger ($a \to 0$). Moreover, the reader will notice three color-shaded vertical stripes in each of these panels (blue, yellow, and green) defining three representative time intervals of the dynamical evolution. In panels (f)–(h) we plot in black the trajectory of the binary in the (ω, e) phase space (familiar from Papers II and III) during those respective segments, starting at the green dot and ending at the red dot. The black arrows indicate the direction of phase-space evolution. Panel (e) shows ϵ_{GR}, the dimensionless measure of the strength of GR precession (defined in Equation (6)). In panels (i) and (j), we show with black dots the secular oscillation timescale t_{sec} and the decay in semimajor axis $|\Delta a|$, respectively, as functions of a, evaluated at the end of each secular oscillation. Panel (k) shows the corresponding orbital decay timescale[6] $\tau_a \equiv |d \ln a/dt|^{-1}$. The binary evolves from right to left in these last three panels, and the blue, yellow, and green shaded segments are again indicated. There are also various critical values and analytical scalings shown with colored lines throughout the Figure, which will be explained in Sections 3–4.

Figure 1 exhibits several striking features, some of which are rarely mentioned in analyses of LK-driven mergers. For instance:

1. Despite the fact that a is changing with time (panel (a)), for most of the evolution there exist two approximately conserved quantities, namely the minimum inclination i_{min} and the minimum pericenter distance p_{min} reached during each secular cycle—see the red dashed lines in panels (c) and (d).
2. The timescale for secular oscillations t_{sec} changes by several orders of magnitude throughout the evolution, and exhibits a highly nontrivial dependence on semimajor axis (panel (i)). Viewed as a function of time (right to left in that panel), it first increases with time (up until the yellow shaded stripe), then is almost constant (up to the green stripe), and thereafter *decreases* with time, ultimately becoming much smaller than its initial value. This decrease is counterintuitive, since naively one might expect a more tightly bound binary to exhibit slower tide-driven secular evolution.
3. The binary's phase-space trajectory evolves as its semimajor axis shrinks. At early times it follows a librating trajectory around the fixed point at $\omega = \pi/2$ (panel (f)), then it transitions to a circulating trajectory (panel (g)), which is ultimately pushed to very high minimum eccentricity (panel (h)).

Of these three observations, the first has been discussed by, e.g., Wen (2003). The second was mentioned by Randall & Xianyu (2018), although their explanation of this phenomenon was incomplete, as we show in Appendix E. The third was briefly signposted by Blaes et al. (2002) and Antonini et al. (2016)—see Section 4. There is very little literature that goes into quantitative detail about these features, and nobody has considered their interplay (for instance, how the changing phase-space structure affects the evolution of the secular timescale).

A central purpose of the present paper is to explain such dynamical characteristics. It will turn out that the qualitatively different behaviors exhibited by $\Delta a(a)$, $t_{sec}(a)$, $\tau_a(a)$, etc. map onto different GR regimes and phase-space features explored in Paper III.

1.2. Plan for the Remainder of This Paper

This paper is organized as follows. In Section 2 we describe briefly the setup of our system without GW emission, and establish notation. We also gather in Appendix A some results from Papers I–III, which we will refer to throughout this work. In Section 3 we introduce the effect of GW emission, and derive two approximate conservation laws, namely conservation of the minimum pericenter distance and the minimum inclination reached by the binary. In Section 4 we outline how a binary evolves through phase space and different GR regimes as its semimajor axis shrinks (much more detail is given in Appendix B). This allows us to establish several asymptotic regimes in which we can make analytic progress. Using these regimes, in Section 4.2 we write down approximate expressions for the secular timescale $t_{sec}(a)$, while in Section 4.3 we derive expressions for $\Delta a(a)$ and $\tau_a(a)$. Again, details of the derivations are relegated to Appendices C–D. In Section 5 we verify our results via several more numerical examples akin to Figure 1, this time for binaries in (non-Keplerian) cluster potentials. In Appendix E we compare our work with the LK calculations of Randall & Xianyu (2018). In Section 6 we

[5] For the precise definitions of these quantities, see Section 2.

[6] We note that the secular timescale was computed by finding the time elapsed between adjacent eccentricity maxima in panel (b), while Δa was computed by calculating the semimajor axis before (a_{bef}) and after (a_{aft}) each peak, and defining $\Delta a(a_{bef}) \equiv a_{aft} - a_{bef}$. To draw the colored contours (which are of constant doubly averaged Hamiltonian H^*; see Equation (3)) in panels (f)–(h), we took the values of Θ and ϵ_{GR} at the midpoint of the corresponding colored stripe (see Equations (1) and (6) for the definitions of these constants), and we set $\Gamma = 1$ (appropriate for the LK case; see the discussion following Equation (5)). Note also that to not overload the plots we refrain from showing explicit separatrices.

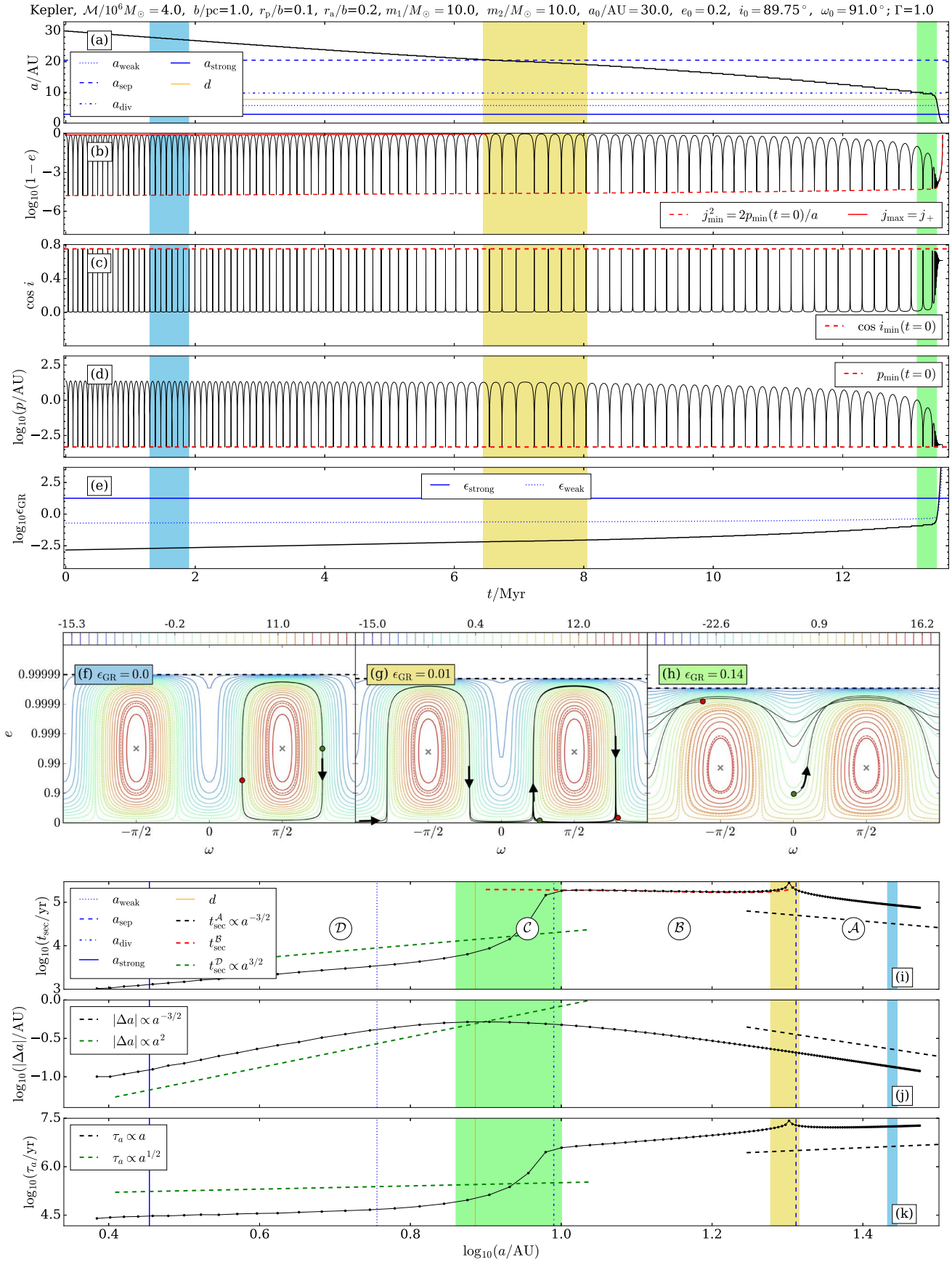

Figure 1. Example 1. A binary ($m_1 = m_2 = 10M_\odot$) orbits an SMBH with mass $\mathcal{M} = 4 \times 10^6 M_\odot$, and undergoes a slow merger. Panels (a)–(e) show the time evolution of a, e, i, p, and $\epsilon_{\rm GR}$, respectively. Panels (f), (g), and (h) show the phase-space evolution during the time intervals shaded in blue, yellow, and green, respectively, in panels (a)–(e). Finally, panels (i), (j), and (k) show with black dots the values of secular period $t_{\rm sec}$, semimajor axis jump $|\Delta a|$ during the eccentricity peak, and characteristic orbital decay time τ_a as functions of a.

discuss our results, including their implications for LK-driven mergers, and we summarize in Section 7.

2. Dynamical Framework

Here we briefly describe our setup in order to establish notation. For more detail see Papers I–III.

Consider a binary with component masses m_1, m_2, orbiting in a fixed, smooth, axisymmetric background potential Φ whose symmetry axis is Z. Let (X,Y) describe the plane perpendicular to Z. Then on long timescales, the orbit of the binary's barycenter (hereafter the "outer" orbit) usually fills an axisymmetric torus. However, if Φ is spherically symmetric then the outer orbit is actually confined to a plane, which we can choose to be the (X,Y) plane. Apart from phase information, the orbit in this plane can be described by its peri/apocenter (r_p, r_a). In this case, on timescales much longer than an outer orbital period the binary will fill an annulus with inner and outer radii (r_p, r_a). In the special case of Keplerian Φ (the LK limit), the outer orbit describes not an annulus but a fixed ellipse with semimajor axis $a_g = (r_a + r_p)/2$ and eccentricity $e_g = (r_a - r_p)/(r_a + r_p)$.

The binary's internal ("inner") Keplerian orbital motion is described by the usual orbital elements: semimajor axis a, eccentricity e, inclination i (relative to the (X,Y) plane), longitude of the ascending node Ω (relative to the X-axis, which is arbitrary but fixed in the cluster frame), argument of pericenter ω, and mean anomaly η. It is also useful to introduce Delaunay actions $L = \sqrt{G(m_1 + m_2)a}$, $J = L\sqrt{1 - e^2}$, and $J_z = J\cos i$, and their conjugate angles η, ω, and Ω, as well as the dimensionless variables

$$\Theta \equiv J_z^2/L^2 = (1 - e^2)\cos^2 i, \tag{1}$$

$$j \equiv J/L = \sqrt{1 - e^2}. \tag{2}$$

Clearly, j must obey $\Theta^{1/2} \leqslant j \leqslant 1$ to be physically meaningful for a given Θ. The minimum/maximum j achieved in a given secular cycle is called $j_{\min\,/\,\max}$.

Ignoring GW emission, the evolution of the inner orbit is dictated by the mutual Newtonian gravitational attraction of the binary components, 1pN GR apsidal precession, and the perturbing tidal influence of the potential Φ. Expanding the tidal force due to the cluster to quadrupole order, and averaging over the inner and outer orbital motion (i.e., performing a weighted integral over the torus, annulus, or ellipse mentioned above) we find the test-particle quadrupole doubly averaged ("DA") equations of motion, given explicitly in Equations (12)–(14) of Paper III. More succinctly, these can be derived from the DA Hamiltonian

$$H = \frac{Aa^2}{8}H^*, \qquad H^* \equiv H_1^* + H_{GR}^*, \tag{3}$$

where A is a constant (with units of s^{-2}) that depends on the potential and outer orbit. In the LK case, $A = G\mathcal{M}/[2a_g^3(1 - e_g^2)^{3/2}]$, where a_g and e_g are, respectively, the semimajor axis and eccentricity of the outer orbit, and $\mathcal{M}$ is the perturber mass. Next,

$$H_1^* = (2 + 3e^2)(1 - 3\Gamma\cos^2 i) - 15\Gamma e^2 \sin^2 i \cos 2\omega, \tag{4}$$

$$H_{GR}^* = -\epsilon_{GR}(1 - e^2)^{-1/2}, \tag{5}$$

are dimensionless Hamiltonians encoding the effects of cluster tides and GR precession, respectively. The quantity Γ is a scalar parameter which, like A, depends on the cluster potential Φ and the choice of outer orbit. It takes values $\in (0,\ 1)$ for binaries in spherically symmetric potentials Φ, and LK theory is recovered in the limit $\Gamma = 1$. Due to a dynamical bifurcation, it turns out that very high eccentricities are much more readily achieved by binaries with $\Gamma > 1/5$ than those with $\Gamma < 1/5$ (Paper II). Finally, the parameter ϵ_{GR} measures the strength of GR precession:

$$\epsilon_{GR} \equiv \frac{24G^2(m_1 + m_2)^2}{c^2 A a^4} \tag{6}$$

$$= 0.258 \times \left(\frac{A^*}{0.5}\right)^{-1}\left(\frac{\mathcal{M}}{10^5 M_\odot}\right)^{-1}\left(\frac{b}{\mathrm{pc}}\right)^3 \times \left(\frac{m_1 + m_2}{M_\odot}\right)^2\left(\frac{a}{20\ \mathrm{au}}\right)^{-4}. \tag{7}$$

In the numerical estimate of Equation (7) we have assumed a spherical cluster of mass $\mathcal{M}$ and scale radius b, and $A^* \equiv A/(G\mathcal{M}/b^3)$—see Paper I.[7] The physical effect of GR precession is typically to quench the cluster tide-driven eccentricity oscillations, as we explored in detail in Paper III, and as has been long established in LK theory (Miller & Hamilton 2002; Fabrycky & Tremaine 2007; Bode & Wegg 2014). As we have shown in Paper III, there are typically no large e oscillations in the "strong GR" regime $\epsilon_{GR} \gtrsim \epsilon_{strong} \equiv 3(1 + 5\Gamma)$. On the other hand, one can ignore GR precession in the "weak GR" regime $\epsilon_{GR} \ll \epsilon_{weak}$ (see Equation (A8)). In the intermediate regime of "moderate GR," $\epsilon_{weak} \lesssim \epsilon_{GR} \ll \epsilon_{strong}$, high-eccentricity excitation is modified but is not prohibited.

In Papers I–III we have gone into great detail about the phase-space dynamics that follows from the Hamiltonian H, shown how phase-space trajectories are split into librating and circulating families, and so on. We will not repeat the arguments here. However, it will be important throughout that we have expressions for various key quantities characterizing secular evolution in terms of Γ and the three dimensionless numbers, Θ, ϵ_{GR}, j_{min}, which are constants of motion when GW emission is ignored. It will also be important to split circulating phase-space trajectories into two asymptotic regimes, characterized by $j_{max} \sim 1$ ("high-j_{max}") and $j_{max} \ll 1$ ("low-j_{max}"), which is a distinction we did not make in previous papers. We gather all of the relevant results in Appendix A, and draw upon them freely hereafter.

3. Secular Dynamics Including Gravitational Wave Emission

GW emission modifies the conservative dynamical picture described in Section 2, allowing compact object binaries to merge. Throughout this paper we concentrate on those binaries whose merger timescale is significantly shortened by secular eccentricity excitation—the so-called "cluster tide-driven mergers" (which includes LK-driven mergers as a special case). Following Randall & Xianyu (2018) we can

[7] In the LK limit, one should set $A^* = 0.5$ and $b = a_g(1 - e_g^2)^{1/2}$.

separate these binaries further into "fast mergers" and "slow mergers." Fast mergers are those that occur after only one (or at most a few) secular eccentricity cycles. Slow mergers occur after many secular eccentricity cycles, and inevitably involve a gradual transition of the binary from the weak-to-moderate GR regime to the strong GR regime (with $e_{\max}$ and other important quantities changing over time)—see Figure 1. In this paper we will focus on understanding the physics of slow mergers.

We will assume throughout this paper that $\Gamma > 0$, and that the binary's maximum eccentricity $e_{\max}$ is achieved at $\omega = \pm\pi/2$. As we have seen in Papers I–III these conditions cover almost all cases of practical interest (at least in spherical clusters), and are always satisfied in the important special case of LK dynamics ($\Gamma = 1$).

3.1. Slow Mergers

Consider a binary with initial eccentricity not close to unity, and suppose that, unless excited to high $e_{\max} \approx 1$, it will not merge within a Hubble time. For the required cluster tide-driven eccentricity excitation to be possible, we know from Paper III that the binary must begin its life in the weak-to-moderate GR regime. Supposing this is the case, and that the binary does indeed achieve high values of e periodically, then during each high-eccentricity episode its semimajor axis a will be decreased by some amount $|\Delta a|$ because of GW emission. For a slow merger, by assumption, each individual decrease is small, $|\Delta a| \ll a$, though of course Δa itself depends on the value of a (see Figure 1(j)). Away from $e \approx 1$, GW emission will be completely negligible, so we can treat a as constant there. Next, the time between these high-eccentricity episodes is $t_{\rm sec}(a)$. Therefore, on timescales longer than a few secular periods we can approximate the slow decay of $a(t)$ as

$$\frac{da}{dt} \approx \frac{\Delta a(a)}{t_{\rm sec}(a)}. \tag{8}$$

Equation (8) is an implicit equation for $a(t)$ for slow-merging binaries. We can use it to define a characteristic orbital decay timescale at a given a:

$$\tau_a(a) \equiv \left| \frac{d\ln a}{dt} \right|^{-1} \approx \left| \frac{a \times t_{\rm sec}(a)}{\Delta a} \right|. \tag{9}$$

This is the quantity we plotted in Figure 1(k).

Eventually, a will become small enough that the binary reaches the strong GR regime and gets "stuck" at high eccentricity. When this happens, Equation (8) breaks down, and we must use a different prescription to follow $a(t)$ accurately all the way to merger.

3.2. Equations of Motion

The main goals of Section 4 will be to understand the behavior of $\Delta a(a)$, $t_{\rm sec}(a)$ and $\tau_a(a)$ during a slow merger, and to appreciate how this behavior is intimately linked with the binary's phase-space trajectory (librating or circulating) and the strength of GR precession (value of $\epsilon_{\rm GR}$). To achieve these goals we must first consider how GW emission affects our equations of motion.

Our DA theory without GW emission consists of Equations (12)–(14) of Paper III, which govern the evolution of ω, j, and Ω, respectively, under the combined effect of secular cluster tides and GR precession. In addition to this, to 2.5th post-Newtonian-order GW emission causes the binary's semimajor axis and eccentricity to evolve according to the Peters (1964) equations[8]:

$$\left(\frac{da}{dt}\right)_{\rm GW} = -\frac{64G^3 m_1 m_2 (m_1+m_2)}{5c^5} \times \frac{1}{a^3(1-e^2)^{7/2}}\left(1 + \frac{73}{24}e^2 + \frac{37}{96}e^4\right), \tag{10}$$

$$\left(\frac{de}{dt}\right)_{\rm GW} = -\frac{304G^3 m_1 m_2 (m_1+m_2)}{15c^5} \times \frac{1}{a^4(1-e^2)^{5/2}}\left(1 + \frac{121}{304}e^2\right). \tag{11}$$

To include GW emission in our theory we therefore add the following terms to our equations of motion for L, j, and $j_z \equiv j\cos i$, respectively:

$$\left(\frac{dL}{dt}\right)_{\rm GW} = \frac{1}{2}\sqrt{\frac{G(m_1+m_2)}{a}}\left(\frac{da}{dt}\right)_{\rm GW}, \tag{12}$$

$$\left(\frac{dj}{dt}\right)_{\rm GW} = -\frac{e}{j}\left(\frac{de}{dt}\right)_{\rm GW}, \tag{13}$$

$$\left(\frac{dj_z}{dt}\right)_{\rm GW} = \frac{j_z}{j}\left(\frac{dj}{dt}\right)_{\rm GW}. \tag{14}$$

We see that GW emission affects both L and j_z (and hence also $\Theta \equiv j_z^2$), which were constants when GW emission was ignored. Note also that at this order, GW emission does not directly affect the argument of pericenter ω or longitude of ascending node Ω (so Equations (12) and (14) of Paper III are unchanged), nor does it affect inclination i.

3.3. Conserved Quantities

When GW emission is switched off, the DA dynamics respects three exact conservation laws. The first is the conservation of a, which results from the "adiabatic" assumption that the binary's inner orbital period is much shorter than the timescale of variation of the weak cluster perturbation (i.e., the outer orbital timescale), allowing us to average over the inner orbit ("single averaging"). The second is the conservation of the dimensionless Hamiltonian H^* (Equation (3)), which follows from the fact that cluster tides are sufficiently weak that we can also average over the outer orbital period and thus treat the perturbation as time-independent ("double averaging"). The third is conservation of the z-component of the binary's angular momentum, or equivalently Θ, which follows from the axisymmetry of the DA time-averaged perturbation as viewed from the binary frame.

Now that we are including GW emission, the binary's binding energy and inner orbital angular momentum will be dissipated according to Equations (10)–(14), and so none of a, H^*, nor Θ will be strictly conserved. On the other hand, it is clear from Equations (10) and (11) that for a fixed semimajor axis a, changes in orbital elements due to GW emission are very strongly concentrated around peak eccentricity $e \to 1$, as we already anticipated. Thus, in the weak-to-moderate GR regime we expect these GW contributions to the equations of

[8] Of course, in principle one can carry the post-Newtonian expansion to higher order, and higher-order terms are important for e.g., LIGO/Virgo templates of inspiralling binary waveforms. However, since this is only important at very late times when compared to the long secular evolution that we are considering here, we always truncate at 2.5pN.

motion to be completely negligible, except in the vicinity of $e \approx 1$. This means that binding energy and angular momentum *can* be treated as roughly conserved away from eccentricity peaks, since in the weak-to-moderate GR regime GW emission is negligible except for short bursts around $e \approx e_{\max}$.

Additionally, as we saw in Figure 1, two new (approximate) conservation laws emerge that are valid on much longer timescales ($\gg t_{\rm sec}$)—these are the conservation of the minimum pericenter distance $p_{\min} \equiv a(1-e_{\max})$, and the conservation of the minimum inclination $i_{\min}$. The fact that these conservation laws hold almost all the way to merger will facilitate our analytical understanding. We now derive each of them in turn.

3.3.1. Conservation of $p_{\min}$

Approximate conservation of $p_{\min}$ follows from the fact that, as $e \to 1$, GW emission dissipates orbital energy of the binary much more efficiently than its angular momentum.[9] As mentioned above, because of the steep dependence on $1-e^2$ in Equations (10) and (11), GW emission is most effective at changing a, e, (and p) only at the eccentricity "peak," when $e \to 1$ (which is the case only during a small fraction $\sim j_{\min} \ll 1$ of each secular period), and can be neglected during the rest of the secular cycle. Thus, during each eccentricity peak GW emission causes changes of the binary orbital elements predominantly over the short time intervals $\delta t \ll T_{\rm b}$ during the periastron passages (where $T_{\rm b} = 2\pi\sqrt{a^3/[G(m_1+m_2)]}$ is the inner orbital period of the binary).

Let the characteristic relative velocity of the binary components at periastron be $v_p \sim \sqrt{G(m_1+m_2)/p}$, which follows from energy conservation and the fact that $p = a(1-e) \ll a$. Representing the effect of the GW radiation reaction as an impulsive force F acting over time δt, we can estimate the change in the binary orbital energy E over each periastron passage to be $\delta E \sim F v_p\, \delta t$. Thus, the characteristic timescale t_E on which E evolves is

$$t_E \sim \frac{E}{\delta E} T_{\rm b} \sim \frac{G(m_1+m_2)}{a F v_p\, \delta t} T_{\rm b}. \tag{15}$$

Similarly, during each periastron passage, GW emission changes the binary angular momentum $J \sim p v_p$ by $\delta J \sim F p\, \delta t$, so that the characteristic time on which J evolves is

$$t_J \sim \frac{J}{\delta J} T_{\rm b} \sim \frac{v_p}{F\ \delta t} T_{\rm b}. \tag{16}$$

The ratio of these two timescales is

$$\frac{t_E}{t_J} \sim \frac{G(m_1+m_2)}{a v_p^2} \sim \frac{p}{a} = 1-e. \tag{17}$$

Thus for $e \approx e_{\max} \to 1$, we have $t_E/t_J \sim 1 - e_{\max} \ll 1$. In other words, when the binary is near peak eccentricity, its energy is dissipated much more rapidly than its angular momentum, so that one can assume that $J \approx J_{\min}$ is almost constant even though E (and a) evolves substantially. Since $J_{\min} \propto [a(1-e_{\max}^2)]^{1/2} \approx (2p_{\min})^{1/2}$ for high $e_{\max}$, this implies that the minimum periastron distance $p_{\min}$ does not change appreciably as a result of GW emission over a single eccentricity peak (Wen 2003, Section 3.1). And since the system undergoes quasiperiodic secular oscillations, the binary returns to the same value of $J_{\min}$ (and hence the same $p_{\min}$) at the e-peak of the following secular cycle.

To be more precise, let us consider the rate of change of the pericenter distance with respect to semimajor axis:

$$p \equiv a(1-e) \quad \Longrightarrow \quad \frac{dp}{da} = 1 - e - a\frac{de}{da}. \tag{18}$$

Dividing Equation (11) by Equation (10) to get de/da, plugging this in to the right-hand side of Equation (18) and expanding near $e = 1$ gives

$$\frac{dp}{da} = (1-e)^2 + \mathcal{O}((1-e)^3). \tag{19}$$

In other words, the rate of change of pericenter distance p vanishes in the limit $e \to 1$, so that $p = p_{\min}$ is constant at the eccentricity peak. Since each individual secular cycle is symmetric (in time, relative to its eccentricity minimum), $p_{\min}$ would then take the same value at the next eccentricity peak, be preserved there, and so on, just as we observed in Figure 1(d). Thus we arrive at the conservation of the minimum pericenter distance:

$$p_{\min} \equiv a(1-e_{\max}) \approx \frac{1}{2} a j_{\min}^2 = \text{const}, \tag{20}$$

which holds true over multiple secular cycles as long as $e_{\max} \to 1$.

Equation (20) implies a simple scaling for $j_{\min}$ in the weak-to-moderate GR regime:

$$j_{\min}(a) \approx \left(\frac{2p_{\min}}{a}\right)^{1/2}, \tag{21}$$

where $p_{\min} = a(t_{\rm i}) \times (1 - e(t_{\rm i}))$ at some reference time $t_{\rm i}$. This is an important result of this paper and will subsequently allow significant analytical simplification.

We note that an argument similar to the one above also applies when the binary gets stuck at high eccentricity in the strong GR regime. In that case there are no more secular oscillations (the binary having decoupled from cluster tides), but energy is still being dissipated efficiently by GW emission while angular momentum is not (for more details, see Wen 2003 and Antognini et al. 2014). As a result, p (rather than just $p_{\min}$) stays approximately constant, so that $j \propto a^{-1/2}$. We will use this scaling when considering the strong GR regime in Section 4.3.

3.3.2. Conservation of $i_{\min}$

The conservation of $i_{\min}$—the minimum value of the binary inclination that is reached at the eccentricity peak—follows from the fact that the GW emission does not affect the orientation of the orbital plane of the binary and thus does not affect its inclination. Again, because of the time-symmetry of each secular cycle, at the next eccentricity peak the binary will arrive at the same value of $i_{\min}$ (which, again, will not be changed by the GW emission, regardless of the decay in a), and so on. As a result, it follows that over time intervals much longer than each secular cycle

$$i_{\min} = \text{const}. \tag{22}$$

This is precisely what we saw in Figure 1(c).

[9] The following argument is not unique to the GW emission and can be generalized for any short-range dissipative force, e.g., due to the fluid tides acting inside the binary components.

For the remainder of this paper we will take $p_{\rm min}$ and $i_{\rm min}$ as our two primary, a-independent constants of motion, which persist over multiple secular cycles in the weak-to-moderate GR regime. We can then rewrite the key secular evolution parameter Θ in terms only of a and these conserved quantities as

$$\Theta(a) \approx \frac{2p_{\rm min}}{a}\cos^2 i_{\rm min}, \tag{23}$$

which will greatly simplify our analytical understanding. We will discuss the circumstances in which the conservation of $p_{\rm min}$ and $i_{\rm min}$ breaks down, invalidating Equation (23), in our detailed discussion of a numerical example in Section 5.2.

4. Evolution of a Shrinking Binary through Time and Phase Space

The binary in Figure 1 started its life in the weak GR regime (see panel (e)). Then as its semimajor axis shrank, it entered the moderate GR regime and finally ended up in the strong GR regime before merging. On a related note, we also saw that the binary's phase-space trajectory evolved from librating (panel (f)), to circulating with $e_{\rm min} \ll 1$ (panel (g); we call this a "high-$j_{\rm max}$" circulating trajectory), to circulating with $e_{\rm min} \approx 1$ (panel (h); we call this a "low-$j_{\rm max}$" circulating trajectory). This pattern of behavior is rather general, and in fact has been discussed briefly in the case of LK-driven mergers by Blaes et al. (2002) and Antonini et al. (2016).

Physically, this evolution of phase-space trajectory follows from the way GR precession modifies the phase-space structure, by encouraging rapid pericenter precession at high e and hence expanding the region of circulating trajectories at the expense of librating trajectories. Hence, one can roughly think of increasing $\epsilon_{\rm GR}$ as pushing the separatrix "down" to lower eccentricity in the (ω, e) plane. Since the increases in $\epsilon_{\rm GR}$ occur only when the binary is at $e \approx e_{\rm max}$ (because this is where GW emission, and hence the decay of a, is concentrated), these downward shifts of the separatrix coincide with the binary's highest eccentricity. In this way a binary on a librating trajectory inevitably moves "toward" the separatrix from below (or rather the separatrix moves closer to it from above) and ultimately ends up crossing the separatrix onto a circulating trajectory. Eventually the binary gets confined to a low-$j_{\rm max}$ circulating trajectory in the strong GR regime, where asymptotically there are no eccentricity oscillations at all.

In this section we wish to understand more quantitatively how the binary's phase-space trajectory and GR regime evolve as a function of semimajor axis a. The details of the various transitions between these regimes are quite technical, so we relegate most of our discussion to Appendix B and retain here only the salient points.

4.1. Characteristic Length Scales

As shown in Appendix B, the orbital evolution of a binary en route to merger admits a key length scale

$$d \equiv \left(\frac{4G^2(m_1+m_2)^2}{5c^2A\Gamma(2p_{\rm min})^{1/2}}\right)^{2/7} \tag{24}$$

$$\approx 7.1\ {\rm au} \times \Gamma^{-2/7}\left(\frac{A^*}{0.5}\right)^{-2/7}\left(\frac{\mathcal{M}}{10^6 M_\odot}\right)^{-2/7} \times \left(\frac{b}{\rm pc}\right)^{6/7}\left(\frac{m_1+m_2}{M_\odot}\right)^{4/7}\left(\frac{p_{\rm min}}{10^{-2}{\rm au}}\right)^{-1/7}, \tag{25}$$

which is independent of a. There are then four critical semimajor axis values to contend with.

A binary that starts its life in the weak GR regime will inevitably move into the moderate GR regime at some point as its semimajor axis shrinks. Thus the first critical value is $a_{\rm weak}$, which we define to be the semimajor axis corresponding to $\epsilon_{\rm GR} = \epsilon_{\rm weak}$, i.e., to the transition between the weak and moderate GR regimes. Using Equation (B4) and the fact that $\epsilon_{\rm weak} \ll \epsilon_{\rm GR}$, this is

$$a_{\rm weak} \equiv \left(\frac{\sqrt{2}-1}{2\cos^2 i_{\rm min}}\right)^{2/7} d \approx 0.63(\cos i_{\rm min})^{-4/7} d. \tag{26}$$

Note that $a_{\rm weak} \sim d$, unless $\cos i_{\rm min} \ll 1$.

Next we define $a_{\rm strong}$, which demarcates the inevitable transition between moderate and strong GR regimes, i.e., it corresponds to $\epsilon_{\rm GR} = \epsilon_{\rm strong} \equiv 3(1+5\Gamma)$. Using Equation (6) we get

$$a_{\rm strong} \equiv \left(\frac{8G^2(m_1+m_2)^2}{c^2A(1+5\Gamma)}\right)^{1/4} \tag{27}$$

$$\approx 6.1{\rm au} \times \left(\frac{1+5\Gamma}{6}\right)^{-1/4}\left(\frac{A^*}{0.5}\right)^{-1/4} \left(\frac{\mathcal{M}}{10^6 M_\odot}\right)^{-1/4}\left(\frac{b}{\rm pc}\right)^{3/4}\left(\frac{m_1+m_2}{M_\odot}\right)^{1/2}. \tag{28}$$

After entering the strong GR regime, the binary gets stuck at high eccentricity, and its semimajor axis decays while $p = a(1-e)$ remains roughly constant. See Section 4.3.3 for more details.

A binary initially on a librating phase-space trajectory will transition to a circulating trajectory once a drops below some threshold value, as we saw in Figure 1. This threshold value is $a_{\rm sep}$, which is given by[10]

$$a_{\rm sep} \equiv \left(\frac{1+5\Gamma}{10\Gamma} - \cos^2 i_{\rm min}\right)^{-2/7} d. \tag{29}$$

Typically $a_{\rm sep} \gtrsim d$. Of course, $a_{\rm sep}$ has physical meaning only if $\cos^2 i_{\rm min} < (1+5\Gamma)/10\Gamma$. This is because for $\cos^2 i_{\rm min} > (1+5\Gamma)/10\Gamma$, the binary is already on a circulating trajectory even for $a \gg d$, so it never crosses a separatrix on its way to $a \to 0$.

There is one further critical semimajor axis value, which we call

$$a_{\rm div} \equiv (\sin i_{\rm min})^{-4/7} d. \tag{30}$$

At $a = a_{\rm div}$ the dimensionless numbers σ and κ, which play a role in setting time spent at highest eccentricity (see Paper III), diverge. This divergence will become important when we discuss the evolution of the secular timescale (Section 4.2). Moreover, if $a_{\rm sep} > a_{\rm div}$, then $a = a_{\rm div}$ corresponds approximately to the

[10] To derive this, we set $j_0^2 = j_+^2 = 1$ in Equations (B1) and (B3)—see Papers II–III.

transition between high-$j_{\max}$ ($a_{\rm div} < a < a_{\rm sep}$) and low-$j_{\max}$ ($a < a_{\rm div}$) circulating trajectories—see Appendix B for details. For future reference we write down the ratio:

$$\frac{a_{\rm sep}}{a_{\rm div}} = \left(1 + \frac{5\Gamma - 1}{1 + 5\Gamma - 10\Gamma\cos^2 i_{\min}}\right)^{2/7}. \tag{31}$$

To summarize, we have defined four critical semimajor axis values $a_{\rm weak}$, $a_{\rm strong}$, $a_{\rm sep}$, $a_{\rm div}$. The weak GR regime corresponds to $a > a_{\rm weak}$, while the moderate GR regime corresponds to $a_{\rm strong} < a < a_{\rm weak}$. We emphasize that we have purposely written e.g., $a > a_{\rm weak}$ rather than $a \gg a_{\rm weak}$ here: it turns out that different dynamical regimes are not very well separated in semimajor axes (in fact we usually have $a_{\rm weak} \sim a_{\rm sep} \sim a_{\rm div} \sim d$), despite being well separated in $\epsilon_{\rm GR}$—see, e.g., Figure 1. In Appendix B we show in more detail how the binary passes through these different regimes as a shrinks for several different values of $\Gamma > 0$. The results are quite complex: transitions between different regimes do not always happen in the same order, and the $\Gamma > 1/5$ case has to be considered separately from $0 < \Gamma \leqslant 1/5$. Nevertheless, the asymptotic regimes defined here will allow us to make analytical progress, and will give us a qualitative understanding of the behavior of $t_{\rm sec}$, Δa, and τ_a throughout slow mergers, which is what we turn to next.

4.2. The Secular Timescale

One crucial quantity in any study of LK-driven or cluster tide-driven mergers is the period of secular eccentricity oscillations, $t_{\rm sec}$, since this gives the time elapsed between each episode of GW emission. In Figure 1(i) we plotted $t_{\rm sec}$ as a function of semimajor axis a for a binary undergoing LK oscillations as it shrank and ultimately merged (time runs from right to left in that panel). We labeled four regimes of the $t_{\rm sec}(a)$ curve, $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$, $\mathcal{D}$, and each regime exhibited a different a-dependence. In this section we will explain the behavior in each of these regimes in turn.

In DA theory without GW emission, the definition of $t_{\rm sec}$ is

$$t_{\rm sec} = 2\int_{j_{\min}}^{j_{\max}} \left(\frac{dj}{dt}\right)^{-1} dj, \tag{32}$$

with dj/dt given in Equation (A1). The right-hand side of Equation (32) of course depends on a, which decays throughout a slow merger.

4.2.1. Regime $\mathcal{A}$: Librating Trajectories, $a > a_{sep}$

Regime $\mathcal{A}$ in Figure 1(i) corresponds approximately to $a > a_{\rm sep}$, i.e., to librating phase-space trajectories (see Figure 1(f) and Section 4.1). Binaries on librating phase-space trajectories spend most of their time far away from $e \approx 1$ (see Figure 1(b)). In the weak-to-moderate GR regime, this means that the explicitly $\epsilon_{\rm GR}$-dependent terms in Equation (A1) are unimportant for most of the evolution, and so in this regime a good approximation to the secular period is found by evaluating Equation (32) ignoring the explicitly $\epsilon_{\rm GR}$-dependent terms. Technically speaking, such an approximation becomes exact in the asymptotic limit $\epsilon_{\rm GR} \ll \epsilon_{\rm weak}$ and $\sigma \to 0$ (Equation (B5))—see Paper III.

Moreover, for librating trajectories, we know that $j_{\max} \approx j_+ \sim 1$ (Equation (A12)). Defining $\Delta \equiv \max[j_+^2, j_-^2, j_0^2] - \min[j_+^2, j_-^2, j_0^2]$, and assuming $j_{\min}^2 \ll j_{\max}^2$, we find (see Equations (32)–(34) of Paper II):

$$t_{\rm sec}^{\mathcal{A}} \approx \frac{8}{3A}\sqrt{\frac{G(m_1+m_2)}{|25\Gamma^2 - 1|}} \times \frac{1}{a^{3/2}\sqrt{\Delta}} K\left(\frac{j_+}{\sqrt{\Delta}}\right), \tag{33}$$

where $K(x) \equiv \int_0^{\pi/2} d\alpha/\sqrt{1 - x^2\sin^2\alpha}$ is plotted in Figure 2. We emphasize that GR is still implicitly present in Equation (33) because it affects the values of j_+ and Δ that must be plugged into the right-hand side.

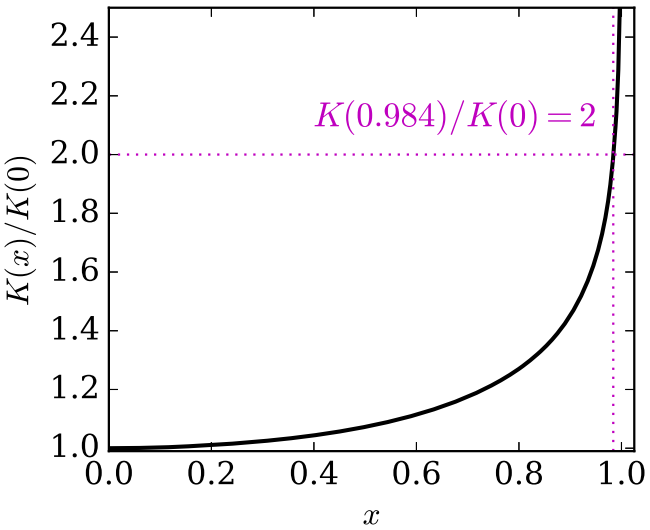

Figure 2. Value of the elliptic integral $K(x)$ normalized by $K(0) \equiv \pi/2$ as a function of x. The scaling is weak except for $x \gtrsim 0.99$.

We cannot simplify Equation (33) further without specifying the ordering of j_0^2, $j_\pm^2$, which itself depends on Γ. However, far from the separatrix between librating and circulating trajectories (i.e., a sufficiently greater than $a_{\rm sep}$), we expect j_+, Δ to depend only weakly on a (Figures 10–11), so that

$$t_{\rm sec}^{\mathcal{A}} \propto a^{-3/2}. \tag{34}$$

This scaling is confirmed in regime $\mathcal{A}$ of Figure 1(i) with a black dashed line. Physically it arises because the binary's inner orbital period is proportional to $a^{3/2}$. It has been noted by many authors when estimating an LK-driven merger timescale (e.g., Wen 2003; Thompson 2011; Randall & Xianyu 2018), although they did not tie it to the librating nature of the phase-space trajectory.

When the binary's phase-space trajectory gets close to the separatrix $a \to a_{\rm sep}$, we know (Appendix A) that $j_+^2, j_0^2 \to 1$ and so $j_{\max}/\sqrt{\Delta} \to 1$. Figure 2 shows that $K(x)$ diverges for $x \to 1$, so $t_{\rm sec}$ should peak around this point. This again matches what we see[11] in regime $\mathcal{A}$ of Figure 1(i).

4.2.2. Regime $\mathcal{B}$: Circulating Trajectories, $a_{\rm div} < a < a_{\rm sep}$

After crossing the separatrix to $a < a_{\rm sep}$, the binary ends up on a circulating phase-space trajectory. One can still use the results for $t_{\rm sec}$ obtained in Paper II (Equations (32)–(34)), but for circulating trajectories the scaling of Equation (34) we found in regime $\mathcal{A}$ no longer holds. The behavior of $t_{\rm sec}(a)$

[11] We should note that the peak in $t_{\rm sec}$ is not centered precisely on $a_{\rm sep}$. This is because the Equation (29) is only approximate, derived in a particular high-eccentricity limit and assuming exact conservation of $p_{\min}$, $i_{\min}$.

becomes significantly more complicated, and $t_{\rm sec}$ does not necessarily increase with decreasing a, as we will see.

The argument we used in deriving Equation (33) relied on $e_{\rm min}$ being much smaller than unity, and the approximation $\sigma \ll 1$. If these approximations are good (which in particular now requires high-$j_{\rm max}$ circulation, i.e., $a_{\rm div} < a < a_{\rm sep}$; see Figure 10), then we can say that we are in regime $\mathcal{B}$. This time we have $j_{\rm max} = j_0$ (Equation (A14)) and $j_+^2 > 1$ (Figures 10–11), so that $\Delta \approx j_+$ and instead of Equation (33), Equations (32)–(34) of Paper II give

$$t_{\rm sec}^{\mathcal{B}} \approx \frac{8}{3A}\sqrt{\frac{G(m_1+m_2)}{|25\Gamma^2-1|}} \times \frac{1}{a^{3/2} j_+} K\left(\frac{j_0}{j_+}\right). \tag{35}$$

Plugging Equations (B1) and (B3) into Equation (35) gives us an explicit expression for $t_{\rm sec}^{\mathcal{B}}(a)$, which we plot with a red dashed line in Figure 1(i). We see that $t_{\rm sec}^{\mathcal{B}}(a)$ diverges as we take $a \to a_{\rm sep}$ from below, as expected. But away from the divergence we expect the elliptic integral K to scale weakly with a, and we get (using Equation (B1))

$$t_{\rm sec}^{\mathcal{B}} \propto \frac{(a/d)^{1/4}}{\sqrt{1+(a/d)^{7/2}\cos^2 i_{\rm min}}}. \tag{36}$$

Thus, $t_{\rm sec}^{\mathcal{B}}$ can either increase or decrease with a, depending on the value of $(a/d)^{7/2}\cos^2 i_{\rm min}$.

4.2.3. Regime $\mathcal{C}$: Circulating Trajectories, $a \lesssim a_{\rm div}$

We see from Figure 1(i) that Equation (35) becomes inaccurate once $a \lesssim a_{\rm div}$. This is expected because $\sigma \to \infty$ as $a \to a_{\rm div}$ (Figure 10), implying that the non-GR approximation for $t_{\rm sec}$ derived in Paper II (which we used for regimes $\mathcal{A}$, $\mathcal{B}$) becomes invalid. Instead, we now need to use Equations (32) and (A1) to evaluate $t_{\rm sec}$.

Figure 1(i) shows that $t_{\rm sec}$ very rapidly diminishes as a decreases below $a_{\rm div}$, which coincides with the rapid increase of $e_{\rm min}$ or, equivalently, rapid decrease of $j_{\rm max}$, which is now in the low-$j_{\rm max}$ limit. This is because for high-$j_{\rm max}$ circulating trajectories, a smaller a leads to a smaller $j_{\rm max}$, which in turn means that the binary spends more time at "high" eccentricities (say with e above 0.9). The cluster tide-driven secular evolution is faster at high e than at $e \sim 0$ because, even though the torque on a binary with, say, $e = 0.1$ is comparable to that on a binary with $e = 0.9$, the angular momentum of the latter is significantly smaller, so the relative change in angular momentum occurs over a much shorter timescale. For that reason $t_{\rm sec}$ decreases with decreasing a, very rapidly (in a non-power-law fashion) for $a \lesssim a_{\rm div}$. We call this interval of rapid $t_{\rm sec}$ decay regime $\mathcal{C}$.

In Appendix C we use Equations (32) and (A1) to derive an approximation of Equation (C3) for $t_{\rm sec}$ valid in regime $\mathcal{C}$. Depending on the relationship between $a_{\rm div}$ and $a_{\rm weak}$, there are two possibilities. For $a_{\rm weak} \ll a \lesssim a_{\rm div}$ (as in Figure 1), Equation (C3) gives

$$t_{\rm sec}^{\mathcal{C}} \approx \frac{2\pi}{15\Gamma A}\frac{\sqrt{2G(m_1+m_2)p_{\rm min}}}{d^2}\cos i_{\rm min} \times \left(\frac{a}{d}\right)^{13/4}\left|1-\left(\frac{a}{a_{\rm div}}\right)^{7/2}\right|^{-3/2}. \tag{37}$$

On the other hand, for $a \lesssim a_{\rm div} \ll a_{\rm weak}$ (as in Figure 5) we find

$$t_{\rm sec}^{\mathcal{C}} \approx \frac{2\pi}{15\Gamma A}\frac{\sqrt{2G(m_1+m_2)p_{\rm min}}}{d^2} \times \left(\frac{a}{d}\right)^{3/2}\frac{2-(a/a_{\rm div})^{7/2}}{|1-(a/a_{\rm div})^{7/2}|^{3/2}}. \tag{38}$$

Both of these expressions show a rapid decay of $t_{\rm sec}$ as a drops even slightly below $a_{\rm div}$ because of the a-dependent term in the denominator. However, this term rapidly becomes constant as a decreases further, switching again to a power-law behavior of $t_{\rm sec}(a)$, which we cover next.

4.2.4. Regime $\mathcal{D}$: Moderate GR, $a \ll a_{\rm div}, a_{\rm weak}$

As a becomes substantially smaller than both $a_{\rm div}$ and $a_{\rm weak}$, and the binary is in the moderate-GR regime, GR precession plays an even more important role in determining $t_{\rm sec}$. We call this situation regime $\mathcal{D}$. In this case we can still use Equation (C3) to find in the appropriate limit that

$$t_{\rm sec}^{\mathcal{D}} \approx \frac{4\pi}{15\Gamma A}\frac{\sqrt{2G(m_1+m_2)p_{\rm min}}}{d^2}\left(\frac{a}{d}\right)^{3/2}. \tag{39}$$

The predicted scaling $t_{\rm sec}^{\mathcal{D}} \propto a^{3/2}$ matches what we observed at the low-a end of Figure 1(i). Note that this expression is valid regardless of the relationship between $a_{\rm div}$ and $a_{\rm weak}$. Again the secular timescale decreases (rather than increasing like one would naively expect) as the semimajor axis shrinks, although not as rapidly as in regime $\mathcal{C}$.

4.3. The Evolution of Semimajor Axis

In this section we aim to understand how the semimajor axis of a binary decays with time in certain asymptotic regimes. To achieve this we first write down expressions for the decay in semimajor axis over one secular cycle, Δa, in terms of a. This allows us to understand the behavior we saw in panel (j) of Figure 1. Then we plug our expressions for $t_{\rm sec}(a)$ and $\Delta a(a)$ into the right-hand side of Equation (8) to calculate (very approximate) expressions for τ_a.

To begin, we integrate Equation (10) over one secular cycle, approximating a as constant to lowest order (which is valid since $|\Delta a| \ll a$ by assumption for a slow merger). The result is[12]

$$\Delta a \approx -\frac{\lambda_0}{a^3}\int_{\rm sec.\ cycle}\frac{dt}{(1-e^2)^{7/2}}\left(1+\frac{73}{24}e^2+\frac{37}{96}e^4\right), \tag{40}$$

where $\lambda_0 \equiv (64/5)G^3c^{-5}m_1m_2(m_1+m_2)$ is independent of a. Assuming the binary reaches very high maximum eccentricity $e_{\rm max} \to 1$, we show in Appendix D that we can ultimately approximate this as (Equation (D4)):

$$\Delta a \approx -\lambda_2 \times \frac{\xi}{a^{3/2}|j_+ j_0|}, \tag{41}$$

[12] Note that Equation (40) is essentially identical to the first line in Equation (55) of Randall & Xianyu (2018)—see Appendix E.

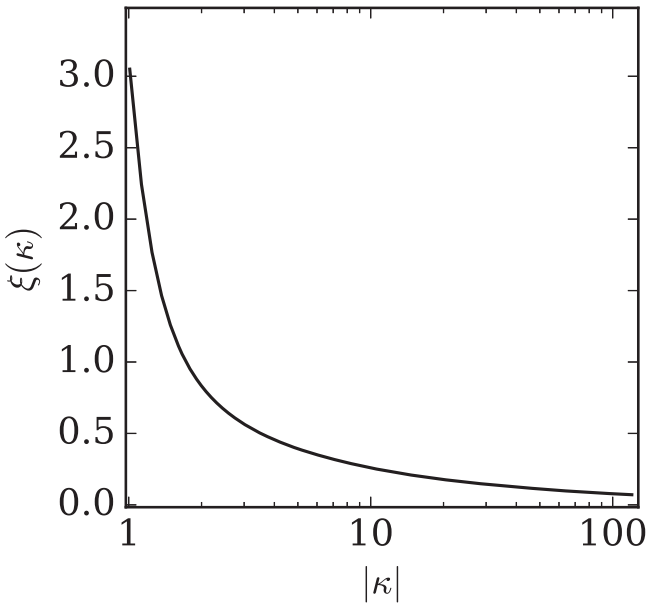

Figure 3. Plot of the function ξ as a function of $|\kappa|$ according to Equation (49), which is valid for low-$j_{\max}$ circulating trajectories in the moderate GR regime. Note that the horizontal axis is logarithmic.

where $\lambda_2 \equiv 1360G^{7/2}m_1m_2(m_1+m_2)^{3/2}/[9c^5A(2p_{\min})^3\sqrt{|25\Gamma^2-1|}]$ is independent of a, and ξ is a complicated function of a and other parameters—see Equation (D5).

To simplify the expressions for ξ, j_+, and j_0 in Equation (41) we need to specify the strength of GR precession. The asymptotic regimes of interest for evaluating Δa are therefore the weak, moderate, and strong GR regimes. Unfortunately these do not map precisely onto the phase-space regimes $\mathcal{A}$-$\mathcal{D}$ that we used to understand $t_{\rm sec}(a)$ behavior in Section 4.2. Nevertheless, there are situations where a binary sits in, for instance, both the weak GR regime and phase-space regime $\mathcal{A}$, and in such situations clean analytic results for τ_a are possible, as we will see.

4.3.1. Weak GR

In Appendix D we show that for weak GR ($a \lesssim a_{\rm weak}$),

$$\xi \approx \frac{8}{15}\frac{1}{\sqrt{1+|\sigma|}}. \tag{42}$$

From Figures 10 and 11 (or by inspection of Equation (B5)), we know that for weak GR we normally have $\sigma \lesssim 1$; thus, ξ will also be $\mathcal{O}(1)$ and scale weakly with a in the weak GR regime. Since j_+ and j_0 also scale weakly with a in this regime (see Equations (B1), (B3)), we find from Equation (41) that

$$\Delta a \propto a^{-3/2}. \tag{43}$$

This matches what we saw at the high-a end of Figure 1(j). Physically, the scaling of Equation (43) just reflects the fact that the time spent in the high-eccentricity state is proportional to the secular timescale, and that in the weak GR regime $t_{\rm sec} \propto a^{-3/2}$ (Equation (34)). Indeed, one might have guessed the result of Equation (43) a priori by noting from Equation (10) that for $j \approx j_{\min}$ we have $(da/dt)_{\rm GW} \propto a^{-3}j_{\min}^{-7}$. Since the time spent at high eccentricity is $t_{\min} \sim j_{\min}\,t_{\rm sec}$ (Paper III), we get $\Delta a \sim (da/dt)_{\rm GW} \times t_{\min} \propto a^{-3}j_{\min}^{-6}\,t_{\rm sec}$. Plugging in Equation (21) for $j_{\min}$, we simply get $\Delta a \propto t_{\rm sec} \propto a^{-3/2}$.

We can now evaluate the characteristic decay timescale τ_a if we assume not only that the binary is in the weak GR regime but also that it is on a librating phase-space trajectory (i.e., that we are in regime $\mathcal{A}$). In this case we can plug Equations (33), (41), and (42) into Equation (9) to find

$$\tau_a = \Lambda_{\rm weak}U^{-1}a, \tag{44}$$

where

$$\Lambda_{\rm weak} \equiv \sqrt{1+\sigma}\,\frac{|j_+j_0|}{\sqrt{\Delta}}K\left(\frac{j_+}{\sqrt{\Delta}}\right), \tag{45}$$

is dimensionless and typically $\mathcal{O}(1)$ away from separatrices, and we defined a "decay rate"

$$U \equiv \frac{272G^3m_1m_2(m_1+m_2)}{3(2p_{\min})^3c^5} \tag{46}$$

$$=1.4\ {\rm au\ Gyr^{-1}} \times \left(\frac{p_{\min}}{10^{-2}{\rm au}}\right)^{-3}\left(\frac{m}{10M_\odot}\right)^3. \tag{47}$$

(In the numerical estimate of Equation (47), we put $m_1 = m_2 = m$). The scaling $\tau_a \propto a$ predicted by Equation (44) is exactly what we see at the high-a end in Figure 1(k).

4.3.2. Moderate GR

In the moderate GR regime we can get a scaling for Δa if we assume the $(d/a)^{7/2}$ terms dominate in Equations (B1) and (B3). Then from Equation (41):

$$\Delta a \approx -\frac{\lambda_2\xi}{d^{7/2}}\frac{\sqrt{|25\Gamma^2-1|}}{10\Gamma} \times a^2. \tag{48}$$

We can simplify the expression for ξ (Equation (D5)) if we further assume the binary is on a low-$j_{\max}$ circulating trajectory (which is inevitably true at some point before the strong GR regime is reached). In this case $j_{\max} = -\sigma j_-$ (Equation (A15)) so that $x_{\max} = x_\sigma \approx (-\sigma j_-)/(\gamma j_-) = -\sigma/\gamma \equiv -\kappa$. For moderate GR it is also easy to show that $|x_\alpha| \approx \gamma^{-2} \ll 1$, so we can ignore $|x_\alpha|$ compared to x in Equation (D5). As a result we find

$$\xi(\kappa) \approx \int_1^{|\kappa|}\frac{x^{-6}dx}{\sqrt{(x-1)(|\kappa|-x)}}. \tag{49}$$

In Figure 3 we plot ξ as a function of $|\kappa|$ according to Equation (49); in particular, we see that $\xi \sim 1$ except for very large $|\kappa| \gg 1$.

Also, away from $a \approx a_{\rm div}$ (i.e., in regime $\mathcal{D}$) we know κ is almost never large compared to unity, and scales weakly with a (Figures 10–11), so we can treat ξ as an order-unity constant for a rough analysis. As a result, Equation (48) predicts a scaling

$$\Delta a \propto a^2. \tag{50}$$

We see from Equation (50) that, unlike for weak GR (Equation (43)), here the individual decrements in semimajor axis Δa get *smaller* as the semimajor axis a shrinks. The scaling of Equation (50) matches what we saw in Figure 1(j) toward the low-a end.

One can understand the result of Equation (50) qualitatively as follows. Like for weak GR, at very high e, we again have

roughly $(da/dt)_{\rm GW} \propto a^{-3} j_{\rm min}^{-7}$, and again using Equation (21) this is $\propto a^{1/2}$. This time, since the binary spends a large fraction of its secular period in the vicinity of $j_{\rm min}$, we get a rough estimate of Δa by multiplying $(da/dt)_{\rm GW}$ not by $t_{\rm min}$, but by $t_{\rm sec}$. Using the scaling of Equation (39) we get $\Delta a \propto a^{1/2} \times a^{3/2} \propto a^2$. Loosely speaking, the factor of ξ in Equation (48) accounts for the fraction of time that the binary actually spends in the vicinity of $j_{\rm min}$ during each secular cycle.

We can also compute the decay timescale τ_a in the moderate GR, low-$j_{\rm max}$ circulating regime. To do so we plug Equations (39) and (41) into Equation (9) to find[13]

$$\tau_a = \Lambda_{\rm mod}\, U^{-1}(2p_{\rm min})^{1/2}\, a^{1/2}, \tag{51}$$

where

$$\Lambda_{\rm mod} \equiv \frac{4\pi(1-\kappa)}{5}\left[\int_1^{|\kappa|} \frac{x^{-6}dx}{\sqrt{(x-1)(|\kappa|-x)}}\right]^{-1}, \tag{52}$$

which is typically $\mathcal{O}(1)$. The scaling $\tau_a \propto a^{1/2}$ is confirmed at the low-a end of Figure 1(k). The moderate GR decay time of Equation (51) is shorter than the analogous weak GR result decay time of Equation (44) by a factor $\sim\sqrt{2p_{\rm min}/a} \sim j_{\rm min} \ll 1$.

4.3.3. Strong GR

Once the strong GR regime is reached, Equation (8) ceases to be valid because the binary decouples from cluster tides and so no longer undergoes secular eccentricity oscillations (Paper III). The evolution of a, e is then dictated purely by Equations (10) and (11). Supposing the transition to the strong GR regime happens at some reference time $t_{\rm i}$, we know from Section 3.3 that for $t > t_{\rm i}$ the binary conserves its value of $p = a(1-e) = p_{\rm min}$, meaning $j = (2p_{\rm min})^{1/2} a^{-1/2}$ (although $p_{\rm min}$ can be a factor of $\sim$2 larger than its initial value $p_{\rm min}(t=0)$—see Figure 7(d) and the final paragraph of Section 5.2.). With this we can eliminate eccentricity from Equation (10), resulting in

$$\frac{da}{dt} \approx -\frac{\lambda_{\rm l}}{(2p_{\rm min})^{7/2}} a^{1/2}. \tag{53}$$

We now use the definition $\tau_a \equiv |d\ln a/dt|^{-1}$ (Equation (9)) to calculate τ_a directly from Equation (53) as

$$\tau_a = \frac{8}{5} U^{-1}(2p_{\rm min})^{1/2}\, a^{1/2}. \tag{54}$$

Thus, the only difference between the characteristic decay timescale in the moderate GR regime (Equation (51)) and that during the strong GR regime (Equation (54)) is a factor of order unity which depends very weakly on a. This explains why the behavior of $\tau_a \propto a^{1/2}$ is barely modified in Figure 1(k) once the binary enters the strong GR regime ($a \lesssim 10^{0.45}$ au).

5. Numerical Examples

In this section we will provide further numerical examples akin to Figure 1, in particular for binaries moving in non-Keplerian potentials. Our aim is to verify and elucidate the approximate analytical results derived in Section 3–4. We do this by direct numerical integration of the DA equations of motion, including both GR precession and GW emission, for various binaries that undergo slow mergers. First we give two Examples with $\Gamma > 1/5$ (Section 5.1–5.2), the first of which exhibits all of the hallmark behavior of a slow merger beginning in the weak GR regime, and the second of which allows us to focus on the late-stage (moderate and strong GR) evolution. We then provide one further example, this time for a binary with $0 < \Gamma \leqslant 1/5$ (Section 5.3). Note that we also provide one additional numerical example in the LK limit in Appendix E, when comparing our work with that of Randall & Xianyu (2018).

To be clear, we note that we ran many more numerical experiments of slow mergers than those shown here. We have chosen to present here the minimal number of examples that still capture qualitatively all of the possible interesting evolutionary scenarios. (There are of course uninteresting cases, such as binaries that are so tightly bound the cluster essentially plays no role in their evolution, but we do not include them here.)

5.1. Example 2: $\Gamma = 0.42 > 1/5$. An Initially Librating Trajectory in the Weak GR Regime

In Figure 4 we show the result of integrating the DA equations of motion for a binary that orbits the spherical Hernquist potential $\Phi(r) = -G\mathcal{M}/(b+r)$, where $\mathcal{M} = 10^6 M_\odot$ and $b = 1$ pc. The outer orbital peri/apocenter is chosen to be $(r_{\rm p}/b,\ r_{\rm a}/b) = (1.5, 1.7)$. The resulting Γ value is $0.42 > 1/5$. The figure is set up in precisely the same way as Figure 1.

Let us first focus on the initial $\sim$3000 Myr. From panel (a) we see that at $t = 0$ the binary has $a_0 = 250$ au, and that for the first $\sim$3000 Myr, $a(t)$ exceeds significantly each of the four critical values $a_{\rm weak}$, $a_{\rm sep}$, $a_{\rm div}$, and $a_{\rm strong}$, which were defined in Section 4.1 and which we show with horizontal blue lines (see the legend). It follows that during this time, the binary resides in the weak GR regime: and indeed, we see from panel (e) that $\epsilon_{\rm GR}$ is initially far smaller than $\epsilon_{\rm weak}$ (blue horizontal dotted line). Moreover, panel (b) shows that the binary undergoes the expected secular eccentricity oscillations (initially on a timescale of $\sim$30Myr), and reaches a very high maximum eccentricity of $1 - e_{\rm max} \approx 10^{-5}$. Concomitantly there are secular oscillations in inclination i (panel (c)) and pericenter distance p (panel (d)), though as predicted in Section 3.3 the values of $\cos i_{\rm min}$ and $p_{\rm min}$ reached at the peak of each secular eccentricity cycle are very nearly conserved (see the dashed red horizontal lines in these panels). Similarly, from panel (b) we see that the maximum eccentricity of the binary is well described by $j_{\rm min}^2 = 2p_{\rm min}/a$ (Equation (21)), while its minimum is well described by $j_{\rm max} = j_+$. The latter fact implies that the binary is initially on a librating phase-space trajectory (Equation (A12)), and this is confirmed by panel (f), in which we show the phase-space evolution during the time interval denoted by the blue shaded stripe.

So, we have a binary on a librating trajectory (regime $\mathcal{A}$) in the weak GR regime, whose semimajor axis is slowly decaying with time. From panels (i), (j), and (k) we see that the binary obeys all of the expected scalings for $t_{\rm sec}$, Δa, and τ_a in this regime, namely Equations (33), (43), and (44), respectively. As we know from Section 4 there are two key things that happen next to such a binary: one is that it enters the moderate GR

[13] Note that we do *not* take Δa from Equation (48): instead we used the more general Equation (41), which allows us to take advantage of the cancellation of the factors $|j_+ j_0|$ *without* having to assume the dominance of the $(d/a)^{7/2}$ terms in j_+^2, j_0^2. Similarly, ξ as given in Equation (49) does not rely on this assumption.

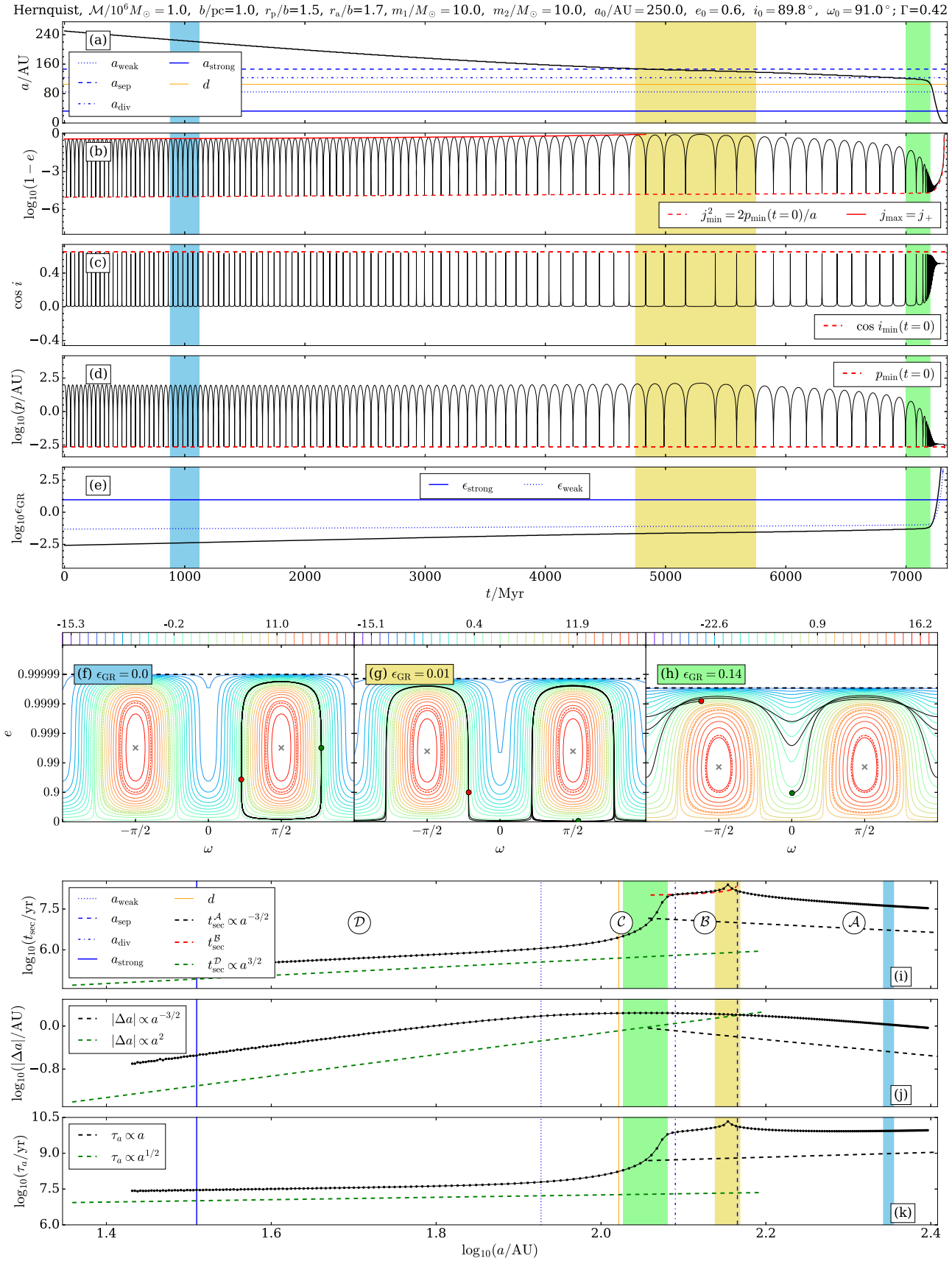

Figure 4. Example 2. The binary orbits a Hernquist cluster $\Phi(r) = -G\mathcal{M}/(b + r)$ with $\mathcal{M} = 10^6 M_\odot$ and $b = 1$ pc, and the outer orbital peri/apocenter is chosen to be $(r_p/b,\ r_a/b) = (1.5,\ 1.7)$. See Section 5.1 for a detailed discussion.

regime, and the other is that its (ω, e) phase-space trajectory crosses the separatrix and becomes circulating (ultimately a low-$j_{\rm max}$ circulating trajectory). We also know from Figure 10 that these two occurrences can happen in any order. In this particular case the binary crosses the separatrix first: we see from panel (a) that a crosses $a_{\rm sep}$ around $t = 4800$ Myr, and from panel (b) that around this time the minimum eccentricity gets very close to zero and then starts to increase and ceases to be well described by $j_{\rm max} = j_+$. This inference is confirmed in panel (g), in which we see explicitly the evolution from libration to circulation that occurs during the time interval denoted by the yellow shaded stripe. Panel (i) confirms the expected scalings of $t_{\rm sec}$, Δa, and τ_a as the binary moves through regime $\mathcal{B}$.

At around $t = 7100$ Myr, the binary's dynamical evolution changes dramatically: the semimajor axis a approaches $a_{\rm div}$ and its decay accelerates, whereas the secular timescale becomes very short (decaying in the non-power-law fashion expected of regime $\mathcal{C}$—see Section 4.2.3). Furthermore we see that $j_{\rm min}^2 = 2p_{\rm min}/a$ is still a good approximation for the maximum eccentricity and, though we do not show it here, the minimum eccentricity at this stage is fairly well described by $j_{\rm max} = -\sigma j_-$ (we defer a more careful, "zoomed-in" discussion of the late stages of a slow merger to Section 5.2). These characteristics are the hallmarks of low-$j_{\rm max}$ circulating trajectories in the moderate GR regime. To confirm this, we look at panel (h), which shows the phase-space evolution during the green striped time interval. We see clearly that the binary rapidly evolves toward a purely high-eccentricity circulating trajectory (into regime $\mathcal{D}$), whereafter it soon enters the strong GR regime and then merges.

5.2. Example 3: $\Gamma = 0.42 > 1/5$. A Binary Initially in the Moderate GR Regime

In Figure 5 we show the evolution of a lower-mass binary ($m_1 = m_2 = 1.4 M_\odot$), but on the same outer orbit in the same cluster as in Example 2; thus we again have $\Gamma = 0.42$. We choose different initial conditions for the inner orbit, in particular $a_0 = 49$ au. We see from panel (a) that because of this choice, $a < a_{\rm weak}$ initially, putting the binary just inside the moderate GR regime (see also panel (e)). On the other hand, initially $a > a_{\rm sep}$ so the phase-space trajectory librates. This is different to Example 1 since in that case, by the time the binary reached the moderate GR regime, its phase-space trajectory was already circulating. Thus with Example 3 we will not only be able to focus on the "late-time" behavior of a slow merger, i.e., its evolution through the moderate and strong GR regimes (as promised in Section 5.1), but also to see a phase-space transition from librating to circulating within the moderate GR regime.

From panel (b) we see that while the binary is on a librating phase-space trajectory ($t \lesssim 4500$ Myr) its maximum eccentricity is well described by $j_{\rm max} = j_+$, and its secular period increases with time. Once it enters the circulating region, its secular period begins to decrease with time. Panel (i) shows that the $t_{\rm sec}(a)$ behavior is again split cleanly into four regimes $\mathcal{A} - \mathcal{D}$. Panels (j) and (k) show that the binary also exhibits the expected asymptotic behavior for $\Delta a(a)$ and $t_{\rm sec}(a)$ at small semimajor axes (i.e., in regime $\mathcal{D}$). This is unsurprising since we know from panel (g) that the binary has reached low-$j_{\rm max}$ circulation by this stage. On the other hand, the large-a scalings $|\Delta a| \propto a^{-3/2}$ (Equation (43)) and $\tau_a \propto a$ (Equation (44)) are never cleanly realized, because the binary does not begin its life in the weak GR regime.

Let us now turn to Figure 6, in which we zoom in on panels (a)–(d) of Figure 5, focusing on $t \gtrsim 6000$ Myr. In panel (b) we no longer show the $j_{\rm max} = j_+$ solid red line, since we know that for this time range the binary is certainly on a circulating trajectory. However, we have added a green dashed line that shows the minimum eccentricity that would be obtained if the binary was on a high-$j_{\rm max}$ circulating trajectory, i.e., with $j_{\rm max} = j_0$ (Equation (A14)). We have also added a blue dashed line showing the low-$j_{\rm max}$ solution $j_{\rm max} = -\sigma j_-$ (Equation (A15)). We see that until around 8000 Myr the evolution is best described as a high-$j_{\rm max}$ circulating trajectory with $j_{\rm max} = j_0$. There is then a transitional stage around $t \approx 9000$ Myr wherein neither high-$j_{\rm max}$ nor low-$j_{\rm max}$ is a good description (this corresponds to j_0^2 approaching and then crossing zero from above in Figures 9 and 10). After $t \approx 9400$ Myr the evolution is quite well described as a low-$j_{\rm max}$ circulating trajectory, $j_{\rm max} = -\sigma j_-$. On the other hand there is small systematic error in this prediction, which we will explain momentarily.

We finish our discussion of Example 3 by zooming in on the very latest stage of the evolution, $t \gtrsim 9400$ Myr, which we plot in Figure 7. It is clear from this figure that at these late times the conservation of $p_{\rm min}$ (panel (d)) begins to fail, evolving from its value on the red dashed line ($p_{\rm min}(t = 0) \approx 10^{-3.11}$ au) toward a slightly larger value. The reason for the evolution of $p_{\rm min}$ is that by this stage $e_{\rm min}$ has gotten so large that one cannot think of GW emission as being confined to just the peak-eccentricity portion of a secular cycle and negligible elsewhere. Instead there is a nonnegligible amount of emission throughout the whole cycle, so a decays continuously rather than in a step-like fashion, as can be confirmed by zooming in on panel (a). As a result, the binary's semimajor axis upon entering the peak of its next secular cycle is slightly smaller than when it left the peak of the previous cycle, so $\epsilon_{\rm GR}$ is slightly larger, and hence the binary achieves a larger value of $p_{\rm min} \approx a j_{\rm min}^2/2$ (see also Section 3.3 of Wen 2003). The magnitude of the oscillations in p also diminishes with time until, after around $t = 10{,}600$ Myr, the binary reaches the strong GR regime and the oscillations are quenched. As predicted in Section 3.3 the value of p itself then remains effectively constant almost all the way to merger, taking a value[14] $p_{\rm strong} \approx 10^{-2.86}$au $\approx 1.8 p_{\rm min}(t = 0)$. We see from panel (b) that the resulting underestimate of $p_{\rm min}$ at these late times leads to a slight overestimate of both the maximum and minimum eccentricities (the blue and red dashed lines each sit slightly too low in this panel)—hence the systematic error in the prediction $j_{\rm max} = -\sigma j_-$ which was computed using the $p_{\rm min}$ value from $t = 0$. Finally, we note that the conservation of $i_{\rm min}$ also fails in these later stages of the evolution (panel (c)), with $\cos i_{\rm min}$ undergoing a decrease from $\cos i_{\rm min}(t = 0) = 0.265$ to roughly $\cos i_{\rm strong} \approx 0.20$. This change—which occurs for the same reason as that in $p_{\rm min}$, and is also discussed in Section 3.3 of Wen (2003)—does not make a significant difference to our analysis because the $\cos^2 i_{\rm min}$, $\sin^2 i_{\rm min}$ terms in, e.g.,

[14] Wen (2003) estimated that $p_{\rm strong}/p_{\rm min}$ should lie in the approximate range (1, 3)—see Equation (31) in that paper and the surrounding discussion. In fact, the following simple physical argument suggests the value ought to be ≈ 2 (Ford & Rasio 2006). Since GW emission is very poor at dissipating angular momentum, $J \equiv \sqrt{G(m_1 + m_2)a(1 - e^2)} = \sqrt{G(m_1 + m_2)p(1 + e)}$ is roughly constant during this phase. When $p = p_{\rm min}$ we have $e \approx 1$, whereas upon circularization we have $p = p_{\rm strong}$ and $e \sim 0$; thus, $2p_{\rm min} \approx p_{\rm strong}$, i.e., $p_{\rm strong}/p_{\rm min} \approx 2$. For all numerical examples presented in this paper, $p_{\rm strong}/p_{\rm min} \in (1.4, 1.8)$.

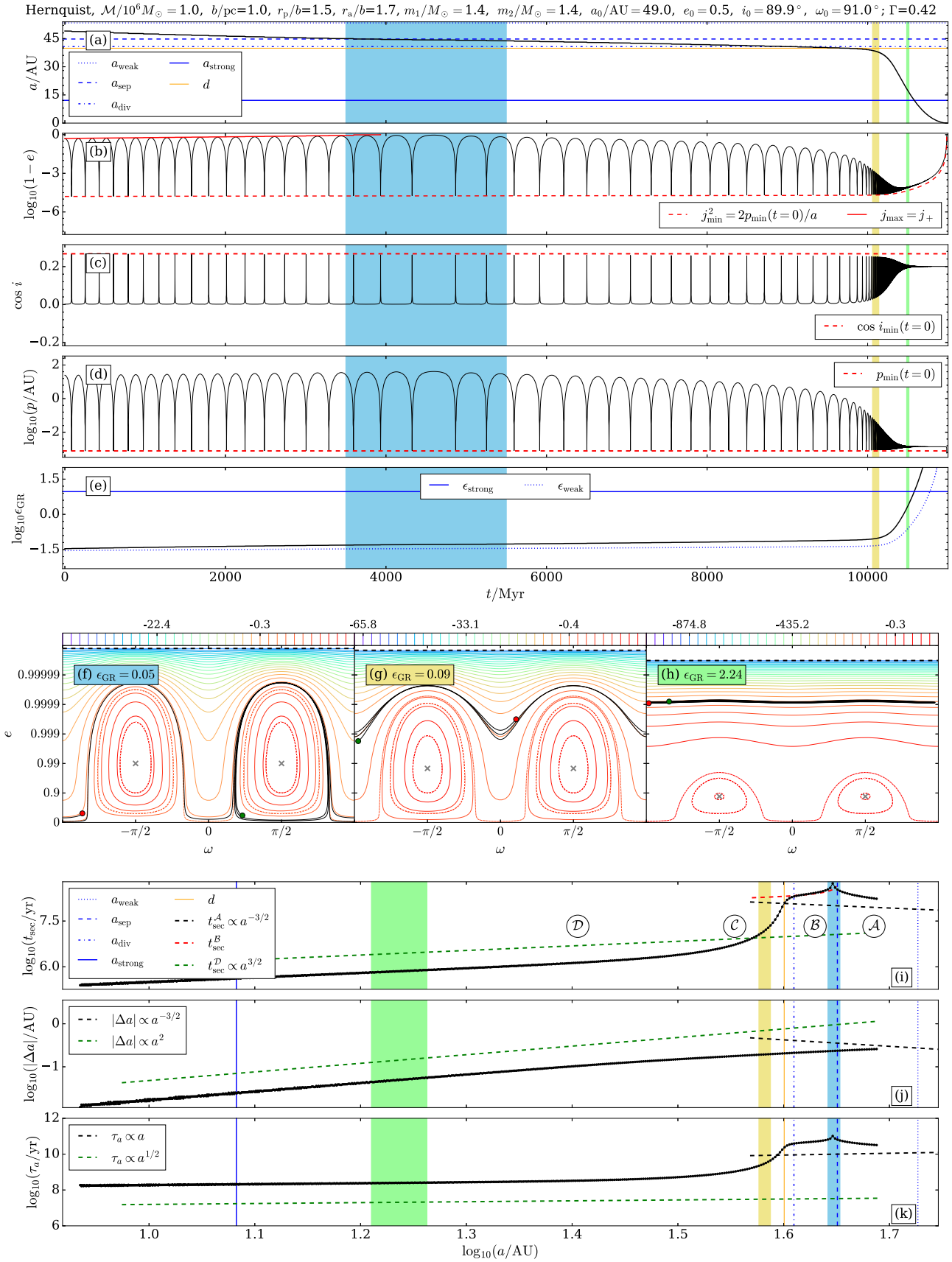

Figure 5. Example 3. The potential and outer orbit are the same as in Example 2 (Figure 4), so that again $\Gamma = 0.42$, but the binary constituent masses and inner orbit initial conditions are different. In this case the binary begins in the moderate GR regime on a librating phase-space trajectory.

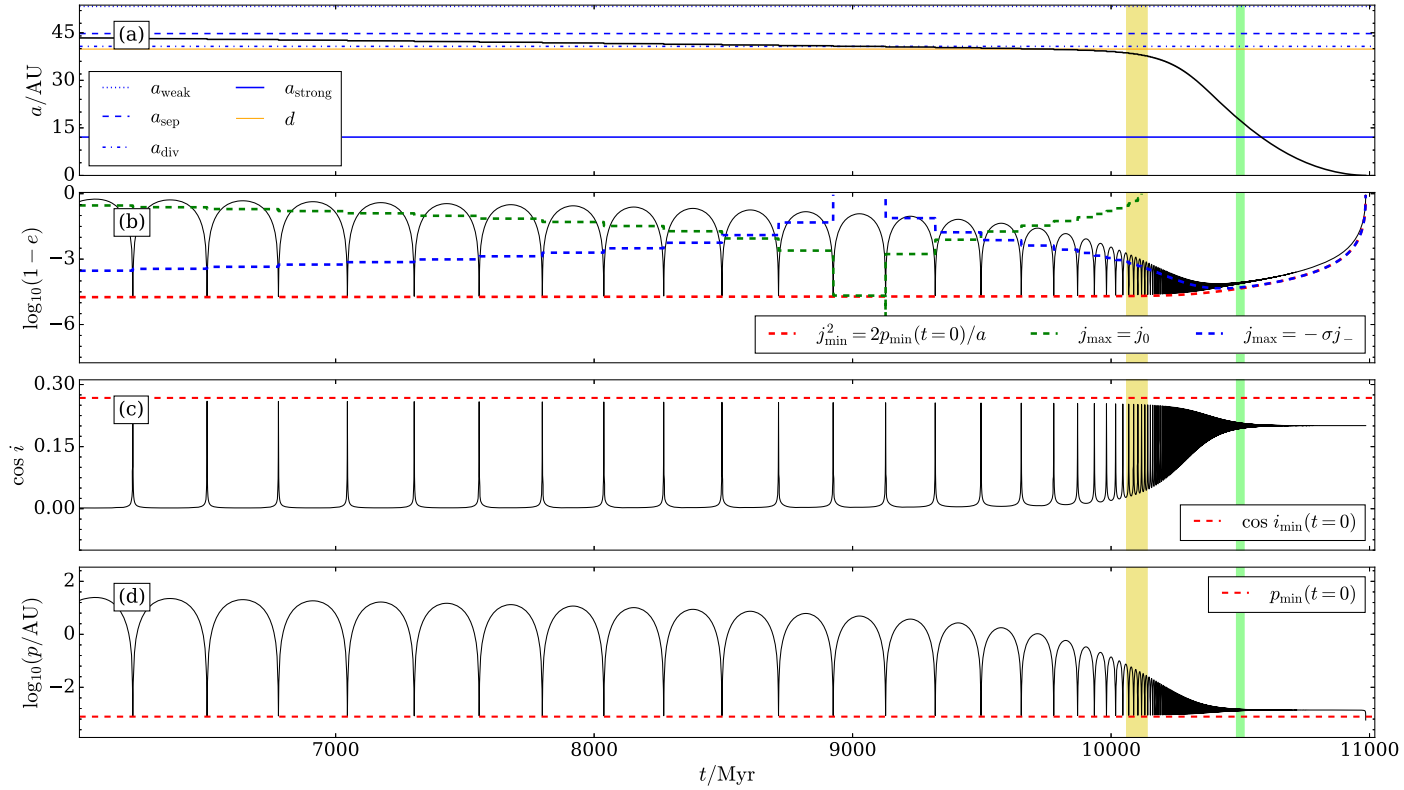

Figure 6. Zoomed-in version of panels (a)–(d) from Figure 5, focusing on $t > 6000$ Myr.

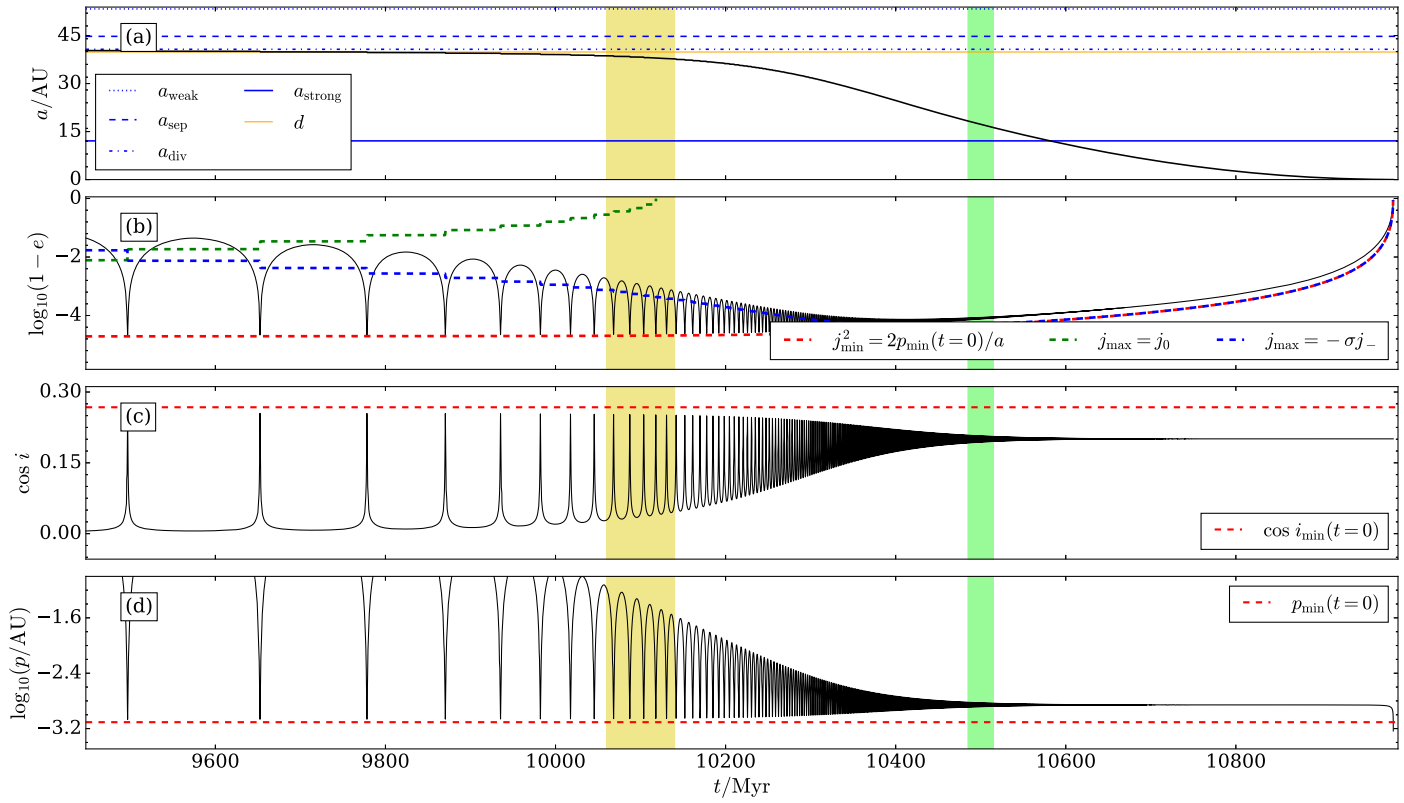

Figure 7. Further zoomed-in version of panels (a)–(d) from Figure 5, this time showing from $t \approx 9400$ Myr to merger.

Equations (B1)–(B3) are already dominated by the $(d/a)^{7/2}$ terms by this stage.

5.3. Example 4: $\Gamma = 0.176 < 1/5$

In Figure 8 we provide one more example, this time of an $m_1 = m_2 = 10M_\odot$ binary orbiting an $\mathcal{M} = 10^7 M_\odot$ Hernquist cluster. We take new initial parameters for the inner orbit, as well as a much smaller outer orbit $(r_p/b,\ r_a/b) = (0.1,\ 0.5)$, resulting in $\Gamma = 0.176 < 1/5$.

This binary begins its life in the weak GR regime ($a > a_{\rm weak}$), and is initially on a librating phase-space trajectory ($a > a_{\rm sep}$). It moves into the circulating regime ($a < a_{\rm sep}$) at around $t = 600$ Myr and then into the moderate regime ($a < a_{\rm weak}$) at around 720 Myr, and by around 860 Myr it has merged. Unsurprisingly, the secular period increases with time while the phase-space orbit librates (regime $\mathcal{A}$ in panel (i)), and decreases with time once it begins to circulate (regimes $\mathcal{C}$ and $\mathcal{D}$). There is no regime $\mathcal{B}$ in this example, which follows from the fact that in this case $a_{\rm sep} < a_{\rm div}$, so the divergence at $a = a_{\rm div}$ does not occur while the binary is on a circulating trajectory. In fact we always have $a_{\rm sep} < a_{\rm div}$ for $0 < \Gamma < 1/5$, as guaranteed by Equation (31). However we emphasize that this behavior is not limited to $\Gamma < 1/5$, and in fact it is easy to find examples of binaries with $\Gamma > 1/5$ that have $a_{\rm sep}/a_{\rm div} < 1$ and hence also have no $\mathcal{B}$ regime.

It is notable in this example that compared to Examples 1–3, the transition in the behavior of e, i, and p around 600 Myr is very abrupt. At the beginning of the yellow shaded time interval the binary is on a librating orbit (starting at the green dot in panel (g)); then at around 600 Myr it drifts through the separatrix and joins the family of circulating trajectories. By the end of the yellow interval it is on a very high-eccentricity circulating trajectory (ending at the red dot). In other words, at some point around the middle of the yellow time interval, the binary effectively "jumps" from libration to low-$j_{\rm max}$ circulation. The reason for this jump is that like in other examples, the binary inevitably ends up on a high-eccentricity circulating trajectory, and for $0 < \Gamma < 1/5$, the eccentricity of such a trajectory is forced to be larger than that of the saddle point at $\omega = 0$, namely $e_{\rm f,0}$—see Paper III. Thereafter the evolution matches the usual low-$j_{\rm max}$ behavior as seen for $\Gamma > 1/5$, followed by a merger.

Finally, at the extremes of panels (i)–(k) we see the expected asymptotic behavior for regimes $\mathcal{A}$ and $\mathcal{D}$ taking shape. However the proper scalings are never fully developed at large a, simply because at $t = 0$ the binary is too close to the separatrix and to $a_{\rm div}$ for j_+, j_0, etc. to be considered near-constant.

6. Discussion

In this paper, we have extended our theory of secular dynamics of binaries in stellar clusters by accounting for the effect of GW emission. We have demonstrated that cluster tides are capable of driving binaries to very high eccentricity, where they can emit GW bursts, shrink in semimajor axis, and ultimately merge. Our results also encompass—and in several ways extend—the theory of (test-particle, quadrupole, doubly averaged) LK-driven compact object binary mergers, which is recovered exactly in the limit $\Gamma = 1$.

Throughout this paper we have focused on understanding the physics of "slow mergers," i.e., those mergers that require many secular periods, but would not have occurred within a Hubble time had the cluster-tidal perturbation not been present to excite the binary eccentricity. Since we made the quadrupole-tide approximation, this meant focusing on initially high-inclination systems for which eccentricity excitation is possible. We ignored octupole effects, which can induce high-eccentricity excitation in triples without such high initial inclination (Naoz 2016). However, in the cluster tide-driven context, octupolar effects are so weak as to be completely ignorable, owing to the very small ratio between binary semimajor axis a and the distance to the cluster center $|\boldsymbol{R}_g|$ (see Appendix E of Paper I for details). We have also ignored short timescale fluctuations, stellar flybys, and the like (see Paper II for discussion), which can both excite eccentricity to high values and also complicate the delicate high-eccentricity dynamics (Grishin et al. 2018; Samsing et al. 2019). Yet even in this relatively simple setting we have seen that the evolution of a binary from "birth" to merger can be rather complex and is, in general, analytically intractable. The key to making analytical progress was our identification of several asymptotic regimes both in GR strength (weak, moderate, and strong) and in phase-space trajectory (librating, high-$j_{\rm max}$ circulating, and low-$j_{\rm max}$ circulating). We emphasize that the analytical results derived in these regimes are only approximations. In practice the boundaries between asymptotic regimes are blurry and poorly separated, especially in a-space. Nevertheless, they have been sufficient for our purpose, which was to gain analytical—and consequently, physical—insight into an important class of problems that have traditionally been outsourced to a computer. In Table 1 we summarize our approximate asymptotic results.

To conclude this paper, we first discuss in Section 6.1 the implications of our results for the calculation of the total merger timescale. Finally in Section 6.2 we discuss our work more broadly in the context of previous studies of LK-driven mergers.

6.1. Merger Timescale

Secular dynamics of binaries including GW emission is a problem that has been considered many times in the LK context for hierarchical triple systems. Many LK studies that include GW emission are focused on the resulting observable merger rate, i.e., the number of binaries that merge per cubic Gpc per year in the local universe. To compute such a rate—as we did for cluster tide-driven compact object mergers in Hamilton & Rafikov (2019a)—one needs to know the time it takes for a given binary to merge as a function of its initial conditions. There are basically two ways to approach this problem. One can either integrate the equations of motion (DA, SA, or N-body) directly and read off the merger time from the simulation (e.g., Antonini et al. 2014; Banerjee 2018; Bub & Petrovich 2020), or one can seek an approximate (semi)analytic formula that parameterizes the merger time in terms of those initial conditions (which can be checked using direct numerical integration for a small number of cases). The latter approach is obviously much faster when one is dealing with millions or billions of binary initial conditions in a Monte Carlo population synthesis (e.g., Liu & Lai 2018; Hamilton & Rafikov 2019b; Britt et al. 2021).

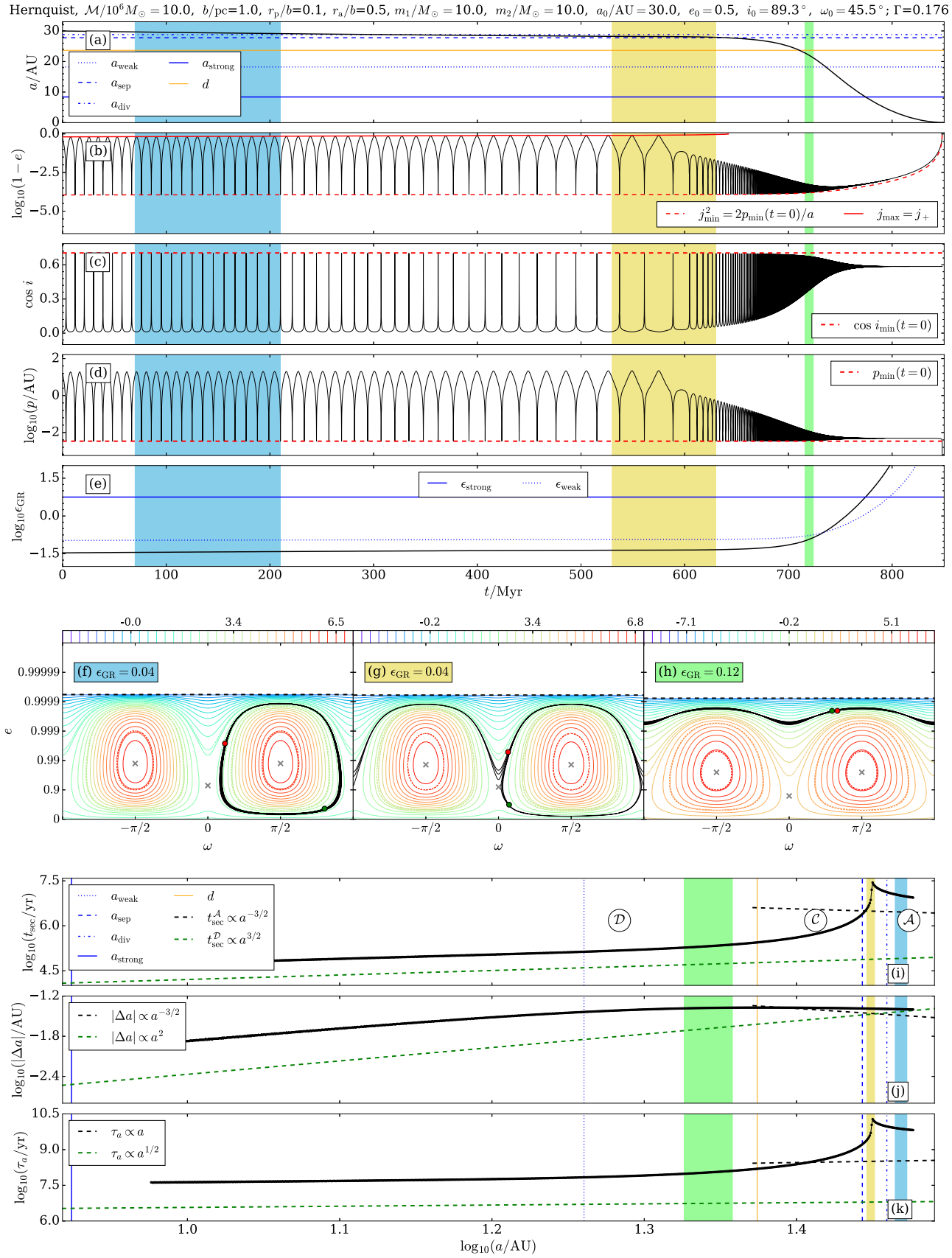

Figure 8. Example 4. Slow merger in the $0 < \Gamma \leqslant 1/5$ regime. The binary orbits an $\mathcal{M} = 10^7 M_\odot$ Hernquist cluster. The outer orbit is much smaller than in Examples 2 and 3, giving $\Gamma = 0.176 < 1/5$.

Table 1
Summary of Key Asymptotic Results

	Weak GR & Librating (Regime $\mathcal{A}$)	Moderate GR & Low-$j_{\rm max}$ Circulating (Regime $\mathcal{D}$)	Strong GR
Conserved Quantities	$p_{\rm min}, i_{\rm min}$	$p_{\rm min}, i_{\rm min}$	p, i
$t_{\rm sec}$	$\propto a^{-3/2}$	$\propto a^{1/2}$	N/A
Δa	$\propto a^{-3/2}$	$\propto a^{2}$	N/A
τ_a	$\Lambda_{\rm weak} U^{-1} a$	$\Lambda_{\rm mod} U^{-1} (2p_{\rm min})^{1/2} a^{1/2}$	$(8/5) U^{-1} (2p_{\rm min})^{1/2} a^{1/2}$

The merger time formula usually used in compact object merger calculations in the LK literature is

$$T_{\rm m} \equiv T_{\rm m}^{\rm iso}(a(0), e_{\rm max}(0)) \times (1 - e_{\rm max}^2(0))^{-1/2} = \frac{3c^5 a(0)^4}{85G^3(m_1+m_2)m_1m_2}(1 - e_{\rm max}(0))^3, \quad (55)$$

where

$$T_{\rm m}^{\rm iso}(a, e) = \frac{3c^5 a^4}{85G^3(m_1+m_2)m_1m_2}(1-e)^{7/2}, \quad (56)$$

is simply the merger time of an isolated binary with initial semimajor axis a and very high initial eccentricity $e \approx 1$ (Peters 1964). Equation (55) is typically justified via the following heuristic argument (Miller & Hamilton 2002; Thompson 2011; Liu & Lai 2018; Randall & Xianyu 2018). First, one assumes that the GW emission is negligible except around $e \approx e_{\rm max}$, and so the total amount of time that needs to be spent at $e \approx e_{\rm max}$ before the binary merges is $\approx T_{\rm m}^{\rm iso}(e_{\rm max})$. But the amount of time that the binary actually spends in the vicinity of $e_{\rm max}$ in each secular cycle is $\approx j_{\rm min}\, t_{\rm sec} \equiv (1 - e_{\rm max}^2)^{1/2} t_{\rm sec}$; thus, the number of secular cycles required until the time spent around $e_{\rm max}$ accumulates to $T_{\rm m}^{\rm iso}(e_{\rm max})$ is $T_{\rm m}^{\rm iso}(e_{\rm max})/[(1 - e_{\rm max}^2)^{1/2} t_{\rm sec}]$. To get the total merger time, we multiply this by $t_{\rm sec}$. Finally, evaluating everything at $t = 0$ we get Equation (55).

Of course, this heuristic derivation can be criticized on several levels. For instance, it makes no distinction between the values of $e_{\rm max}$, $t_{\rm sec}$ at $t = 0$ and their values at later times, even though we know (Section 3–4) that both of these quantities vary with a. Also, it does not accurately treat the behavior of e around $e_{\rm max}$, instead assuming that e is precisely equal to $e_{\rm max}$ within a discrete time window that lasts for $(1 - e_{\rm max}^2)^{1/2} t_{\rm sec}$, and that GW emission is negligible outside that window.

In reality, we know from Paper II that even in the absence of GR precession, the fraction of each secular period spent in the vicinity of high eccentricity is not precisely proportional to $t_{\rm sec}$. Indeed, combining Equations (34) and (59) of Paper II with Equations (33) and (45) of the present paper, we find that the time for j to change from $j_{\rm min}$ to $\sqrt{2} j_{\rm min}$ in the $\epsilon_{\rm GR} \to 0$ limit (in regime $\mathcal{A}$) is

$$t_{\rm min} \approx t_{\rm sec}^{\mathcal{A}} j_{\rm min} \times (2\Lambda_{\rm weak})^{-1}. \quad (57)$$

The factor $(2\Lambda_{\rm weak})^{-1}$ can be significantly different from unity if the binary is near a separatrix—see Paper II. Finally, the expression for $t_{\rm min}$ becomes even more complicated when we do include GR precession, especially for large values of σ and/or κ.

Despite these shortcomings, Equation (55) actually works reasonably well in practice (to within a factor of order unity) when compared to direct numerical integration of the (DA, test-particle quadrupole) equations of motion for triple systems (Thompson 2011; Liu & Lai 2018; Randall & Xianyu 2018). To see why this might be the case, we now show that one can actually derive Equation (55) in a slightly less hand-waving fashion using the results of this paper. For slow mergers, a rather general formula for the merger time $t_{\rm m}$ is found by integrating Equation (8) from $t = 0$ to $t = t_{\rm m}$:

$$t_{\rm m} \approx \int_{a(0)}^{0} da' \frac{t_{\rm sec}(a')}{\Delta a(a')} = \int_{a(0)}^{0} da' \frac{\tau_a(a')}{a'} \quad (58)$$

(see Equation (57) of Randall & Xianyu 2018). Of course, as it stands that Equation (58) is an entirely impractical formula given the complexity of the general analytic expression for τ_a that must then be integrated over. To make progress we assume that the majority of a slow merger is spent in the weak GR regime, and that by ignoring the time spent in the moderate and strong GR regimes, we do not impart any major error (though we note that this approximation would fail in Figure 5, for example). Then a decent approximation to τ_a is given by Equation (44). Plugging this into Equation (58) gives

$$t_{\rm m} \approx \Lambda_{\rm weak} U^{-1} a(0), \quad (59)$$

with $\Lambda_{\rm weak}$ and U given in Equations (45) and (46), respectively. Of course, $\Lambda_{\rm weak}$ accounts for the fact that the time spent at highest eccentricity is not precisely proportional to $t_{\rm sec}$—see Equation (57). Since $p_{\rm min}(t) \equiv a(t) \times (1 - e_{\rm max}(t))$ is conserved throughout a slow merger, we can substitute in Equation (46) the expression

$$(2p_{\rm min})^3 = 8(1 - e_{\rm max}(0))^3 a(0)^3 \approx (1 - e_{\rm max}^2(0))^3 a(0)^3, \quad (60)$$

where in the second line we assumed $e_{\rm max}(0) \approx 1$. Comparing the result to Equation (55), we find that in this approximation the merger occurs at time

$$t_{\rm m} \approx (5\Lambda_{\rm weak}/16) \times T_{\rm m}. \quad (61)$$

Thus provided $\Lambda_{\rm weak} \sim 1$, we recover the standard estimate of the merger timescale of Equation (55) to within a factor of order unity.

If anything, one might expect that $T_{\rm m}$ will be an overestimate of the "true" merger time (even if one calculates this "true" time by integrating the DA quadrupolar equations, i.e., ignoring SA effects, octupolar terms, and so on). That is because, as we saw in Section 4.2 and Section 5, the decay of $a(t)$ speeds up substantially once the binary reaches its low-$j_{\rm max}$ circulating phase in the moderate GR regime (see Equation (51)). Thus, approximating the entire decay using the weak GR, Equation (44) may seem overly conservative. It is therefore surprising to note Figure 8 of Thompson (2011) and Figure 5 of Randall & Xianyu (2018), both of which suggest that $T_{\rm m}$ typically *underestimates* the true (DA) merger

time by a factor ~ 2 for compact object binaries in hierarchical triple systems. In future work it might be interesting to understand more deeply the reason for this trend. It may also be profitable to try to use the results of this paper to calibrate a merger timescale formula that is more accurate than Equation (55)—even an estimate with typical in error at the level of only a few tens of percent would be a significant improvement. On the other hand, for realistic calculations, such a formula may be of limited interest, since the true merger time can be greatly shortened when one includes subsecular (e.g., "singly averaged") effects, octupolar terms, and so on (see, e.g., Antonini et al. 2014; Grishin et al. 2018).

Finally, throughout this paper we have been completely agnostic about *how* the black hole binary formed or how long that process took, and we have not worried about the likelihood of achieving the initial conditions we have chosen. Of course such considerations are crucial in population synthesis studies, meaning that the merger time formula is a necessary but insufficient piece of information for understanding whole populations of compact object mergers.

6.2. Relation to Studies of LK-driven Mergers

As discussed in Section 6.1, most LK studies "solve" the problem of GW-assisted mergers either by direct numerical integration or by stating and then evaluating the merger time of Equation (55) after calculating $e_{\max}$ from simple theory. There does not exist much in the literature that lies in between these extremes, in which an attempt is made to understand in detail the physics of each stage of the merger or to derive analytic results in specific asymptotic regimes as we have done here. Nevertheless, some of the key ideas covered in this paper have been considered by other authors, as we now describe.

A central result of Section 3 was the approximate conservation of $p_{\min}$ and $i_{\min}$ during slow mergers: this was ultimately what allowed us to express various important quantities ($j_{\min}$,$t_{\rm sec}$, etc.) as functions of a. These conservation laws (as well as their breakdown in the late stages of a slow merger) seem to have first been described in the LK limit by Wen (2003). The behavior of p during slow mergers has subsequently been appreciated as an important diagnostic of different regimes; for instance, Antonini has followed the p evolution in order to distinguish between "LK-dominated" and "GW-dominated" regimes (Antonini & Perets 2012; Antonini et al. 2014, 2017). Some basic scalings of $j_{\min}$, $t_{\rm sec}$, etc. with a were also written down by, e.g., Miller & Hamilton (2002), Wen (2003), and Thompson (2011), although none of these authors ventured beyond the weak GR regime in their analytical efforts, and so did not derive the peculiar results in the moderate GR regime that we have found here. No other studies have progressed beyond these simple scaling relations, to write down explicit formulae like we did in Section 4.

Another main achievement of the present paper has been to understand the interplay between the time evolution of key dynamical quantities like (a, e), and the underlying phase-space structure. The fact that a binary initially on a librating phase-space trajectory necessarily transitions into the circulating regime as it shrinks was first mentioned by Blaes et al. (2002; although they did not note the accompanying qualitative change in $t_{\rm sec}$ behavior). Of course, since LK theory corresponds to $\Gamma = 1 > 1/5$, no previous authors have noted the new behavior that arises in the $0 < \Gamma \leqslant 1/5$ regime, e.g., the abrupt changes in phase-space trajectory and the associated sharp "kink" in $t_{\rm sec}(a)$—see Section 5.3. Furthermore, to our knowledge, no LK study has distinguished between high-$j_{\max}$ and low-$j_{\max}$ circulating trajectories.

The only LK study we know of to have written down a formula for the decay in semimajor axis Δa over one secular cycle is Randall & Xianyu (2018)—see their Equation (55). These authors also wrote down an expression (their Equation (57)) that is essentially the same as our Equation (8), pertaining to the slow evolution of a. In addition, Randall & Xianyu (2018) seem to be the only authors who mentioned that $t_{\rm sec}$ can sometimes decrease as a shrinks, even though every author who has integrated the equations of motion numerically must have encountered this phenomenon. In Appendix E we look in detail at some of the calculations of Randall & Xianyu (2018). As we show there, Randall & Xianyu (2018) implicitly assumed weak GR and $\sigma \ll 1$ when deriving certain analytical results, so their calculations are not valid outside of this regime.

7. Summary

In this paper we studied the (2.5pN) GW-driven orbital decay and subsequent merger of binary systems, which are torqued to high eccentricity by cluster tides on secular timescales. We worked in the DA, test-particle quadrupole approximation and included the effect of (1pN) GR precession in our calculations. Our results may be summarized as follows.

1. Cluster tides are capable of torquing binaries to sufficiently high eccentricity that they emit bursts of GWs and ultimately merge. Cluster tide-driven eccentricity excitation is therefore a viable mechanism for producing LIGO/Virgo mergers, similar to LK-driven mergers that have been widely explored in the past. In fact (test-particle quadrupole DA) LK-driven mergers are simply a special case of the cluster tide-driven mergers considered here.
2. For slow mergers (those that take place over many secular periods), there are two approximate conservation laws that hold as the semimajor axis a decays, namely conservation of the minimum pericenter distance $p_{\min} = a(1 - e_{\max})$ and conservation of the minimum inclination reached $i_{\min}$. The evolution of a decaying binary through phase space can be understood in terms of these conserved quantities.
3. We uncovered several asymptotic regimes both in terms of GR strength and phase-space morphology. The different regimes exhibit different characteristic behaviors of secular timescale $t_{\rm sec}(a)$, decay in semimajor axis per cycle $\Delta a(a)$, and consequently the decay timescale $\tau_a(a)$.
4. We re-derived a formula for the merger timescale that has been frequently used in LK theory, and provided a more detailed justification for it than those that have been offered previously.

The insights from this paper will inform future studies of LK-driven and cluster tide-driven binary mergers.

We thank Ulrich Sperhake and Bence Kocsis for comments on an earlier version of this work. This work was supported by a grant from the Simons Foundation (816048, CH), STFC grant ST/T00049X/1, and Ambrose Monell Foundation (RRR).

Appendix A
High-eccentricity Results without Gravitational Wave Emission

In this section we gather some results from Papers II–III concerning cluster tide-driven secular dynamics *without* gravitational wave (GW) emission (but including GR precession). Though there is nothing strictly new here, it will be useful to have these results gathered in one place and written in a form that makes their meaning transparent.

Without GWs, the entire eccentricity evolution is dictated by Equation (15) of Paper III:

$$\frac{dj}{dt} = \pm\frac{6C}{Lj^2}\{(25\Gamma^2 - 1)[(j_+^2 - j^2)(j^2 - j_-^2) - \frac{\epsilon_{\rm GR}}{3(1+5\Gamma)}j]\left[j^2(j_0^2 - j^2) + \frac{\epsilon_{\rm GR}}{3(5\Gamma-1)}j\right]\}^{1/2}, \tag{A1}$$

where $j_\pm^2$, j_0^2 are given in Equations (16)–(19) of that Paper. It will be important that we are able to derive simple expressions for $j_\pm$, j_0, etc. in the high-eccentricity limit. If the binary initially has e not close to unity, then to reach high $e_{\rm max}$ it is necessary to have both $\Theta \ll 1$ and the binary initially in the weak-to-moderate GR regime. Making these assumptions, we can use Equations (46) of Paper III, which we repeat here:

$$j_+^2 \approx \frac{2\Sigma}{1+5\Gamma} \sim 1, \qquad j_-^2 \approx \frac{5\Gamma\Theta}{\Sigma} \sim \Theta \ll 1. \tag{A2}$$

Moreover, evaluating Equations (18)–(19) of Paper III at $\omega = \pm\pi/2$, $j = j_{\rm min} \approx 1$ we get the following equation for Σ which did not appear explicitly in Paper III:

$$\Sigma = 5\Gamma\left(\frac{\Theta}{j_{\rm min}^2} + \frac{\epsilon_{\rm GR}}{30\Gamma j_{\rm min}}\right). \tag{A3}$$

Using the approximations of Equations (A2) and (A3), we find that we can write $j_\pm^2$ and j_0^2 exclusively in terms of Γ and the three dimensionless numbers Θ, $\epsilon_{\rm GR}$, $j_{\rm min}$, which are constants when GW emission is ignored:

$$j_+^2 \approx \frac{10\Gamma}{1+5\Gamma}\left(\frac{\Theta}{j_{\rm min}^2} + \frac{\epsilon_{\rm GR}}{30\Gamma j_{\rm min}}\right), \tag{A4}$$

$$j_-^2 \approx j_{\rm min}^2\left(1 + \frac{\epsilon_{\rm GR} j_{\rm min}}{30\Gamma\Theta}\right)^{-1}, \tag{A5}$$

$$j_0^2 \approx \frac{10\Gamma}{5\Gamma-1}\left(1 - \frac{\Theta}{j_{\rm min}^2} - \frac{\epsilon_{\rm GR}}{30\Gamma j_{\rm min}}\right). \tag{A6}$$

We can write down an expression for $j_{\rm min}$ by assuming $\Gamma > 0$ and that maximum e is achieved at $\omega = \pm\pi/2$. Then (see Equation (52) of Paper III):

$$j_{\rm min} = \frac{\gamma j_-}{2}[1 + \sqrt{1+4\gamma^{-2}}] = \frac{1}{2j_+^2\epsilon_{\rm strong}}[\epsilon_{\rm GR} + \sqrt{\epsilon_{\rm GR}^2 + \epsilon_{\rm weak}^2}], \tag{A7}$$

where

$$\epsilon_{\rm weak} \equiv 6(1+5\Gamma)j_+^2 j_- \approx (720\Gamma\Sigma)^{1/2}\Theta^{1/2}. \tag{A8}$$

It follows that for weak GR ($\epsilon_{\rm GR} \ll \epsilon_{\rm weak}$),

$$j_{\rm min} \approx j_- \sim \Theta^{1/2} \gg \epsilon_{\rm GR}, \qquad \text{(weak GR)}. \tag{A9}$$

Similarly, for moderate GR ($\epsilon_{\rm weak} \ll \epsilon_{\rm GR} \ll \epsilon_{\rm strong}$) and using the constancy of $\cos i_{\rm min}$ (Section 3.3.2), we have

$$j_{\rm min} \sim \epsilon_{\rm GR} \sim \Theta^{1/2}, \qquad \text{(moderate GR)}. \tag{A10}$$

It follows from Equations (A4), (A5), (A9), and (A10) that in the weak-to-moderate regime, provided $\Gamma \sim 1$, we always have $j_+^2 \sim 1$ and $j_-^2 \lesssim j_{\rm min}^2 \sim \Theta \ll 1$.

Note that in Section 4 of Paper III we already arrived at the weak GR result of Equation (A9). However we did not arrive at the same moderate GR result (A10); instead, we found $j_{\rm min} \sim \epsilon_{\rm GR} \gg \Theta^{1/2}$. The reason for the discrepancy is that in Paper III we implicitly assumed that Θ was kept fixed while $\epsilon_{\rm GR}$ was increased. However, when the decay of a is due to GW emission, one cannot change $\epsilon_{\rm GR}$ without also changing Θ—see Equation (23). Accounting for this fact leads to Equation (A10).

A.1. Maximum Angular Momentum

For all of the phase-space trajectories in which we are interested, $j_{\rm min}$ is given by the same formula, Equation (A7). However it turns out (Section 4.2–5) that to understand the behavior of slow mergers, one must distinguish between qualitatively different trajectories, and in particular to know their maximum angular momentum $j_{\rm max}$ (corresponding to minimum eccentricity $e_{\rm min}$), so we will devote some effort to this now.

The maximum j can either be found at $\omega = \pm\pi/2$ (if the phase-space trajectory librates) or at $\omega = 0$ (if it circulates). In the librating case, we find[15]

$$j_{\rm max} \approx j_+ \qquad \text{(librating orbits)}. \tag{A12}$$

Finding an approximate expression for $j_{\rm max}$ for circulating trajectories is more complex, because there are two qualitatively different regimes of circulating trajectory to consider. The first type of circulating trajectory, which we call "high-$j_{\rm max}$," corresponds to $j_{\rm max} \sim 1$ or $e_{\rm min} \sim 0$, i.e., the binary undergoes an order-unity oscillation in eccentricity during each secular cycle. This is the classic type of circulating trajectory undergone by, for instance, a binary with $\Gamma > 1/5$ in the weak GR regime starting out with small eccentricity at $\omega \approx 0$ (see, e.g., Figure 12(f) for an illustration). The second type of circulating solution, which we call "low-$j_{\rm max}$," corresponds to $j_{\rm max} \ll 1$ or $e_{\rm min} \sim 1$, so that the oscillation in eccentricity is actually rather small despite $e_{\rm max}$ being large (see, e.g.,

[15] To see this, recall that for librating trajectories $j_{\rm max}$ is a solution to the quartic found by setting the first square bracket in Equation (A1) to zero. We can simplify this quartic by noting that, since librating trajectories loop around fixed points at $\omega = \pm\pi/2$, they necessarily have $j_{\rm max}^2 > j_{{\rm f},\pi/2}^2$. We know from Figure 3 of Paper III that $j_{{\rm f},\pi/2}^2 \gg \Theta$, and from Equations (A9)–(A10) that $\Theta \gtrsim j_-^2$, so we can ignore j_- in the quartic and write

$$j_{\rm max}^3 - j_+^2 j_{\rm max} + \epsilon_{\rm GR}/\epsilon_{\rm strong} \approx 0. \tag{A11}$$

Since in the weak-to-moderate GR regime we have $j_+^2 \sim 1$ and $\epsilon_{\rm GR} \ll \epsilon_{\rm strong}$, the solution is obviously $j_{\rm max} \approx j_+$.

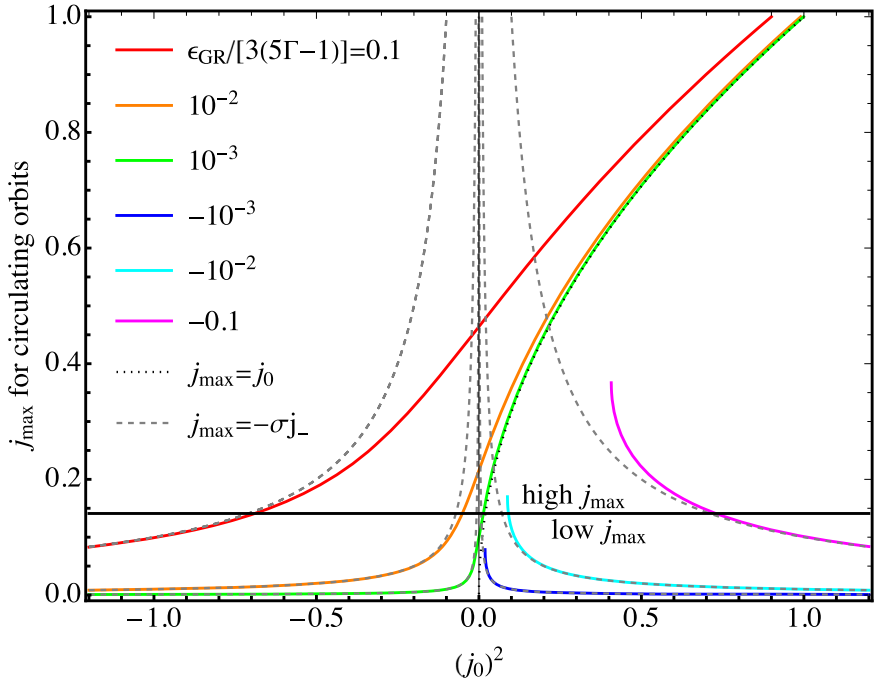

Figure 9. Solid colored lines show j_{max} (the physical solutions to Equation (A13)) for circulating trajectories as a function of j_0^2, for different values of $\epsilon_{GR}/[3(5\Gamma-1)]$, shown with different colors. (We know from Papers II–III that j_0^2 can be negative.) A horizontal black line at $j_{max}=0.141$ ($e_{min}=0.99$) separates "high-j_{max}" and "low-j_{max}" circulating solutions. Gray dashed lines show the solution $j_{max}=-\sigma j_-$, valid for low-j_{max} circulating trajectories (see Equation (A15)). The black dotted line shows $j_{max}=j_0$, valid for high-j_{max} circulating trajectories (Equation (A14)).

Figure 12(h)). In this case we can say that the binary is stuck at high eccentricity. Low-j_{max} circulating trajectories are important because every binary passes through this stage while in the moderate GR regime during a slow merger, as a precursor to the strong GR regime.[16]

To make the distinction between high-j_{max} and low-j_{max} trajectories quantitative, recall from Appendix A3 of Paper III that for all circulating trajectories j_{max} is a solution to the cubic equation

$$j_{max}(j_{max}^2 - j_0^2) - \frac{\epsilon_{GR}}{3(5\Gamma-1)} = 0. \quad \text{(A13)}$$

One can solve this cubic analytically, but for simplicity here we will just plot the solution. Figure 9 shows j_{max} as a function of j_0^2 (which can be positive or negative) for different values of $\epsilon_{GR}/[3(5\Gamma-1)]$, shown with different colored solid lines. In particular, red, orange, and green lines correspond to $\Gamma>1/5$ while blue, cyan, and purple lines are for $\Gamma<1/5$. We see from Figure 9 that for $\Gamma>1/5$, circulating solutions exist for all values of j_0^2. However for $\Gamma<1/5$, no solution exists below some (positive) value of j_0^2, consistent with what we found in Appendix B of Paper III.

We have chosen to split Figure 9 into two asymptotic regions, "high-j_{max}" ($j_{max}>0.141$, i.e., $e_{min}<0.99$) and "low-j_{max}" region ($j_{max}<0.141$, i.e., $e_{min}>0.99$). For high-j_{max} trajectories, provided j_0^2 is positive and $\mathcal{O}(1)$, we expect they are well approximated by ignoring the ϵ_{GR} term in Equation (A13), so that

$$j_{max} \approx j_0 \sim 1 \quad (\text{high}-j_{max}\ \text{circulating orbits}). \quad \text{(A14)}$$

In Figure 9 we plot this solution with a dotted black curve. For low-j_{max} trajectories, as long as $j_{max} \ll |j_0|$, we find from Equation (A13) that

$$j_{max} \approx \frac{\epsilon_{GR}}{3(5\Gamma-1)(-j_0^2)} = -\sigma j_- \ll 1 \quad (\text{low}-j_{max}\ \text{circulating orbits}). \quad \text{(A15)}$$

We plot this solution with different gray dashed curves, using the same values of $\epsilon_{GR}/[3(5\Gamma-1)]$ that we used for the colored solid lines. We see that for $\Gamma>1/5$ (red, orange, and green lines) the true solution interpolates between the two asymptotic solutions Equations (A14) and (A15) as j_0^2 is varied. For $\Gamma<1/5$ (blue, cyan, and purple lines), Equation (A15) provides a good approximation for sufficiently positive $j_0^2>0$. Overall we see that different j_{max} curves touch the $-\sigma j_-$ solution approximately at $j_{max}\approx 0.141$. In other words, binaries on circulating trajectories transition from high-j_{max} to low-j_{max} circulation around this point. Expressing j_{max} in Equation (A15) through e_{max} and using Equations (B2) and (B5) derived in the next section, we can calculate the semimajor axis at which this occurs, with the result

$$a \approx a_{div}\left[1+\frac{1-e_{max}}{10^{-2}}\right]^{2/7}, \quad \text{(A16)}$$

which is $\approx a_{div}$ in most cases of interest since typically $1-e_{max} \ll 10^{-2}$. Thus a good rule of thumb is that circulating

[16] For $0<\Gamma\leqslant 1/5$, we know from Paper III that high-e circulating trajectories are immediately formed once ϵ_{GR} exceeds $6(1-5\Gamma)\Theta^{3/2} \ll \epsilon_{weak}$. Thus, one does not necessarily need to be in the moderate GR regime to have low-j_{max} circulating orbits. However most of our focus in this paper will be on low-j_{max} circulating orbits that exist in the moderate GR regime, which occur for all Γ.

trajectories with $a > a_{\rm div}$ are high-$j_{\rm max}$, and circulating trajectories with $a < a_{\rm div}$ are low-$j_{\rm max}$.

In this discussion we have ignored one possible regime, namely that of low-$j_{\rm max}$ circulating trajectories with small $|j_0|$, i.e., $|j_0| \lesssim j_{\rm max} \ll 1$. However as Figure 9 shows, such solutions only exist for a narrow range of j_0^2 values centered around zero. This regime is typically short-lived in the sense that a shrinking binary passes through it rather quickly on the way to merger (equivalently it is centered on a very narrow semimajor axis range around $a \approx a_{\rm div}$). Throughout the rest of the paper we ignore this intermediate case, i.e., we always assume that low-$j_{\rm max}$ circulating trajectories have $|j_0| \gg j_{\rm max}$.

Appendix B
Phase-space Evolution and GR Regimes for Shrinking Binaries

In Section 3.3 we have seen how $j_{\rm min}$ and Θ depend on a—see Equations (21) and (23). We can use these results to understand how a binary moves through phase space as its semimajor axis shrinks. To begin, we substitute Equations (21), (23), and (6) into Equations (A4)–(A6) to get $j_\pm^2$, j_0^2 as explicit functions of semimajor axis:

$$j_+^2 \approx \frac{10\Gamma}{1+5\Gamma}\left[\cos^2 i_{\rm min} + \left(\frac{d}{a}\right)^{7/2}\right] = \frac{10\Gamma}{1+5\Gamma}\cos^2 i_{\rm min}\left[1 + \left(\frac{\ell a_{\rm weak}}{a}\right)^{7/2}\right], \tag{B1}$$

$$j_-^2 \approx \frac{2p_{\rm min}}{a}\left[1 + \frac{1}{\cos^2 i_{\rm min}}\left(\frac{d}{a}\right)^{7/2}\right]^{-1}, \tag{B2}$$

$$j_0^2 \approx \frac{10\Gamma}{5\Gamma-1}\left[\sin^2 i_{\rm min} - \left(\frac{d}{a}\right)^{7/2}\right] = \frac{10\Gamma}{5\Gamma-1}\sin^2 i_{\rm min}\left[1 - \left(\frac{a_{\rm div}}{a}\right)^{7/2}\right], \tag{B3}$$

where d is defined in Equation (24) and $\ell = [(\sqrt{2}-1)/2]^{-2/7} \approx 1.57$—see the definition of Equation (26). Next we write down the important dimensionless quantities γ, σ, and κ (familiar from Equations (49), (50), and (63) of Paper III, respectively) as functions of a, as follows. First, by combining Equations (A8), (B1), and (B2) and the definition of Equation (24), it is straightforward to show that

$$\gamma(a) \equiv \frac{2\epsilon_{\rm GR}}{\epsilon_{\rm weak}} \approx \frac{1}{\sqrt{\zeta(\zeta+1)}}, \quad \times \text{ where } \quad \zeta \equiv (a/d)^{7/2}\cos^2 i_{\rm min}. \tag{B4}$$

Second, plugging Equations (6), (B2), and (B3) into Equation (50) of Paper III we get:

$$\sigma(a) \approx \left[\left(\frac{d}{a}\right)^{7/2}\frac{1}{\cos^2 i_{\rm min}} + 1\right]^{1/2}\left[\left(\frac{a}{d}\right)^{7/2}\sin^2 i_{\rm min} - 1\right]^{-1}. \tag{B5}$$

Third, we can take the ratio of Equations (B5) and (B4) to get $\kappa \equiv \sigma/\gamma$:

$$\kappa(a) \approx \left[\left(\frac{a}{d}\right)^{7/2}\cos^2 i_{\rm min} + 1\right]\left[\left(\frac{a}{d}\right)^{7/2}\sin^2 i_{\rm min} - 1\right]^{-1} = \left[\left(\frac{a}{\ell a_{\rm weak}}\right)^{7/2} + 1\right]\left[\left(\frac{a}{a_{\rm div}}\right)^{7/2} - 1\right]^{-1}. \tag{B6}$$

These results lead naturally to the definitions of the critical semimajor axis values $a_{\rm sep}$ and $a_{\rm div}$ that we gave in Section 4.1. We now use these results, as well as the quantities $a_{\rm weak}$ (Equation (26)) and $a_{\rm strong}$ (Equation (27)), to understand more precisely how binaries move through phase space and different GR regimes as a decays. We begin with the regime $\Gamma > 1/5$, and then discuss $0 < \Gamma \leqslant 1/5$.

B.1. Phase-space Evolution for $\Gamma > 1/5$

In Figure 10 we plot j_+^2, j_0^2, $|\sigma|$, and $|\kappa|$ as functions of a/d for various fixed values of Γ and $\cos i_{\rm min}$, according to Equations (B1), (B3), (B5), and (B6), respectively. The choices of Γ and $\cos i_{\rm min}$ in panels (a)–(d) are chosen to coincide with the examples shown in Figures 1, 12, 4, and 5, respectively. We also show the critical values $a_{\rm weak}$ (dotted vertical line), $a_{\rm sep}$ (dashed vertical line), and $a_{\rm div}$ (dotted–dashed vertical line), defined in Equations (26)–(30). Additionally, in the upper panels we show with blue shading the region $|j^2| < (0.141)^2$, within which the split into "high-$j_{\rm max}$" and "low-$j_{\rm max}$" circulating trajectories is invalid (see the final paragraph of Appendix A). We show with orange shading the region $(0.141)^2 < j^2 < 1$. In particular, by looking at the runs of j_+^2 and j_0^2 and whether they lie in this orange region, we will be able to infer the value of $j_{\rm max}$ and hence infer what type of phase-space trajectory the binary is on. Without loss of generality, for each example (a)–(d), we can consider a binary that starts at the extreme right of each panel, i.e., with $a \gg a_{\rm weak}$ (the weak GR regime), and follow it as a decreases.

First we focus on panels (a) and (b), which are for $\Gamma = 1$ (the LK limit). In panel (a) the binary "begins" at large a with $0 < j_+^2 < 1$ and $j_0^2 > 1$; this means that it is on a librating trajectory in the weak GR regime, with $j_{\rm max} \approx j_+$ (Equation (A12)). Of course as a is decreased j_+^2 is always increased, while j_0^2 is decreased, and when $a = a_{\rm sep}$ the two cross over, $j_0^2 = j_+^2 = 1$. At this point the binary switches to a high-$j_{\rm max}$ circulating trajectory with $j_{\rm max} \approx j_0$. In this case $a_{\rm sep} < a_{\rm weak}$, so that the separatrix crossing occurs while the binary is still in the weak GR regime. Once a becomes smaller than $a_{\rm div}$ we quickly get j_0^2 values that are strongly negative, and the binary transitions to a low-$j_{\rm max}$ circulating trajectory (Figure 9) with $j_{\rm max} \approx -\sigma j_- \ll 1$. It will remain on such a trajectory until it gets trapped at high eccentricity in the strong GR regime around $a \sim a_{\rm strong}$ (not shown here).

Example (b) shows very similar behavior to Example (a), except that the smaller value of $\cos i_{\rm min}$ means that the three values $a_{\rm weak}$, $a_{\rm sep}$, and $a_{\rm div}$ are now even more closely clustered together around $a/d \approx 1$ (note also that $a_{\rm sep}$ is now very slightly smaller than $a_{\rm weak}$). Because of this clustering, Example (b) is perhaps "cleaner" than Example (a): for a significantly larger than d, the binary is clearly on a librating

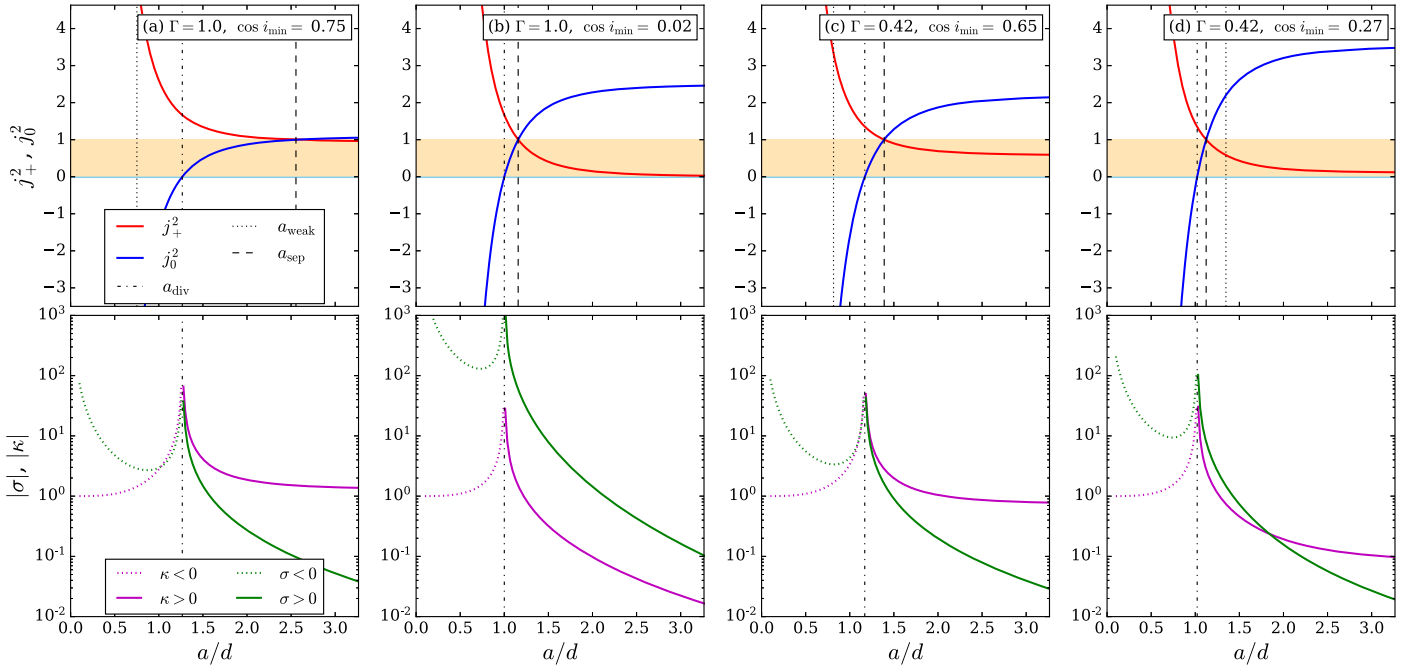

Figure 10. Plots of the key quantities j_+^2, j_0^2, σ, and κ as functions of a/d, for different values of $\Gamma > 1/5$ and $\cos i_{min}$. In the lower panels, for which the vertical axis is on a logarithmic scale, we show negative values of σ, κ with dotted curves and positive values with solid curves. In each panel we also show a_{weak} (vertical dotted line), a_{sep} (vertical dashed line), and a_{div} (vertical dotted–dashed line). Finally in the upper row we show with blue shading the (very thin) region $|j^2| < (0.141)^2 \approx 0.02$, within which the split into "high-j_{max}" and "low-j_{max}" circulating trajectories is invalid (Section A), and with pale orange shading the region $(0.141)^2 < j^2 < 1$. The values of Γ and $\cos i_{min}$ in panels (a)–(d) are chosen to coincide with the examples shown in Figures 1, 12, 4, and 5, respectively.

trajectory in the weak GR regime, whereas for a significantly smaller than d, it is clearly on a low-j_{max} circulating trajectory in the moderate GR regime. In practice, the transitions between these two various regimes are not always so well demarcated.

At this stage it is worth noting how different quantities scale with a in each regime. From Examples (a)–(b), we see that in the weak GR regime ($a > a_{weak}$) we nearly always have $|j_+^2|$, $|j_0^2| \gg 0.1$, and both of these quantities scale very weakly with a. In the moderate GR regime ($a < a_{weak}$) the scaling of j_+^2 and j_0^2 with a is much stronger, as we would expect from Equations (B1) and (B3). Moreover, in every case it is clear that $|j_0|^2$ lies in the blue shaded region only for a very narrow range of semimajor axes surrounding a_{div} (Equation (A16)), and so we were justified in ignoring the small j_0 regime when discussing low-j_{max} circulating trajectories in Appendix A. Turning to the bottom panels, we see that $|\sigma|$ and $|\kappa|$ both vary over several orders of magnitude as a is decreased. However, it is noteworthy that for a far away from a_{div}, the value of $|\kappa|$ is usually $\mathcal{O}(1)$ and scales weakly with a.

Finally we turn to Examples (c) and (d), which are for $\Gamma = 0.42$. The physical interpretation of these examples is identical to those of Examples (a) and (b), demonstrating a broad uniformity of evolution for all binaries in the $\Gamma > 1/5$ regime. In fact, this broad-brush picture can break down very close to $\Gamma = 1/5$, but we ignore this complication here.

B.2. Phase-space Evolution for $0 < \Gamma \leqslant 1/5$

In Figure 11 we plot the same quantities as in Figure 10, except this time we focus on the regime $0 < \Gamma \leqslant 1/5$. In particular the choices of Γ and $\cos i_{min}$ in panel (a) coincide with those from Figure 8. We see that a rather different phase-space evolution emerges for $0 < \Gamma \leqslant 1/5$ compared to $\Gamma > 1/5$.

First we consider panel (a), which is for $\Gamma = 0.176$ and $\cos i_{min} = 0.7$. In this case, for large $a \gg d$ we have $j_+^2 \lesssim 1$, while j_0^2 is large and negative. This means that in the asymptotic weak GR regime, the binary is on a librating trajectory, with $j_{max} \approx j_+$. However, once a decreases below a_{div} in this plot, we see that j_0^2 becomes positive (though still smaller than j_+^2). Soon a reaches a_{sep}, below which both j_+^2 and j_0^2 are greater than unity: the binary has transitioned onto a low-j_{max} circulating orbit (Figure 9). We note that all of this happens well before the binary reaches the moderate GR regime. This is not surprising because we know that a family of high-eccentricity circulating trajectories (i.e., low-j_{max}) naturally arises in the $0 < \Gamma \leqslant 1/5$ regime as soon as ϵ_{GR} exceeds $6(1 - 5\Gamma)\Theta^{3/2} \ll \epsilon_{weak}$ (Paper III). The binary stays on its low-j_{max} circulating trajectory as a shrinks into the moderate GR regime $a < a_{weak}$ and onward to the strong GR regime.

A very similar story holds in panels (b)–(d). The only important difference is that as we decrease Γ or $\cos i_{min}$, or both, the value of $j_{max} \approx j_+$ for asymptotically weak GR ($a \gg d$) decreases. This means that librating trajectories with high $e_{max} \to 1$ in the very weak GR regime do not reach low[17] e_{min} (recall that we have assumed $j_{min} \ll 1$ in deriving our expression for j_+^2). Said differently, for $0 < \Gamma < 1/5$, binaries that initially have $e \sim 0$ do not tend to reach $e \to 1$—that is, low minimum eccentricities are not typically associated with high maximum eccentricities, so this is typically not the type of situation in which we are interested.

[17] This is essentially because the fixed points at $\omega = \pm\pi/2$, $j = j_f = (10\Gamma\Theta/(1 + 5\Gamma))^{1/4}$ sit at too high an eccentricity.

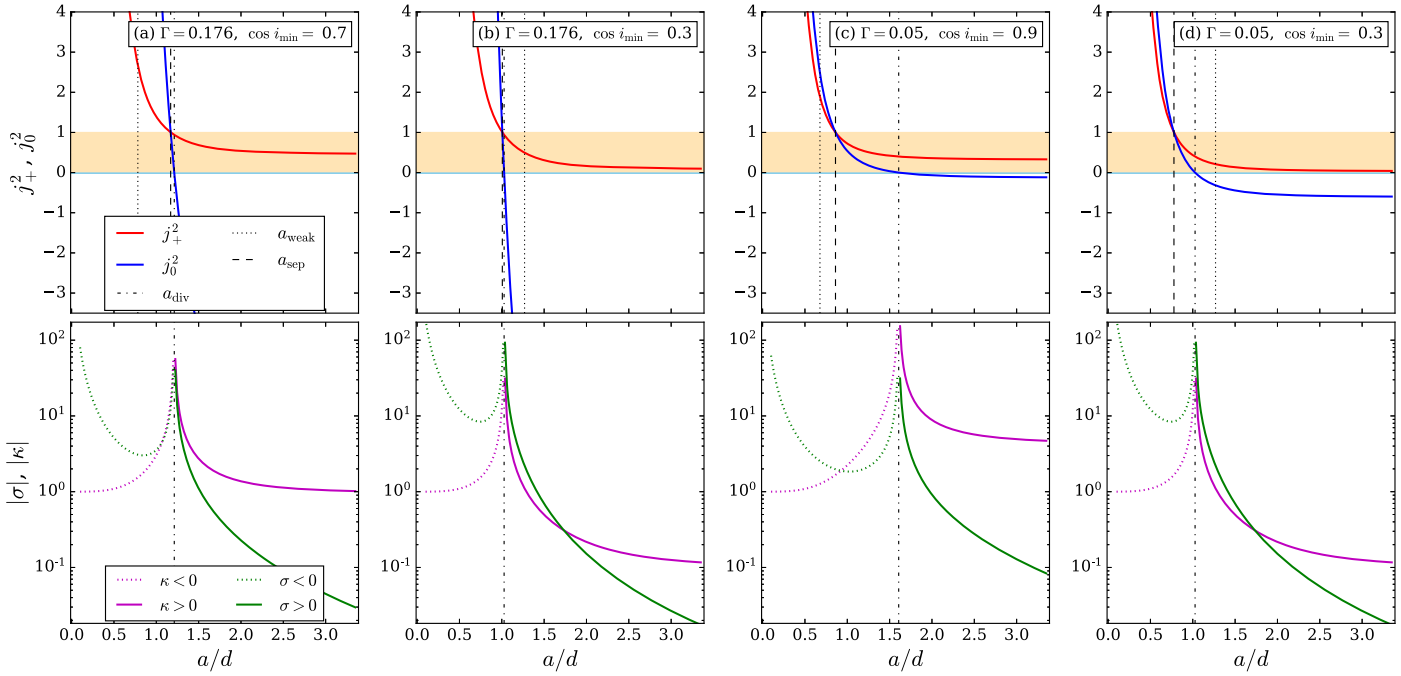

Figure 11. The same as in Figure 10 but for the regime $0 < \Gamma \leqslant 1/5$. The choices of Γ and $\cos i_{\rm min}$ in panel (a) coincide with those from Figure 8.

Finally we mention that in all examples shown in Figure 11, for a sufficiently far from $a_{\rm div}$, we again have $|\kappa| \sim \mathcal{O}(1)$ or smaller, and κ varies only weakly with a.

Appendix C
Derivation of $t_{\rm sec}$ Formula in Regimes $\mathcal{C}$ and $\mathcal{D}$

As a crosses $a_{\rm div}$ and the binary enters regime $\mathcal{C}$, j_0^2 passes through zero and rapidly becomes strongly negative—see Equation (B3) and Figures 10–11. Those Figures also show that j_+^2 is typically large in amplitude for $a \lesssim a_{\rm div}$, certainly larger than j^2, which is limited by $j_{\rm max}^2$, and which is already in the low-$j_{\rm max}$ regime. Also, for most of the secular cycle, we can neglect j_-^2 compared to j^2 since $j_{\rm max} = -\sigma j_-$ is well separated from $j_{\rm min} \approx \gamma j_-$. Taking the limit $j_-^2 \ll j^2 \ll j_+^2, j_0^2$ (which is most accurate in regime $\mathcal{D}$ with moderate GR), Equation (A1) reduces to

$$\frac{dj}{dt} \approx \pm\frac{6C}{Lj}\sqrt{|25\Gamma^2 - 1|}\,|j_+ j_0|\sqrt{(j - j_{\rm min})(j_{\rm max} - j)}. \quad \text{(C1)}$$

Plugging this into Equation (32) and performing the integral, we get

$$\begin{aligned} t_{\rm sec} &\approx \frac{L\pi(j_{\rm min} + j_{\rm max})}{6C\sqrt{|25\Gamma^2 - 1|}\,|j_+ j_0|} \\ &\approx \frac{8}{3A}\sqrt{\frac{G(m_1+m_2)}{|25\Gamma^2 - 1|}} \times \frac{\pi}{2a^{3/2}}\frac{(1-\kappa)j_{\rm min}}{|j_+ j_0|}, \end{aligned} \quad \text{(C2)}$$

where in the second line we used $j_{\rm min} + j_{\rm max} \approx \gamma j_- - \sigma j_- = (1-\kappa)j_{\rm min}$ (see Equation (A15)). Using Equations (21), (B1), (B3), and (B6) for $j_{\rm min}$, j_+^2, j_0^2, and κ, correspondingly, we find

$$t_{\rm sec} \approx \frac{2\pi}{15\Gamma A}\frac{\sqrt{2G(m_1+m_2)p_{\rm min}}}{a^2 \cos i_{\rm min} \sin i_{\rm min}} f(a),$$

$$\text{with} \quad f(a) = \left(\frac{a}{a_{\rm div}}\right)^{7/4} \frac{2 + \left(\frac{a}{\ell a_{\rm weak}}\right)^{7/2} - \left(\frac{a}{a_{\rm div}}\right)^{7/2}}{\sqrt{1 + \left(\frac{\ell a_{\rm weak}}{a}\right)^{7/2}}\left|1 - \left(\frac{a}{a_{\rm div}}\right)^{7/2}\right|^{3/2}} \quad \text{(C3)}$$

Different limits of this expression in regimes $\mathcal{C}$ and $\mathcal{D}$ are explored in Sections 4.2.3 and 4.2.4.

Appendix D
Derivation of an Approximate Formula for Semimajor Axis Decay

Assuming the binary reaches very high maximum eccentricity $e_{\rm max} \to 1$, we can approximate Equation (40) as

$$\Delta a \approx -\frac{2\lambda_1}{a^3}\int_{j_{\rm min}}^{j_{\rm max}} \frac{dj}{j^7}\left(\frac{dj}{dt}\right)^{-1}, \quad \text{(D1)}$$

where $\lambda_1 \equiv (1 + 73/24 + 37/96)\lambda_0 = (170/3)G^3 c^{-5} m_1 m_2 (m_1 + m_2)$. In general, dj/dt—given in Equation (A1)—is so complicated that even this approximate integral is intractable. However, noting the very strong j^{-7} dependence in Equation (D1), we expect the integral to be dominated by the contributions from very high eccentricity, i.e., $j \ll j_+, |j_0|$. In this limit, we can approximate dj/dt using Equation (48) of Paper III. Moreover, since we know that the minimum $j_{\rm min}$ is a zero of the first square bracket in that equation, we can write

it as

$$\frac{dj}{dt} \approx \pm \frac{3Aa^{3/2}}{4\sqrt{G(m_1+m_2)}j^{3/2}} \times \sqrt{(25\Gamma^2-1)j_+^2(-j_0^2)(j-j_{\rm min})(j+|j_\alpha|)(j_\sigma-j)}, \tag{D2}$$

where $j_\alpha \equiv \gamma j_-[1-\sqrt{1+4\gamma^{-2}}]/2 < 0$ is the other root of the first square bracket in Equation (48) of Paper III, and $j_\sigma \equiv -\sigma j_-$.[18]

We now take Equation (D2) and plug it into Equation (D1). Defining

$$x_{\rm max} \equiv j_{\rm max}/j_{\rm min}, \quad x_\alpha \equiv j_\alpha/j_{\rm min}, \quad x_\sigma \equiv j_\sigma/j_{\rm min}, \tag{D3}$$

and using Equation (21), the result is

$$\Delta a \approx -\lambda_2 \times \frac{\xi(x_{\rm max}, x_\alpha, x_\sigma)}{a^{3/2}|j_+ j_0|}, \tag{D4}$$

where $\lambda_2 \equiv 1360G^{7/2}m_1m_2(m_1+m_2)^{3/2}/[9c^5A(2p_{\rm min})^3\sqrt{|25\Gamma^2-1|}]$ is independent of a, and

$$\xi(x_{\rm max}, x_\alpha, x_\sigma) \equiv \int_1^{x_{\rm max}} \frac{dx}{x^{11/2}\sqrt{(x-1)(x+|x_\alpha|)|x_\sigma - x|}}. \tag{D5}$$

We can simplify this result in the limit of weak GR. In this limit we have $j_{\rm min} \ll j_{\rm max}$ so that $x_{\rm max} \gg 1$. We also have $j_\sigma < 0$, so that $|x_\sigma - x| = x + |x_\sigma|$. In this case, the integral in Equation (D5) is completely dominated by the contribution from $x \approx 1$, and so we may take the upper limit of the integral to $x_{\rm max} \to \infty$ with impunity. Since $j_{\rm min} \approx j_-$ and $\gamma \ll 1$ in this limit (see Equation (B4)), we can simply replace x_σ with σ and $x_\alpha \to 1$. An excellent approximation to the resulting integral (accurate to within a few percent over several decades of $|\sigma|$) is then given by Equation (42).

Appendix E
Relation to Randall & Xianyu (2018)

Throughout the main text we referred to the paper by Randall & Xianyu (2018)—hereafter RX18—several times. The RX18 paper largely inspired the present work, since those authors are among the few who have attempted to gain an analytical understanding of LK-driven slow mergers (indeed it is from their paper that we have taken the terminology "slow merger"). In particular, to our knowledge, RX18 were the first to (i) calculate Δa explicitly, and (ii) comment upon the decrease in $t_{\rm sec}$ as the binary shrinks and offer an explanation thereof. On the other hand, we feel that both (i) and (ii) as presented in RX18 can be improved. In this Appendix Section we explain how our calculations differ from those of RX18 regarding points (i) and (ii) (Appendices E.1 and E.2, respectively).

To begin, we present Figure 12. This Figure reproduces exactly the numerical example shown in Figure 3 of RX18, from which those authors drew several of their conclusions. Specifically, it follows the evolution of a binary of $m_1 = m_2 = 10M_\odot$ and $a_0 = 0.1$ au as it orbits an SMBH of mass $4 \times 10^6 M_\odot$. We see that in this example the binary sits from the start in the moderate (rather than weak) GR regime on a circulating phase-space trajectory, and that the secular time-scale does indeed decrease as the binary shrinks. The merger occurs after around $t = 7000$ yr. We will refer to this Figure frequently throughout the remainder of this section.

[18] Using the results of Appendix A, one can check that the sign of the quantity inside the square root is positive. For instance, for $\Gamma > 1/5$ we recall that low-$j_{\rm max}$ circulating trajectories have $j_0^2 < 0$ and $j_{\rm max} = -\sigma j_- = j_\sigma$, while Type 1 circulating trajectories have $j_0^2 > 0$, $j_\sigma < 0$, and $j_{\rm max} \sim 1$.

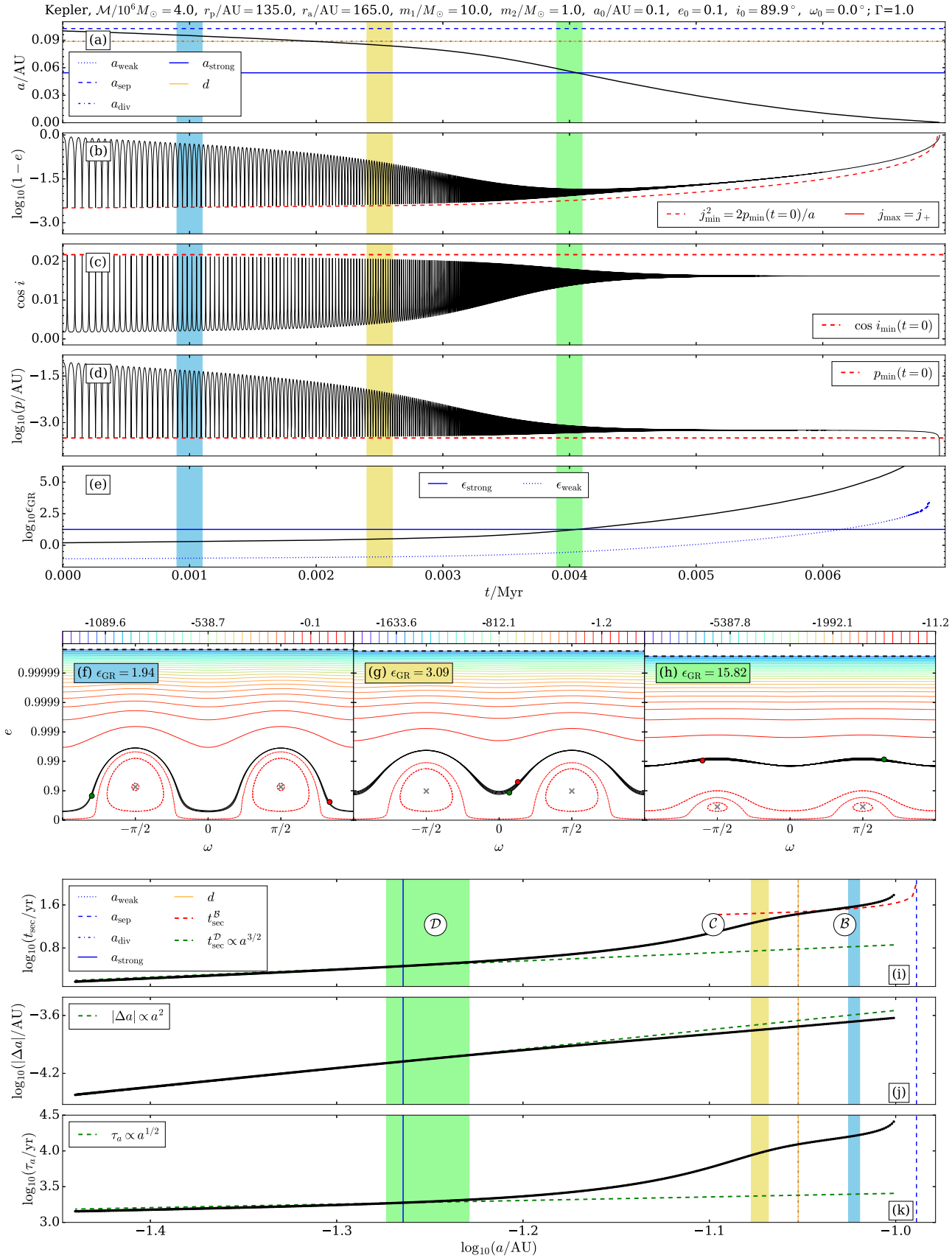

Figure 12. Reproduction of Figure 3 of Randall & Xianyu (2018). In this case, a binary with $m_1 = m_2 = 10M_\odot$ orbits an SMBH (i.e., Kepler potential, $\Gamma = 1$) of mass $\mathcal{M} = 4 \times 10^6 M_\odot$. The outer orbit has semimajor axis $a_g = (r_a + r_p)/2 = 150$ au and eccentricity $e_g = (r_a - r_p)/(r_a + r_p) = 0.1$. Note that $a_{\rm div}$ and d overlap almost exactly, which follows from the fact that in this example $\cos i_{\rm min} \ll 1$ (see Equation (30)).

E.1. Calculation of Δa

RX18 began their calculation of Δa by writing down their Equation (55), the first two lines of which are identical to our Equation (40) if we evaluate the final bracket at $e = e_{\max}$. One is then faced with the computation of an integral, $\Delta a \propto \int dt(1 - e^2(t))^{-7/2}$, over one secular cycle. To perform this integral, in Section 4.3 we changed variables from $t \to j \in (j_{\min}, j_{\max})$ and hence wrote down Equation (D1). On the other hand, RX18 chose to compute the integral by first approximating $e(t)$ as a quadratic in time (see their Equation (53)). In particular, using our notation and letting the maximum eccentricity occur at $t = 0$ without loss of generality, their Equation (52) reads

$$e(t) = e_{\max} + \frac{1}{2}\left(\frac{d^2e}{dt^2}\right)_{t=0} t^2. \tag{E1}$$

RX18 then plugged this into $\int dt(1 - e^2(t))^{-7/2}$ and integrated over $t \in (-\infty, \infty)$ to get Δa. The result is their second Equation (55), which in our notation and evaluating at $e_{\max} \approx 1$ reads

$$\Delta a_{\rm RX18} \approx -\frac{544 G^3 m_1 m_2 (m_1 + m_2)}{9 c^5 a^3 j_{\min}^6} \times \left|\frac{d^2e}{dt^2}\right|_{t=0}^{-1/2}. \tag{E2}$$

Finally, RX18 evaluated $\ddot{e}|_{t=0}$ using their Equation (53).

However, RX18's method for computing Δa implicitly makes two assumptions that are not true in general, as we now explain.

1. The assumption that $e(t)$ is quadratic for small t is equivalent to the assumption that $j(t)$ is quadratic for small t. We know from Paper III that this quadratic approximation is only good if the binary is in the weak GR regime ($\epsilon_{\rm GR} \ll \epsilon_{\rm weak}$) and it has $\sigma \ll 1$ (Equation (B5)). While these conditions do hold for many binaries of interest (i.e., see the early stages of Figures 1 and 4, for which $\sigma \approx 0.02$ and 0.08, respectively), they are not true for the RX18 calculation shown in Figure 12—this example begins in the moderate GR regime (panel (e)) and has $\sigma \approx 14.8$.
2. The equation that RX18 quote for $\ddot{e}|_{t=0}$—namely their Equation (53)—is a poor approximation in general. To see this, we compute the "exact" value of $\ddot{e}|_{t=0}$ directly by differentiating $e = (1 - j^2)^{1/2}$ twice, using the DA equations of motion (see Equations (12)–(13) of Paper III), and demanding that at $t = 0$, $j = j_{\min}$, $dj/dt = 0$ and $\omega = \pm\pi/2$. Without any approximations, we find

$$\left(\frac{d^2e}{dt^2}\right)_{t=0} = -\frac{60\Gamma C}{L}\frac{(j_{\min}^2 - \Theta)e_{\max}}{j_{\min}}\left(\frac{d\omega}{dt}\right)_{t=0}. \tag{E3}$$

For this to coincide with Equation (53) of RX18 in the LK ($\Gamma = 1$) limit, one must have $(j_{\min}^2 - \Theta) \approx j_{\min}^2 = (1 - e_{\max}^2)$, which is only true if[19]

$$j_{\min}^2 \gg \Theta, \qquad \text{i.e.} \quad \cos^2 i_{\min} \ll 1. \tag{E4}$$

The condition in Equation (E4) *does* happen to be true in the specific numerical example shown in Figure 12, but it is certainly not true in general, as we have seen in several numerical examples (Figures 1, 4, 5, and 8). In fact, we know from Appendix A that if a slow-merging binary is initially in the weak GR regime, then it has $j_{\min}^2 \sim \Theta$ all the way into the moderate GR regime and beyond, so in general one should use the formula of Equation (E3).

We can make a direct comparison between our method of computing Δa and that of RX18 as follows. Let us follow the RX18 method and use Equations (E2) and (E3), evaluating $d\omega/dt$ at maximum eccentricity using Equation (12) of Paper III—we call the result $\Delta a_{\rm RX18}$. We then compare the result to our equation for Δa, namely Equation (D4). Using $j_{\min}^2 = 2p_{\min}/a$ and after some algebra, we arrive at

$$\frac{\Delta a_{\rm RX18}}{\Delta a} = \frac{8}{15}\frac{\sqrt{|25\Gamma^2 - 1|}}{10\Gamma}\frac{|j_+ j_0|}{\xi}\frac{j_{\min}^2}{\sqrt{j_{\min}^2 - \Theta}} \times \left(10\Gamma\Theta - (1 + 5\Gamma)j_{\min}^4 + \frac{\epsilon_{\rm GR} j_{\min}}{6}\right)^{-1/2}. \tag{E5}$$

We can make sense of Equation (E5) by evaluating the right-hand side in the weak and moderate GR regimes.

In the weak GR regime, we have $\epsilon_{\rm GR} j_{\min} \ll \Theta$ (Equation (A9)). If we also assume $j_{\min}^4 \ll \Theta$ (see Appendix C of Paper III for justification) and ignore the a-dependent terms in Equations (B1) and (B3), we get

$$\frac{\Delta a_{\rm RX18}}{\Delta a} \approx \frac{8}{15\xi} \approx \sqrt{1 + |\sigma|}, \tag{E6}$$

where to get the second equality we used Equation (42). Note that for $\sigma \ll 1$ we recover $\Delta a_{\rm RX18} = \Delta a$, i.e., our calculation coincides precisely with that of RX18 when we make the approximations that they (implicitly) did, namely weak GR and $\sigma \ll 1$. However, we emphasize that neither of these approximations is actually valid for the example shown in Figure 12.

In the moderate GR regime, we assume that the $\epsilon_{\rm GR}$ term dominates the final bracket in Equation (E5), and that the a-dependent terms dominate Equations (B1) and (B3). With these assumptions, we get

$$\frac{\Delta a_{\rm RX18}}{\Delta a} = \frac{8\sqrt{60\Gamma}}{15}\left(\frac{d}{a}\right)^{7/2}\frac{1}{\xi}\frac{j_{\min}^2}{\sqrt{j_{\min}^2 - \Theta}}\sqrt{\frac{6}{\epsilon_{\rm GR} j_{\min}}} \sim \frac{\sqrt{10\Gamma}}{\xi}\left(\frac{d}{a}\right)^{7/2}\left(\frac{j_{\min}}{\epsilon_{\rm GR}}\right)^{1/2}, \tag{E7}$$

with ξ given in Equation (49) (and plotted in Figure 3). All three fractions in Equation (E7) are $\mathcal{O}(1)$ or larger. Thus, we typically have $\Delta a_{\rm RX18}/\Delta a \gg 1$, meaning that the method of RX18 can seriously overestimate the value of Δa in the moderate GR regime.

E.2. Decrease in $t_{\rm sec}$ with Time

As we mentioned in Section 1, the decrease in $t_{\rm sec}$ with time during a slow merger was first pointed out by RX18 in their Section 3.1, when discussing the example shown in Figure 12. When interpreting this counterintuitive scaling of $t_{\rm sec}(a)$ physically, RX18 noted that smaller a (larger $\epsilon_{\rm GR}$) promotes

[19] It is easy to show that the condition of Equation (E4) is also required to make Equation (54) of RX18 agree with Equation (12) of Paper III at maximum eccentricity.

faster apsidal precession, which is obviously true. They then claimed that this faster precession directly leads to a shorter secular period. They also claimed that it was directly responsible for the corresponding increase in maximum eccentricity with time and decrease in minimum eccentricity with time as the binary shrinks (Figure 12(b)).

This interpretation is not quite right, and also does not explain why in the librating regime $t_{\rm sec}$ *increases* with shrinking a. In reality, in the weak-to-moderate regime, GR precession is unimportant except during an extremely high-eccentricity episode. Typically these extreme eccentricity episodes last a very short time compared to the secular period. In other words, for most phase-space trajectories, the second (GR) term in Equation (56) of RX18 is completely negligible during the majority of the evolution, so it barely affects $t_{\rm sec}$. What GR precession *does* do, when coupled with GW emission, is to alter the phase-space morphology, and to periodically nudge the binary onto a new phase-space trajectory every time it reaches high eccentricity (note how closely the contours of H^* are bunched at these high eccentricities in Figures 12(f)–(h)). As a is decreased and $\epsilon_{\rm GR}$ is increased, after passing from libration to circulation, the binary gets pushed ever farther away from the separatrix, toward the low-$j_{\rm max}$ circulating region where Equation (39) applies. As long as this process continues, the binary gets pushed to higher minimum eccentricity (smaller and smaller $j_{\rm max}$), even though its $e_{\rm max}$ is getting smaller. On average, the binary spends more and more time at "high" (say $e \gtrsim 0.9$) eccentricities where cluster tide-driven secular evolution is fast (since the binary angular momentum is small). We emphasize that this last statement is true regardless of GR precession: indeed, the binary typically does not care about GR precession directly when, say, $e = 0.9$. Thus, whereas RX18 attributed the evolution of $t_{\rm sec}$ and $e_{\rm min\,/\,max}$ to fast GR-aided ω precession during the whole secular cycle, both of these phenomena are present even in the weak GR regime where apsidal precession is nearly always negligible—see Figure 4.

ORCID iDs

Chris Hamilton https://orcid.org/0000-0002-5861-5687
Roman R. Rafikov https://orcid.org/0000-0002-0012-1609

References

Abbott, R., Abbott, T., Acernese, F., et al. 2021, arXiv:2111.03606
Antognini, J. M., Shappee, B. J., Thompson, T. A., & Amaro-Seoane, P. 2014, MNRAS, 439, 1079
Antonini, F., Hamers, A. S., & Lithwick, Y. 2016, AJ, 152, 174
Antonini, F., Murray, N., & Mikkola, S. 2014, ApJ, 781, 45
Antonini, F., & Perets, H. B. 2012, ApJ, 757, 27
Antonini, F., Toonen, S., & Hamers, A. S. 2017, ApJ, 841, 77
Arca Sedda, M. 2020, ApJ, 891, 47
Banerjee, S. 2018, MNRAS, 473, 909
Blaes, O., Lee, M. H., & Socrates, A. 2002, ApJ, 578, 775
Bode, J. N., & Wegg, C. 2014, MNRAS, 438, 573
Britt, D., Johanson, B., Wood, L., Miller, M. C., & Michaely, E. 2021, MNRAS, 505, 3844
Bub, M. W., & Petrovich, C. 2020, ApJ, 894, 15
Fabrycky, D., & Tremaine, S. 2007, ApJ, 669, 1298
Ford, E. B., & Rasio, F. A. 2006, ApJ, 638, L45
Grishin, E., Perets, H. B., & Fragione, G. 2018, MNRAS, 481, 4907
Hamilton, C., & Rafikov, R. R. 2019a, ApJL, 881, L13
Hamilton, C., & Rafikov, R. R. 2019b, MNRAS, 488, 5512
Hamilton, C., & Rafikov, R. R. 2019c, MNRAS, 488, 5512
Hamilton, C., & Rafikov, R. R. 2021, MNRAS, 505, 4151
Kozai, Y. 1962, AJ, 67, 591
Lidov, M. 1962, P&SS, 9, 719
Liu, B., & Lai, D. 2017, ApJL, 846, L11
Liu, B., & Lai, D. 2018, MNRAS, 483, 4060
Miller, M. C., & Hamilton, D. P. 2002, ApJ, 576, 894
Naoz, S. 2016, ARA&A, 54, 441
Peters, P. C. 1964, PhRv, 136, B1224
Randall, L., & Xianyu, Z.-Z. 2018, ApJ, 864, 134
Samsing, J., Hamers, A. S., & Tyles, J. G. 2019, PhRvD, 100, 043010
Silsbee, K., & Tremaine, S. 2017, ApJ, 836, 39
Thompson, T. A. 2011, ApJ, 741, 82
Wen, L. 2003, ApJ, 598, 419