Overspinning a Kerr black hole: The effect of the self-force

Marta Colleoni and Leor Barack

Mathematical Sciences, University of Southampton, Southampton SO17 1BJ, United Kingdom
(Received 29 January 2015; published 20 May 2015)

We study the scenario in which a massive particle is thrown into a rapidly rotating Kerr black hole in an attempt to spin it up beyond its extremal limit, challenging weak cosmic censorship. We work in black-hole perturbation theory, and focus on nonspinning, uncharged particles sent in on equatorial orbits. We first identify the complete parameter-space region in which overspinning occurs when backreaction effects from the particle's self-gravity are ignored. We find, in particular, that overspinning can be achieved only with particles sent in from infinity. Gravitational self-force effects may prevent overspinning by radiating away a sufficient amount of the particle's angular momentum ("dissipative effect"), and/or by increasing the effective centrifugal repulsion, so that particles with suitable parameters never get captured ("conservative effect"). We analyze the full effect of the self-force, thereby completing previous studies by Jacobson and Sotiriou (who neglected the self-force) and by Barausse, Cardoso and Khanna (who considered the dissipative effect on a subset of orbits). Our main result is an inequality, involving certain self-force quantities, which describes a necessary and sufficient condition for the overspinning scenario to be overruled. This "censorship" condition is formulated on a certain one-parameter family of geodesics in the limit of an extremal Kerr geometry. We find that the censorship condition is insensitive to the dissipative effect (within the first-order self-force approximation used here), except for a subset of perfectly fine-tuned orbits, for which a separate censorship condition is derived. We do not obtain here the self-force input needed to evaluate either of our two conditions, but discuss the prospects for producing the necessary data using state-of-the-art numerical codes.

I. INTRODUCTION

The cosmic censorship conjecture [1] has over the years become a cornerstone of classical general relativity. In its weak version it states, in essence, that curvature singularities arising in solutions to the Einstein's field equations must be cloaked behind event horizons, so that they are prevented from being in causal contact with distant observers. Despite being strongly motivated on physical grounds, the conjecture's precise extent of validity remains unclear. A notable counterexample involves finely tuned initial conditions [2]. The formulation of the conjecture may be refined to exclude such examples [3].

In a 1974 paper [4] Wald proposed a simple but powerful framework for testing weak cosmic censorship, using the gedanken experiment of a particle thrown into a Kerr-Newman black hole. If parameters can be chosen such that the postcapture mass M_f, charge Q_f and spin J_f satisfy $M_f^2 < (J_f/M_f)^2 + Q_f^2$, then a naked singularity would presumably form, in direct violation of weak censorship. Whether the equations of classical general relativity permit such a process has since been subject of much investigation. It is usually assumed that the particle's energy and electric charge are much smaller than those of the black hole, which then places the problem within the realm of black-hole perturbation theory.

In [4] Wald showed that the over-extremality scenario is ruled out when the configuration is that of a pointlike test particle captured by an extremal Kerr-Newman black hole. Electrostatic and centrifugal repulsion, he showed, would prevent a particle carrying sufficient charge and/or angular momentum from entering the black hole. The same conclusion was shown to hold true also for a spinning test particle dropped from rest at infinity along the symmetry axis of an extremal Kerr black hole, with its spin aligned along the axis. In this case, it is the repulsion force from spin-spin coupling that prevents suitable particles from ever entering the black hole.

However, later work has demonstrated that over extremality is achievable when the initial black hole is taken to be *nearly* extremal—if backreaction effects on the particle's trajectory are ignored. This was first shown by Hubeny [5] for a nearly extremal Reissner-Nordström black hole, and more recently by Jacobson and Sotiriou [6] for a nearly extremal Kerr black hole ("overcharging" and "overspinning" scenarios, respectively). The nearly extremal Kerr-Newman case was subsequently studied in Ref. [7]. In all cases, all orbits identified as capable of driving the black hole beyond the extremal limit lie very close, in the relevant parameter space, to the separatrix between orbits that are captured by the black hole and ones that are scattered off it. In Hubeny's analysis of a radially falling electric charge, electrostatic repulsion only marginally fails to prevent the particle from falling into the hole: The particle's radial velocity upon crossing (what would have been) the event horizon is proportional to the ratio $\tilde{\eta} \ll 1$ between the

particle's energy and the black hole's mass. The amount of postcapture excess charge, $Q_f - M_f$, is found to be *quadratic* in $\tilde{\eta}$. Similarly, in Ref. [6]'s analysis of equatorial-plane captures, overspinning particles clear the peak of the effective potential barrier with radial velocities $\propto \tilde{\eta}$, and the postcapture excess spin, $J_f - M_f^2$, is quadratic in $\tilde{\eta}$.

This suggests strongly that backreaction effects cannot be ignored and may well change the outcome of the gedanken experiment. Heuristically, effects of the (electromagnetic and/or gravitational) self-force enter the analysis in two ways. First, the *dissipative* piece of the self-force continually removes some of the particle's energy and angular momentum, sending them to infinity and down the event horizon in gravitational waves. In the Kerr case, dissipative effects may accumulate as the particle "hovers" above the peak of the effective potential on a nearly circular orbit. Second, the *conservative* piece of the self-force might supply just the right amount of additional repulsive force to prevent would-be overcharging/overspinning particles from ever entering the black hole. For particles sent in from infinity in the Kerr case, this second effect may be formulated in terms of a shift in the critical impact parameter for capture: If the gravitational self-force (GSF) shifts the critical impact parameter inward by a sufficient amount (for a given energy at infinity), then would-be overspinning particles may end up being scattered away rather than captured.

There have been several recent attempts to quantify the effect of backreaction in the problem. Focusing on the Reissner-Nordström case, Isoyama, Sago and Tanaka [8] argued that the full effect can be properly taken into account by considering the quasiequilibrium configuration of a charged particle placed precisely on the capture-scatter separatrix. An exact solution is known for this configuration—the static double Reissner-Nordström spacetime—and the authors calculated that its total energy is always greater than its total charge. They have also established that radiative losses during the final plunge are negligible, hence concluding that (under the assumption that the true capture system does indeed go through a quasiequilibrium state) the final configuration cannot be a naked singularity.

In a later work, Zimmerman, Vega and Poisson [9] took up the challenge of directly calculating the charged particle's trajectory including the full effect of the electromagnetic self-force. Analyzing numerically a large sample of orbits within the domain identified by Hubeny, the authors found no example of successful overcharging: All particles with a combination of charge and energy suitable for overcharging the black hole were found to be repelled before reaching the horizon. This analysis, however, neglected the potentially important effect of backreaction from the gravitational perturbation sourced by the particle's electromagnetic energy-momentum. A complete analysis would require calculation of the corresponding GSF, but

techniques for calculating self-forces in the coupled problem are only now starting to be developed [10,11].

In that respect, the Kerr setup provides a cleaner environment, in which the perturbative problem is purely gravitational (at the obvious cost of abandoning spherical symmetry). Barausse, Cardoso and Khanna [12,13] studied the dissipative GSF effect in the Kerr overspinning problem, focussing on ultrarelativistic particles on equatorial orbits. Using analytic arguments backed by a numerical calculation of the energy and angular momentum carried away in gravitational waves, they showed that dissipation averts the overspinning for some but not all of Jacobson-Sotiriou's orbits. For sufficiently small $\tilde{\eta}$, the dissipative effect is always negligible and cannot prevent overspinning. This result highlights the importance of accounting for the *full* effect of GSF. To reach a definitive conclusion necessitates an actual calculation of the full local GSF acting on the captured particles.

In the past few years, rigorous methods for GSF calculations in Kerr spacetime have advanced enough to allow a more systematic and complete treatment of the overspinning problem. The program initiated with this paper revisits the problem from this new vantage point. It seeks to obtain a more conclusive answer to the question of whether it is indeed the self-force that provides the mechanism by which black holes protect themselves from being overspun.

Our current paper lays the necessary groundwork. Concentrating on equatorial orbits, we first identify the complete "window" in the parameter space in which overspinning occurs if the GSF is ignored. We then formulate a condition for this window to be eliminated by the effect of the full GSF. The condition takes the form of an inequality that is required to hold for each member of a certain one-parameter family of geodesics, and it involves the GSF calculated along such orbits. Here we do not obtain the necessary GSF data, but we discuss methods for computing it (numerically) using existing codes. With collaborators we have began work to obtain the GSF data, and we intend to present the results in a follow-up paper.

The rest of this Introduction summarizes our analysis (also in relation to previous work) and describes its main results.

A. This work: Overview and results

Jacobson and Sotiriou [6] assumed that overspinning occurs if two conditions are met: (i) the geodesic trajectory of the test particle is timelike at the horizon, and (ii) $J_f > M_f^2$. The first condition is very lax. It allows for low-energy orbits that are deeply bound to the black hole and confined to the immediate neighborhood of the horizon. The physics of such orbits becomes very subtle, especially when self-gravity and finite-size effects are included [consider that deeply bound orbits below the innermost stable circular orbit (ISCO) plunge into the black

hole within an amount of proper time that shrinks to zero in the extremal limit [14]], so one would preferably avoid such orbits as candidates (see, however, Hod [15] for a heuristic treatment). Jacobson and Sotiriou acknowledge this issue, and to address it they supplement their analysis with two specific numerical examples of overspinning orbits that are sent in from afar. They stop short of determining the full range of overspinning orbits when deeply bound orbits are disallowed.

Our capture condition will be more stringent, and more in the spirit of Ref. [9]: We will send our particle in from "sufficiently far" (this condition will be made precise in the next section), and deem it "captured" if it has no inner radial turning point outside the black hole. Thus, for a legitimate capture we demand that the particle "clears" the peak of the effective potential on its inward journey.

In Sec. II we revisit the overspinning problem in the geodesic approximation. We identify the precise region in the parameter space of equatorial orbits around a nearly extremal Kerr black hole (excluding deeply bound orbits) where overspinning occurs, and give analytic expressions for the boundaries of that region. The overspinning window is illustrated in Figs. 2 and 3. Perhaps unexpectedly, we find that only particles sent in from infinity are capable of overspinning the black hole. This fact has somehow gone unnoticed in previous work, to the best of our knowledge. We find that for any given value of the particle's energy at infinity, there exists an open range of orbital angular momenta and particle's rest masses for which overspinning occurs. That only orbits coming from infinity are potential overspinners is somewhat fortuitous, because for such orbits it is straightforward to identify the system's total [Arnowitt-Deser-Misner (ADM)] energy and angular momentum even when GSF effects are included.

We then turn to analyze the GSF effect. In Sec. III we first review essential results from GSF theory, and then discuss the determination of the "critical orbit" that separates (in the relevant parameter space) between plunging and scattered orbits. We do this in two steps. First, we ignore the dissipative piece of the GSF, and calculate the correction due to the conservative GSF to the critical value of the angular momentum for a fixed value of the energy at infinity. Then we restore dissipation and consider its consequences. Under the full GSF, all critical orbits merge into a "global attractor" that takes the system adiabatically along a sequence of quasicircular unstable orbits ending at the ISCO, before plunging into the black hole. By fine-tuning the initial value of the angular momentum (for a given initial energy), an orbit can be made to evolve arbitrary far along the global attractor. We make a formal distinction between "generic" and "fine-tuned" captured orbits, based on how the difference between the initial and final values of the particle's specific energy scales with the particle's mass μ (in a procedure whereby μ is taken to zero while holding fixed the initial specific energy and angular

momentum). *Generic* orbits are ones for which that difference vanishes for $\mu \to 0$ (this includes, for example, all of the orbits considered in Refs. [12,13]); *fine-tuned* orbits are ones for which the difference does not vanish even for $\mu \to 0$.

With this preparatory work in place, we move on, in Sec. IV, to obtain the overspinning condition as modified by the full GSF. The end results are two inequalities, one for generic orbits [Eq. (69)] and another for fine-tuned ones [Eq. (81)], which describe conditions for overspinning to be averted under the effect of the full GSF. In the generic case, the condition involves only the conservative piece of the GSF, evaluated along critical geodesics in the extremal Kerr limit. Overspinning can be ruled out if and only if the condition is met for each member of this one-parameter family. The condition for fine-tuned orbits requires, in addition to the conservative GSF, also knowledge of the fluxes of energy and angular momentum radiated to infinity by particles on unstable circular orbits, in the extremal Kerr limit. Overspinning can be ruled out if and only if the condition is met for any values of the initial and final energies.

In Sec. V we propose an alternative form of the over-spinning conditions, based on the framework of the "first law of binary black-hole mechanics," as recently applied to orbits in Kerr [16]. The alternative form, given (for the generic case) in Eq. (104), involves only perturbative quantities calculated along (unstable) circular orbits, which should be more easily computable with existing GSF codes. This simplified formulation relies on explicit expressions, given in [16], for the ADM-like energy and angular momentum of circular-orbit configurations in Kerr, includ-ing leading-order self-interaction terms. The underlying theoretical framework is yet to be firmly established and tested, however, so we regard the simplified condition (104) as somewhat less rigorous than our direct condition (69). Our stance is that it would be desirable to evaluate both forms of the condition, for the sake of establishing confidence in the result.

In Sec. VI we discuss the numerical input required for evaluating our overspinning conditions, and the prospects for obtaining it through adaptation of existing codes. Evaluation of the direct conditions (69) and (81) involves GSF calculations along unbound orbits on a nearly extremal Kerr background, which has not been attempted so far. However, we think the basic computational infra-structure for such calculations is well in place.

Section VII summarizes our results and speculates on what a numerical evaluation of our censorship conditions might yield. The Appendices contain some of the details of calculations done in Secs. III and IV: a derivation of the ADM energy and angular momentum for the system under consideration, a calculation of the GSF-induced shift in the critical value of the angular momentum, and an evaluation

of radiative losses during the final plunge into the black hole.

Throughout this paper we set $G = c = 1$ and use the metric signature $(-, +, +, +)$.

II. OVERSPINNING ORBITS IN THE GEODESIC APPROXIMATION

Our initial configuration features a Kerr black hole of mass M and angular momentum $J = aM < M^2$. A point-like test particle of rest mass $\mu \ll M$ is sent in on a geodesic of the background Kerr geometry. As in [6], we restrict attention to prograde orbits in the equatorial plane, so that the orbital angular momentum is aligned with the spin of the black hole. (Intuitively, this configuration seems most favorable for a successful overspinning.) We denote the particle's *specific* energy and angular momentum by E and L, respectively; these are constants of the geodesic motion. For the geodesic approximation to make sense, we must assume $\mu E \ll M$ and $\mu L \ll J$. Then, clearly, overspinning could only be possible, in principle, if the black hole is nearly extremal. We write

$$a/M = 1 - \epsilon^2, \tag{1}$$

where $\epsilon \ll 1$.[1]

Below we study the overspinning scenario in the above setup, but we begin with a survey of some essential properties of timelike equatorial geodesics of the Kerr metric.

A. Relevant results for Kerr geodesics

Let u^α denote the particle's four-velocity. In Boyer-Lindquist coordinates $\{t, r, \theta, \phi\}$ we have $u^\theta \equiv 0$, and

$$\dot{u}_t = 0, \qquad \dot{u}_\phi = 0, \tag{2}$$

where an overdot denotes differentiation with respect to proper time. The two equalities express the conservation of energy $E = -\xi^\alpha_{(t)} u_\alpha = -u_t$ and angular momentum $L = \xi^\alpha_{(\phi)} u_\alpha = u_\phi$, where $\xi^\alpha_{(t)} := \partial^\alpha_t$ and $\xi^\alpha_{(\phi)} := \partial^\alpha_\phi$ are Killing vectors associated with the time translation and rotational symmetries of the Kerr background. The pair $\{E, L\}$ parametrizes the family of equatorial geodesics (up to initial conditions).

The normalization $u_\alpha u^\alpha = -1$ now gives the radial equation of motion, which we write in the form

$$\dot{r}^2 = B(r)[E - V_-(L, r)][E - V_+(L, r)]. \tag{3}$$

Here r is the Boyer-Lindquist radius of the orbit, $B(r) := 1 + a^2(r + 2M)/r^3$, and (for $MaL \neq 0$)

¹Note, to avoid confusion, that [6] has instead $a/M = 1 - 2\epsilon^2$.

$$V_\pm(L, r) := \frac{2MaL}{Br^3}\left(1 \pm \sqrt{1 + \frac{Br^3[L^2(r - 2M) + r\Delta]}{4M^2 a^2 L^2}}\right), \tag{4}$$

with $\Delta := r^2 - 2Mr + a^2$. For prograde orbits, the potential V_- is manifestly negative definite, so the factor $B(r)(E - V_-)$ in Eq. (3) is manifestly *positive* definite. Thus, V_+ plays the role of an effective potential for the radial motion, which is allowed for $E \geq V_+(L, r)$, with an equality identifying radial turning points.

Stationary points of $V_+(L, r)$ outside the black hole, when they exist, correspond to circular orbits. These satisfy the simultaneous conditions

$$E = V_+, \qquad \partial_r V_+ = 0 \quad \text{(circular orbits)}. \tag{5}$$

Substituting from Eq. (4) and solving for E and L in terms of the circular-orbit radius, $r = R$, gives $E = E_c(R)$ and $L = L_c(R)$, with

$$E_c(R) = \frac{1 - 2\tilde{R}^{-1} + \tilde{a}\tilde{R}^{-3/2}}{\sqrt{1 - 3\tilde{R}^{-1} + 2\tilde{a}\tilde{R}^{-3/2}}}, \tag{6}$$

$$\tilde{L}_c(R) = \frac{\tilde{R}^{1/2}(1 - 2\tilde{a}\tilde{R}^{-3/2} + \tilde{a}^2\tilde{R}^{-2})}{\sqrt{1 - 3\tilde{R}^{-1} + 2\tilde{a}\tilde{R}^{-3/2}}}. \tag{7}$$

Here an overtilde denotes adimensionalization using M, i.e., $\tilde{R} := R/M$, $\tilde{a} := a/M$ and $\tilde{L} := L/M$; we shall adopt this notation throughout the rest of the paper. Timelike circular orbits exist only for $R > R_{\text{ph}}(a)$, the radius of a photon's unstable circular orbit ("light ring"). $R_{\text{ph}}(a)$ is the (unique) root of $1 - 3\tilde{R}^{-1} + 2\tilde{a}\tilde{R}^{-3/2}$ greater than the event horizon's radius, $\tilde{R}_{\text{eh}}(a) = 1 + (1 - \tilde{a}^2)^{1/2}$. The angular velocity $\Omega := u^\phi/u^t$ of any circular geodesic orbit reads

$$\tilde{\Omega}(R) = (\tilde{a} + \tilde{R}^{3/2})^{-1}. \tag{8}$$

The number of stationary points of V_+ and their location depend on L. There are none outside the black hole when L is below a certain critical value $L_{\text{isco}}(a)$, and there are two for $L > L_{\text{isco}}(a)$: a maximum representing an unstable circular orbit, and, further out, a minimum representing a stable one. The critical value $L_{\text{isco}}(a)$ marks the ISCO. It is given by $L_{\text{isco}} = L_c(R_{\text{isco}})$, where the ISCO radius R_{isco} is found by solving Eq. (5) simultaneously with $\partial_r^2 V_+(L, r) = 0$. The ISCO may also be said to represent the outer boundary of the region of unstable circular orbits.

The radii of unstable circular geodesic orbits span the interval $R_{\text{ph}}(a) < R < R_{\text{isco}}(a)$. This one-parameter family of orbits will feature dominantly in our analysis, because it defines the capture-scatter threshold where much of the relevant physics occurs. Members of the family may be parametrized by either E or L, both being monotonically
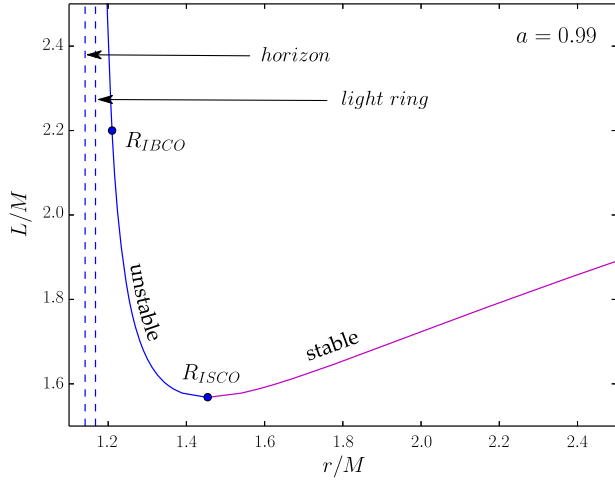
FIG. 1 (color online). Timelike circular equatorial geodesics around a nearly extremal Kerr black hole, shown here for $a = 0.99M$. The plot shows specific angular momentum versus Boyer-Lindquist radius. Orbits with $r > R_{isco}$ (magenta) are stable, while these with $r < R_{isco}$ (blue) are unstable. Also indicated are the IBCO ($E = 1$) and the photon orbit (light ring, $E, L \to \infty$). In the extremal limit, $a \to M$, the radii R_{isco}, R_{ibco} and R_{ph} all coincide with the horizon radius R_{eh}.

decreasing functions of R between R_{ph} (where $E, L \to \infty$) and R_{isco} for any $\tilde{a} < 1$. This monotonicity can be readily established from Eqs. (6) and (7). Hence, the radius R itself is also a valid parameter.

To each unstable circular orbit there correspond non-circular homoclinic-like geodesic orbits [17] that join the circular orbit asymptotically in either their infinite past or their infinite future, or both. Nearly homoclinic orbits exhibit a "zoom-whirl" behavior [18]: an episode of prolonged rotation ("whirl") about the location of the associated unstable circular orbit. We will see that all orbits relevant to the overspinning problem fall in that category. Based on the correspondence with homoclinic orbits, unstable circular orbits may be divided into "bound" ($E < 1$) and "unbound" ($E \geq 1$). The radius of the inner-most bound circular orbit (IBCO) is obtained by solving $E_c(R) = 1$, giving $\tilde{R}_{ibco} = [1 + (1 - \tilde{a})^{1/2}]^2$.

Figure 1 illustrates the range of stable and unstable circular orbits, and the location of the various special orbits mentioned, in a particular example ($\tilde{a} = 0.99$). We note the ordering

$$R_{eh} < R_{ph} < R_{ibco} < R_{isco}, \qquad (9)$$

which applies for any $\tilde{a} < 1$.

Let us now specialize to a near-extremal Kerr background with spin as in Eq. (1). One finds

$$\tilde{R}_{eh} = 1 + \sqrt{2}\epsilon + O(\epsilon^2), \qquad (10)$$

$$\tilde{R}_{ph} = 1 + \sqrt{8/3}\epsilon + O(\epsilon^2), \qquad (11)$$

$$\tilde{R}_{ibco} = (1 + \epsilon)^2 \quad \text{(exact)}, \qquad (12)$$

$$\tilde{R}_{isco} = 1 + (2\epsilon)^{2/3} + O(\epsilon^{4/3}). \qquad (13)$$

The function $E_c(R)$ in Eq. (6) can be inverted perturbatively in ϵ to obtain the radius of an arbitrary unstable circular orbit in terms of its energy E. We find

$$\tilde{R} = 1 + \epsilon\rho_1(E) + \epsilon^2\rho_2(E) + O(\epsilon^3), \qquad (14)$$

where the first two coefficients, needed below, read

$$\rho_1 = \frac{2\sqrt{2}E}{\sqrt{3E^2 - 1}}, \qquad \rho_2 = \frac{2(2E^4 - E^2 + 1)}{(3E^2 - 1)^2}. \qquad (15)$$

Equation (12) is the special case of (14) with $E = 1$, giving $\rho_1 = 2$ and $\rho_2 = 1$.

It follows that, in the extremal limit $\epsilon \to 0$, the Boyer-Lindquist radii of the light ring and the ISCO both coincide with the horizon radius, and so do the radii of all unstable circular orbits enclosed between them. Also peculiar is the fact that the ratio of coordinate differences $(\tilde{R}_{isco} - \tilde{R}_{eh})/(\tilde{R}_{ibco} - \tilde{R}_{eh})$ diverges as $\epsilon \to 0$. A closer look reveals [19] that the light ring, IBCO and ISCO remain separated from the horizon, and from each other, when examined on a Boyer-Lindquist $t = $ const slice: On that slice, the proper radial distance between the light ring and the horizon is finite, and so is the distance between any fixed-E unstable circular orbit and the light ring. The proper radial distance between the ISCO and any fixed-E unstable circular orbit *diverges* on the $t = $ const slice; the geometry of the $t = $ const hypersurface appears to "stretch" infinitely around the ISCO location [19]. The situation, however, is rather different when examined on a horizon-crossing time slice. As emphasized recently by Jacobson [14], on any such slice, the light ring, IBCO and ISCO all actually coincide with horizon generators. From that perspective, they—and all unstable orbits in between them—are "at the same place" in the extremal limit.

These subtleties will not affect our analysis directly: ϵ will be kept (arbitrarily) small but nonzero, and the strict ordering (9) will therefore apply on any time slice. However, we must take note of the degeneracy of R as a parameter for unstable circular orbits when $\epsilon \to 0$. The energy E, on the other hand, remains a good parameter even in this limit, spanning the entire range $\infty > E > \frac{1}{\sqrt{3}}$. We will thus generally adopt E for labeling unstable circular orbits. Given E, the angular momentum $L_c(R(E))$, which we henceforth write as $L_c(E)$, is obtained by substituting Eqs. (1) and (14) in Eq. (7) and then expanding in ϵ. The result is

$$\tilde{L}_c(E) = 2E + (6E^2 - 2)^{1/2}\epsilon + O(\epsilon^2). \qquad (16)$$

We note that to determine the $O(\epsilon)$ term here required the explicit values of both ρ_1 and ρ_2 of Eq. (15).

B. Exclusion of deeply bound orbits

Heuristically, if we assume our point particle represents a compact object—say, a Schwarzschild black hole—then its effective proper "diameter" is $\sim\mu$. Below it will become clear that a successful overspinning requires $\mu \sim \epsilon$, and so relevant objects have proper diameters $\sim\epsilon$. Now consider placing such an object in a deeply bound orbit with an *outer* turning point at $r < R_{\text{isco}}$ [and with $L > L_c(E)$]. Such an object (it can be checked) will plunge through the horizon within a proper time of $O(\epsilon)$ (at most), comparable to its own "light-crossing time." It is not clear whether the object can be made to initially "fit" in its entirety outside the hole. At the very least, it is not clear if the simple model of a point particle and a stationary horizon provides a faithful description of the physics in this case.

To avoid such subtleties, we wish to exclude deeply bound orbits from our analysis. We achieve this by requiring that, if the orbit possesses an outer radial turning point at some $r = r_{\text{out}}$, then

$$r_{\text{out}} > R_{\text{isco}}(\epsilon). \qquad (17)$$

It can be checked that, under this condition, the proper-time interval along any timelike equatorial geodesic connecting $r = r_{\text{out}}$ to $r = R_{\text{eh}}$ is finite (nonzero) even in the limit $\epsilon \to 0$ (taken with fixed E, L). The condition (17) demands that eligible particles must clear the peak of the effective potential (when such a peak exists) as they plunge into the black hole.

C. Overspinning domain

Given the restriction (17), a necessary and sufficient condition for a falling particle of specific energy E to be captured by the black hole is

$$L < L_c(E). \qquad (18)$$

A captured particle would overspin the black hole if and only if

$$(M + \mu E)^2 < aM + \mu L. \qquad (19)$$

Using $\tilde{a} = 1 - \epsilon^2$ and introducing the small mass ratio $\eta := \mu/M$, this condition becomes

$$\epsilon^2 + \eta W + \eta^2 E^2 < 0, \qquad (20)$$

where we have introduced[2]

---

[2]Heuristically, $W/2$ may be interpreted as the specific energy in a corotating frame with $\tilde{\Omega} = 1/2$, i.e., the common angular velocity of all unstable circular geodesics in the extremal limit.

$$W := 2E - \tilde{L}. \qquad (21)$$

Note that Eq. (20) sets a lower bound on L (for given E, η, ϵ), while Eq. (18) sets an upper bound. Also note that Eq. (20) implies the necessary condition $W < 0$ for over-spinning to occur.

Our goal now is to identify the complete domain in the space of $\{\eta, E, L\}$ for which the conditions (18) and (20) are simultaneously satisfied, assuming $\epsilon \ll 1$ and the condition (17). For easy reference, let us call this domain "OS," for "overspinning".

We first show that orbits with $L \leq L_{\text{isco}}$ all fall *outside* OS. To this end, consider first the ISCO itself, where $W = 2E_{\text{isco}} - \tilde{L}_{\text{isco}} =: W_{\text{isco}}$. Using Eqs. (6), (7) and (13) we obtain $W_{\text{isco}} = -\hat{c}\epsilon^{4/3} + O(\epsilon^2)$, where $\hat{c} = 2^{1/3}\sqrt{3} > 0$. Thus, W_{isco} is negative as required, but it can be easily checked that (20) is always violated for sufficiently small ϵ: Replacing $W \to -\hat{c}\epsilon^{4/3}$ in Eq. (20) and considering the left-hand side as a quadratic function of η, we find this function is positive definite for any $\epsilon < (2E_{\text{isco}}/\hat{c})^3$. [Since E_{isco} is bounded from below by $E_{\text{isco}}(\epsilon = 0) = \frac{1}{\sqrt{3}}$, we find that (20) is always violated for $\epsilon < \frac{4}{27}$.] This rules out the ISCO itself, and it immediately rules out also all orbits with $\{E > E_{\text{isco}}, L = L_{\text{isco}}\}$, for which $W > W_{\text{isco}}$. Orbits with $\{E < E_{\text{isco}}, L_{\text{isco}}\}$ *can* potentially satisfy Eq. (20), but they are always deeply bound in the sense of failing to satisfy Eq. (17): For any $E < E_{\text{isco}}$, the orbit has an outer radial turning point at a radius $r_{\text{out}} < R_{\text{isco}}$.

The upshot is that orbits with $L = L_{\text{isco}}$ are all outside OS. For orbits with $L < L_{\text{isco}}$ we would need to require $E < E_{\text{isco}}$ in order for W to be sufficiently negative. But, once again, such orbits are excluded on account of their being deeply bound. We conclude that orbits with $L \leq L_{\text{isco}}$ are all outside OS.

Let us focus therefore on orbits with $L > L_{\text{isco}}$. For such an orbit to be in OS, we require that (given E, η, ϵ) L is bounded from above by $L_c(E)$ and simultaneously from below via Eq. (20):

$$\epsilon^2 + 2\eta E + \eta^2 E^2 < \eta\tilde{L} < \eta\tilde{L}_c(E; \epsilon). \qquad (22)$$

We have made here the ϵ dependence of L_c explicit, for clarity. The span of the permissible range is $\eta\Delta_L := -\epsilon^2 - \eta[2E - \tilde{L}_c(E; \epsilon)] - \eta^2 E^2$, or, using Eq. (16),

$$\eta\Delta_L = -\epsilon^2 + \eta\epsilon\sqrt{6E^2 - 2} - \eta^2 E^2, \qquad (23)$$

where we have omitted terms of $O(\eta\epsilon^2)$. OS is populated if and only if we can find E, η, ϵ for which $\Delta_L > 0$.

A few conclusions can be drawn immediately. First, considering $\eta\Delta_L$ in Eq. (23) as a quadratic function of η, we find it has a maximum value

$$\max_{\eta}\eta\Delta_L = \frac{\epsilon^2(E^2-1)}{2E^2}. \tag{24}$$

This is positive only for $E > 1$. Therefore, all orbits with $E \leq 1$ fall outside OS. Bound orbits cannot overspin.

Second, for any $E > 1$, we can obtain $\Delta_L > 0$ by choosing the mass ratio η from within the interval

$$\epsilon\eta_-(E) < \eta < \epsilon\eta_+(E), \tag{25}$$

where

$$\eta_\pm = \frac{1}{\sqrt{2}E^2}\left[\sqrt{3E^2-1} \pm \sqrt{E^2-1}\right]. \tag{26}$$

In other words, overspinning can be achieved for any $E > 1$, as long as η satisfies (25). Since the condition $\Delta_L > 0$ is both necessary and sufficient, the converse also holds: All orbits in OS satisfy $E > 1$ with Eq. (25).

Third, from Eq. (25) it follows that η must be chosen to be of $O(\epsilon)$ (assuming $E \ll 1/\epsilon$). One can check that η_+ has a maximal value of

$$\max_{E}\eta_+ = \sqrt{3/2}, \tag{27}$$

obtained for $E = 2/\sqrt{3}$. Therefore, the range $\eta \geq \sqrt{3/2}\epsilon$ lies outside OS. The bandwidth of admissible mass ratios, for given E and ϵ, is

$$\Delta_\eta := \epsilon\eta_+ - \epsilon\eta_- = \epsilon\sqrt{2(E^2-1)}/E^2, \tag{28}$$

which is maximal for $E = \sqrt{2}$. Figure 2 depicts the permissible range of η/ϵ as a function of E.

Fourth, from Eqs. (24) and (28) we learn that an $E = \mathrm{const}(> 1)$ slice of OS has maximal dimensions $\Delta_L \sim \epsilon^2/\eta \sim \epsilon$ and $\Delta_\eta \sim \epsilon$. OS is thus a narrow "tube" in the $\{E, L, \eta\}$ parameter space, of a cross section $\sim\epsilon \times \epsilon$, whose boundary is tangent to the surface of unstable circular orbits, $L = L_c(E)$.

To summarize, we have found that OS is a narrow tubelike region of the $\{E, L, \eta\}$ space, described by $E > 1$, $L_c(E; \epsilon) - \Delta_L(E, \eta; \epsilon) < L < L_c(E; \epsilon)$ and $\epsilon\eta_-(E) < \eta < \epsilon\eta_+(E)$, where Δ_L and $\eta_\pm$ are given in Eqs. (23) and (26), respectively. A neater description of the OS window is obtained in terms of the quantity W defined in Eq. (21): Rearranging Eq. (22) and using (16), we find

$$\epsilon W_-(E) < W < \epsilon W_+(E, \eta/\epsilon), \tag{29}$$

where

$$W_- = -\sqrt{6E^2-2}, \qquad W_+ = -\left(\frac{\epsilon}{\eta} + \frac{\eta}{\epsilon}E^2\right). \tag{30}$$

This domain is illustrated in Fig. 3 for a sample of η/ϵ values. To overspin a black hole of given M and $\epsilon \ll 1$, one should pick an E greater than 1, choose any η from within the interval (25), and then choose W (hence L) from within the interval (29).
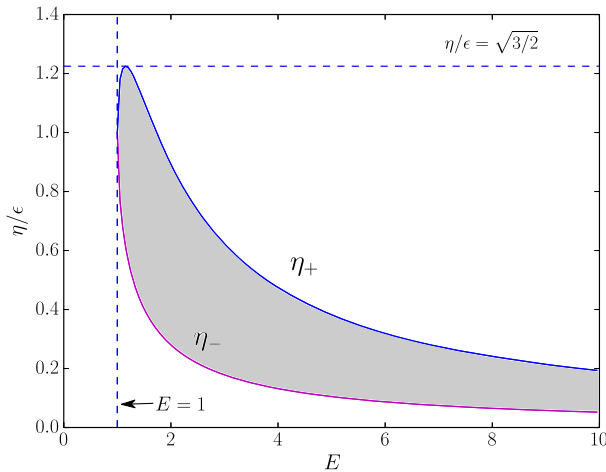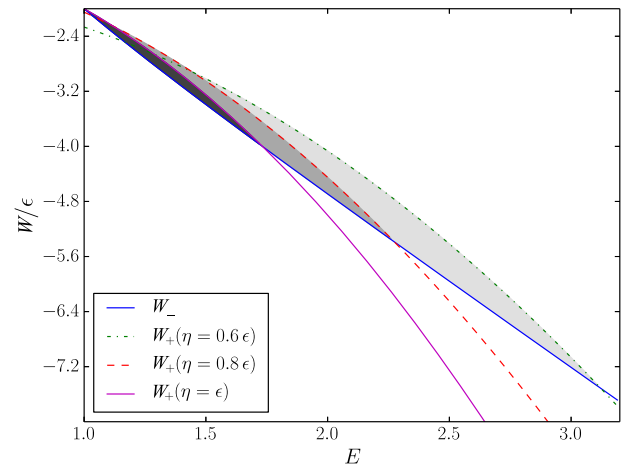

FIG. 2 (color online). Domain of mass ratios η for which overspinning is possible in the geodesic approximation. η is shown divided by the near-extremality parameter $\epsilon = (1 - a/M)^{1/2}$, and E is the particle's specific energy. The boundaries $\eta_\pm(E)$ are given in Eq. (26). Overspinning is not possible for $E < 1$ or $\eta > \sqrt{3/2}\epsilon$. However, for any value $E > 1$ there is a range of η for which the black hole may be overspun. This happens if the particle's angular momentum is chosen from within the range indicated in Eq. (29).

FIG. 3 (color online). The overspinning window, shown in the plane of E, W (where $W = 2E - L/M$) for several values of η/ϵ. Note W is shown divided by ϵ. The boundaries $W_\pm$ are given in Eq. (30). The lower boundary $W_-(E)$ (which does not depend on η) arises from the requirement that the particle is captured by the black hole. The upper boundary $W_+(E, \eta/\epsilon)$ comes from the requirement that the final object is an over-extremal black hole. Overspinning is possible with any $E > 1$, provided η is chosen from within the range shown in Eq. (25).

III. SELF-FORCE PRELIMINARIES

Because the width of the overspinning window is of $O(\eta)$, self-gravity effects may potentially close this window, and they must therefore be included in the analysis. Specifically, the GSF modifies the capture condition (18) by changing the functional relation $L_c(E)$ at $O(\eta)$. It also modifies the overspinning condition (19) by dissipating away some of the system's initial energy and angular momentum. In this section we introduce relevant results from the theory of self-forced motion. In Sec. IV we will then use these results to derive conditions for capture and overspinning under the full-GSF effect.

A. Equation of motion with self-force

There now exists a rigorous formulation of the equations of motion for compact objects in curved spacetime, valid through first postgeodesic order in perturbation theory—see [20,20–22] and references therein, and [23,24] for recent reviews. The formulation applies in situations where all length scales associated with the compact object are much smaller than the typical curvature radius of the background geometry. The motion of the compact object is then determined via a systematic procedure of matched asymptotic expansions, and interpreted as an accelerated motion in the background spacetime, subject to an effective GSF ($\propto \eta^2$). One of the results is that the internal structure of the object does not affect the self-acceleration at $O(\eta)$ (except, if the particle is spinning, through the familiar Mathisson-Papapetrou spin term).

The GSF formalism should be applicable in our setup, since we work under the assumption $\mu E \ll M$. The introduction of the small background-related parameter ϵ should not pose a problem, because the background's curvature radius remains much larger than μE even in the limit $\epsilon \to 0$, and even as the particle approaches the horizon. We will indeed proceed under the assumption that the standard first-order GSF formalism is applicable anywhere along the particle's trajectory until it crosses the horizon.

The equation of motion, including the leading-order GSF, may be written in the form

$$\mu \hat{u}^\beta \nabla_\beta \hat{u}^\alpha = F^\alpha. \tag{31}$$

Here $\hat{u}^\alpha$ is the particle's four-velocity, tangent to the (accelerated) trajectory in the background spacetime (Kerr, in our case) and normalized using

$$g_{\alpha\beta} \hat{u}^\alpha \hat{u}^\beta = -1, \tag{32}$$

where $g_{\alpha\beta}$ is the background (Kerr) metric. The covariant derivative in (31) is taken with respect to $g_{\alpha\beta}$, and F^α is the first-order GSF, proportional to μ^2. The GSF is normal to the four-velocity, $g_{\alpha\beta} \hat{u}^\alpha F^\beta = 0$, so that the rest mass μ remains constant. Methods to compute F^α in Kerr

spacetime are reviewed in [25]. In should be noted that F^α itself is a gauge-dependent notion: A full, gauge-invariant information about the motion is contained only in the combination of the GSF and the metric perturbation with which it is associated [26].

Now consider a particle sent in along the equator of the Kerr black hole, i.e. with $\theta = \pi/2$ and $\hat{u}^\theta = 0$ at the initial moment. In any reasonable gauge, the component F^θ would vanish from symmetry and the motion will remain equatorial. Let us then define

$$\hat{E} := -\hat{u}_t, \qquad \hat{L} := \hat{u}_\phi, \tag{33}$$

in analogy with E and L of the geodesic case. Here, however, $\hat{E}$ and $\hat{L}$ are not constants of the motion. Rather, Eq. (31) tells us they evolve (slowly) according to

$$\mu \frac{d\hat{E}}{d\tau} = -F_t(\tau), \qquad \mu \frac{d\hat{L}}{d\tau} = F_\phi(\tau), \tag{34}$$

where $F_\alpha = g_{\alpha\beta} F^\beta$ and τ is proper time along the self-accelerated orbit. With these definitions, Eq. (32) produces the radial equation of motion

$$\dot{r}^2 = B(r)(\hat{E} - V_-(\hat{L}, r))(\hat{E} - V_+(\hat{L}, r)), \tag{35}$$

whose form is identical to that of Eq. (3)—except that here $\hat{E}$ and $\hat{L}$ are slow functions of τ along the orbit.

The results of the previous section lead us to focus attention on particles sent in from infinity, i.e., ones with $r(\tau \to -\infty) \to \infty$. For such particles, we define

$$E_\infty := \hat{E}(\tau \to -\infty), \qquad L_\infty := \hat{L}(\tau \to -\infty). \tag{36}$$

From Eq. (34) we have

$$\hat{E}(\tau) = E_\infty + \Delta E(\tau), \qquad \hat{L}(\tau) = L_\infty + \Delta L(\tau), \tag{37}$$

where

$$\mu \Delta E(\tau) = -\int_{-\infty}^\tau F_t(\tau') d\tau', \qquad \mu \Delta L(\tau) = \int_{-\infty}^\tau F_\phi(\tau') d\tau'. \tag{38}$$

In principle, the coupled set (35) with (37) determines the self-accelerated orbit, given the initial values E_∞, L_∞ and a method for calculating the GSF along the orbit.

B. Dissipative and conservative pieces of the self-force

The quantities $\Delta E(\tau)$ and $\Delta L(\tau)$ encapsulate both conservative and dissipative effects of the GSF. This terminology refers to a splitting of the GSF in the form

$$F^\alpha = F^\alpha_{\text{cons}} + F^\alpha_{\text{diss}}, \tag{39}$$

where the first and second terms are the self-forces exerted, respectively, by the "time-symmetric" and "time-antisymmetric" pieces of the (regularized) metric perturbation (cf. [25] for a more precise definition). For geodesic motion in the equatorial plane of a Kerr black hole, F^α can be thought of as a function of only r and $\dot{r}$ along the orbit. The particular time symmetry of such geodesics then implies [25]

$$F^\alpha_{\text{cons}}(r, \dot{r}) = \frac{1}{2}[F^\alpha(r, \dot{r}) + s_{(\alpha)}F^\alpha(r, -\dot{r})], \quad (40)$$

$$F^\alpha_{\text{diss}}(r, \dot{r}) = \frac{1}{2}[F^\alpha(r, \dot{r}) - s_{(\alpha)}F^\alpha(r, -\dot{r})] \quad (41)$$

(no summation over α), where $s_{(t)} = -1 = s_{(\phi)}$ and $s_{(r)} = +1$. This gives a simple prescription for constructing F^α_{cons} and F^α_{diss} along geodesics, given the full GSF.

For circular orbits we have $F^\alpha(r, \dot{r}) = F^\alpha(r, -\dot{r})$, meaning F^t, F^ϕ are purely dissipative while F^r is purely conservative. In general, however, each component has both dissipative and conservative pieces. Of particular interest to us will be nearly circular orbits with $|\dot{r}| \ll 1$. Along such orbits we may write, to leading order in $|\dot{r}|$,

$$F^\alpha_{\text{cons}} \simeq \dot{r}F^\alpha_1(r), \qquad F^\alpha_{\text{diss}} \simeq F^\alpha_0(r) \quad (42)$$

for $\alpha = t, \phi$, and

$$F^r_{\text{cons}} \simeq F^r_0(r), \qquad F^r_{\text{diss}} \simeq \dot{r}F^r_1(r), \quad (43)$$

where F^α_0 and F^α_1 are some functions of r only.

Equations (40)–(43) are applicable, at leading order in η, even for an orbit that is slowly evolving under the GSF effect. In that case the GSF depends also on the instantaneous self-acceleration, but that dependence appears only at subleading order in η. At leading order, Eqs. (40)–(43) maintain their form at each point along the orbit.

The GSF integrals ΔE and ΔL can be related, in certain situations, to asymptotic fluxes of energy and angular momentum in gravitational waves. This was established rigorously in Ref. [27] for a trajectory starting and ending at infinity.[3] A similar balance relation has been argued to hold also for adiabatic inspiral orbits around a black hole, subject to a suitable averaging over many orbital periods [28,29]. In both scenarios, the contribution from F^α_{cons} to the integrals ΔE and ΔL (taken from $\tau = -\infty$ to $\tau = +\infty$) vanishes at leading order, by virtue of the orbital symmetry expressed in Eq. (40). This guarantees that the radiated fluxes balance the work done by the *dissipative* piece of the self-force alone, as expected.

---

[3]The configuration considered in Ref. [27] had no black hole in it, but the authors argue convincingly that a similar conclusion would hold also in the black-hole case, if fluxes down the event horizon were accounted for in the balance equation.

C. ADM energy and angular momentum

Our analysis in the next section will require knowledge of the total, conserved ADM energy and angular-momentum contents of the spacetime in the above setup. Specifically, we will need expressions for E_{ADM} and L_{ADM} in terms of E_∞ and L_∞ (and the two masses, M and μ), correct through $O(\eta^2)$. A subtlety is that ADM quantities are most conveniently evaluated in a "center-of-mass" (CoM) system (and, at the required order, would include a contribution from the black hole's "recoil" motion), whereas E_∞ and L_∞ are components of the particle's four-velocity, defined in a coordinate system centered around the black hole.

In our setup, the ADM quantities are most easily evaluated on a hypersurface of constant $t \ll -M$, where the binary separation is $r \gg M$. In the limit $t \to -\infty$ ($r \to \infty$), the gravitational interaction energy vanishes and does not contribute to E_{ADM}. Working at that limit, we assume that, for the purpose of calculating ADM quantities, the black hole–article system may be replaced with that of two relativistic pointlike particles in flat spacetime. E_{ADM} is then simply the sum of the two relativistic energies in the center-of-mass frame, and L_{ADM} is similarly the sum of two angular momenta (with respect to the center of mass), plus the spin of the black hole.

Appendix A gives the details of this calculation, which is straightforward. The result is

$$E_{\text{ADM}} = M\left[1 + \eta E_\infty - \frac{1}{2}\eta^2(E_\infty^2 - 1)\right] + o(\eta^2), \quad (44)$$

$$L_{\text{ADM}} = M(a + \eta L_\infty - \eta^2 L_\infty E_\infty) + o(\eta^2). \quad (45)$$

D. Critical orbits

In the geodesic case we have introduced the function $L_c(E)$, which we now interpret as the *critical value* of the angular momentum for a given energy: Geodesic orbits with $L > L_c(E)$ scatter back to infinity, while ones with $L < L_c(E)$ fall into the black hole. This type of critical behavior carries over to the GSF case, though radiation losses then introduce a subtlety, since orbits that are initially scattered may fall into the black hole at a subsequent approach. However, we may still speak of a critical threshold for an *immediate* capture, which separates (in the space of initial conditions) between orbits that scatter at first approach and orbits that do not. A detailed analysis of this critical behavior was given in Ref. [30] for orbits in Schwarzschild spacetime (working in the first-order GSF approximation, as here), and in the following discussion we assume the same qualitative behavior applies in the Kerr case too.

In particular, we assume there exists a critical value $L_\infty = L_{\infty,c}(E_\infty)$ that separates between the two possible
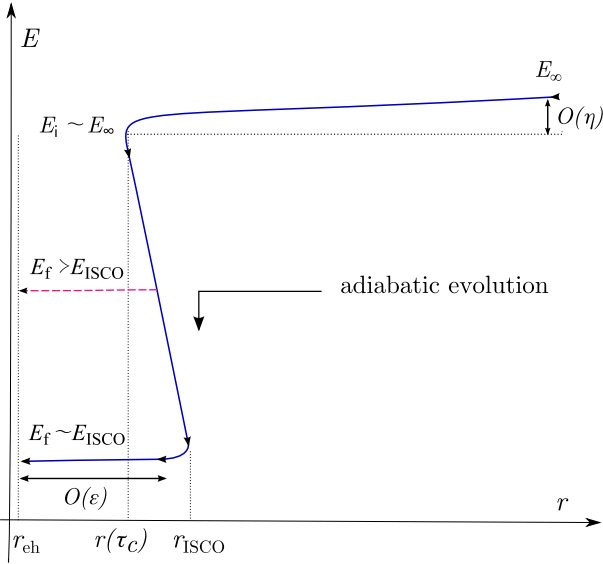
FIG. 4 (color online).   Schematic illustration of the evolution of orbital energy along a perfectly fine-tuned critical orbit (solid blue curve). The orbit approaches from infinity, becomes trapped on an unstable circular orbit, and then evolves adiabatically in a quasicircular fashion before transiting to plunge around the ISCO location. Radiative losses are small during the approach and plunge, but, through fine-tuning, the orbit can be made to lose "all" its energy during the quasicircular stage. Intermediate values of the final energy E_f may also be obtained by fine-tuning (dashed magenta line). Note the orbital radius *increases* through radiation losses during the quasicircular stage. In the near-extremal case, $\epsilon \ll 1$, the quasicircular evolution and final plunge occur within a small range of coordinate radii, $\Delta r = O(\epsilon)$.

outcomes. The initial conditions $\{E_\infty, L_{\infty,c}(E_\infty)\}$ thus define a one-parameter family of "critical orbits." Let us denote by $\hat{E}_c(\tau; E_\infty)$ and $\hat{L}_c(\tau; E_\infty)$ the functions $\hat{E}(\tau)$ and $\hat{L}(\tau)$ corresponding to a critical orbit with a given E_∞ [so that $\hat{E}_c(\tau \to -\infty; E_\infty) = E_\infty$ and $\hat{L}_c(\tau \to -\infty; E_\infty) = L_{\infty,c}(E_\infty)$]. Unlike in the geodesic case where critical geodesics of different E are disjoint, in the GSF case all critical orbits join a *global attractor*, which is the perfectly fine-tuned orbit that evolves radiatively along the sequence of unstable circular orbits starting at the light ring and ending at the ISCO, where it plunges into the black hole. Figure 1 in Ref. [30] illustrates the evolution of the critical orbit along the attractor, and see also Fig. 4 above.

Let us define the "GSF correction"

$$\delta L_c(\tau; E_\infty) := \hat{L}_c(\tau; E_\infty) - L_c(E_\infty), \qquad (46)$$

and then

$$\delta L_\infty(E_\infty) := \delta L_c(\tau \to -\infty; E_\infty). \qquad (47)$$

δL_∞ is the GSF-induced shift in the critical value of L_∞ at a fixed E_∞. It may also be interpreted in terms of a GSF

correction to the critical impact parameter. We assume that the difference $\delta L_c(\tau; E_\infty)$ remains small $[O(\eta)]$ during the approach, which should be the case in any reasonable gauge. However, clearly, that difference ceases to remain small as the critical orbit joins the global attractor and evolves along it; then the meaning of $\delta L_c(\tau; E_\infty)$ as a small GSF correction is lost.

For our analysis of overspinning orbits in the next section, we will require an explicit expression for $\delta L_\infty(E_\infty)$ in terms of GSF quantities. It is instructive to derive this relation first with the dissipative piece of the GSF turned off, i.e. replacing the full GSF with its conservative piece (in which case the global attractor disappears, and critical orbits of different E_∞ remain disjoint). Let us call the resulting quantity $\delta L_\infty^{\rm cons}(E_\infty)$. As a second step we will restore dissipation and consider its effect.

1. Conservative GSF effect

With dissipation turned off, the critical orbit becomes exactly stationary at $\tau \to \infty$, where it joins an unstable (nongeodesic) circular orbit of radius $\hat{R}(E_\infty) = R(E_\infty) + \delta R$. Here $R(E_\infty)$ is the geodesic relation given in Eq. (14), and δR is a conservative GSF correction. To obtain $\delta L_\infty^{\rm cons}$ we first substitute $\hat{E}$ and $\hat{L}$ from Eq. (37) into the radial equation of motion (35), replacing L_∞ with $L_c(E_\infty) + \delta L_\infty^{\rm cons}(E_\infty)$, where $L_c(E_\infty)$ is the geodesic relation given in Eq. (16). We then demand $dr/d\tau = 0$ as well as $d^2r/d\tau^2 = 0$ at $r = \hat{R}$ through $O(\eta)$. At leading order in ϵ this yields two algebraic equations for the two $O(\eta)$ unknowns $\delta L_\infty^{\rm cons}$ and δR, given E_∞ and the GSF. The solution is

$$\delta L_\infty^{\rm cons}(E_\infty) = 2M\Delta E^{\rm cons}(\infty) - \Delta L^{\rm cons}(\infty), \qquad (48)$$

and $\delta R(E_\infty) = O(\epsilon)O(\eta)$. Here $\Delta E^{\rm cons}$ and $\Delta L^{\rm cons}$ are the same as ΔE and ΔL of Eq. (38), but with $F_\alpha \to F_\alpha^{\rm cons}$, and with the GSF integrals evaluated along the critical orbit with energy at infinity E_∞. The precise dependence of δR on the GSF will not be needed, but we note that the $O(\epsilon\eta)$ GSF correction to the radius of the critical circular orbit is reassuringly small compared to the $O(\epsilon)$ radial distance to the light ring.

To simplify the appearance of subsequent equations, let us from now on use units in which $M = 1$. This, in particular, makes our "tilde" notation redundant (with $\tilde{L} = L$, etc.) and μ becomes interchangeable with η. Recalling our W notation from Eq. (21), we rewrite Eq. (48) as

$$\delta L_\infty^{\rm cons}(E_\infty) = \Delta W^{\rm cons}(\infty), \qquad (49)$$

where $\Delta W^{\rm cons}$ represents the conservative piece of

$$\Delta W(\tau) := 2\Delta E(\tau) - \Delta L(\tau)$$

$$= -\eta^{-1} \int_{-\infty}^{\tau} (2F_t + F_\phi) d\tau. \qquad (50)$$

The quantity $\Delta W^{\mathrm{cons}}(\infty)$ is the limit $\tau \to \infty$ of $\Delta W^{\mathrm{cons}}(\tau)$. Does this limit actually exist? The answer is positive, since both F_t^{cons} and F_ϕ^{cons} vanish exponentially fast in τ as the orbit approaches the limiting circular orbit at $\tau \to \infty$.

To make this last statement more precise, let us split the τ integral into an "approach" piece, $\int_{-\infty}^{\tau_c}$, and a "quasicircular" piece, $\int_{\tau_c}^{\infty}$, with τ_c chosen so that $\delta r(\tau_c)$, where $\delta r(\tau) := r(\tau) - \hat{R}$, is already very small. For a small δr we have the form $F_t^{\mathrm{cons}} \simeq \dot{r} F_{1t}(r)$ [Eq. (42)] and similarly for F_ϕ^{cons}. Thus $\int_{\tau_c}^{\infty} F_t^{\mathrm{cons}} d\tau \simeq -F_{1t}(\hat{R})\delta r(\tau_c)$, and similarly for the ϕ component. A local analysis of Eq. (35) near the limiting circular orbit gives $\delta r \sim e^{-\lambda \tau}$, with a Lyapunov exponent $\lambda = M^{-1}(3E_\infty^2 - 1)^{1/2}$ at leading order in ϵ (and ignoring the small effect of the GSF). The choice $\tau_c = -\lambda^{-1} \log \eta$, for example, gives $\delta r(\tau_c) \sim \eta$, and the quasicircular piece of ΔW^{cons} does not contribute to $\delta L_\infty^{\mathrm{cons}}$ at leading order in η.

Our discussion assumes that $\Delta E(\tau_c)$ and $\Delta L(\tau_c)$ [hence also $\Delta W(\tau_c)$] are $O(\eta)$ quantities, i.e. that the accumulated GSF-induced positional shift in the orbit during the approach is a small, $O(\eta)$ quantity. This should be the case in any reasonable gauge. Under this assumption, we note that the value of the integral ΔW remains unchanged, at leading order in η, if in Eq. (50) we replace the integration along the actual, GSF-perturbed orbit, with an integration along the critical *geodesic* of energy E_∞. This can be exploited to simplify actual calculations: To compute $\delta L_\infty^{\mathrm{cons}}$ at leading order in η requires only an evaluation of the GSF along a fixed geodesic, and there is no need to consider the backreaction from the GSF on the orbital trajectory.

2. Full-GSF effect

Now restore dissipation. The fine-tuned critical orbit no longer settles into a strictly stationary motion, but rather it continues to evolve radiatively, in an adiabatic fashion, through a sequence of unstable circular orbits of decreasing energies (hence increasing radii). With a perfect fine-tuning, the orbit can reach the ISCO before plunging into the black hole—a scenario illustrated in Fig. 4. A relation between the degree of fine-tuning and the amount of energy loss was derived in Ref. [30] (for the Schwarzschild case): Rewriting their Eq. (124) in terms of angular momentum, we have the scaling $\delta L / L_{\infty,c} \sim \exp[(E_{\mathrm{f}} - E_{\mathrm{i}})/\eta]$, where δL (not to be confused with the GSF shift δL_∞) is any small perturbation in the value of the initial angular momentum off the critical value $L_{\infty,c}$, and $E_{\mathrm{f}} - E_{\mathrm{i}}$ is the resulting change in specific energy as the orbit progresses along the global attractor. To achieve an $O(1)$ change in the specific

energy requires an "exponentially delicate" fine-tuning, $\delta L / L_{\infty,c} \sim \exp(-1/\eta)$.

For our analysis we do not require knowledge of the perfectly fine-tuned angular momentum at that level. We need $L_{\infty,c}$ through $O(\eta)$ only. Fine-tuning at $O(\eta)$ [corresponding to $\delta L = o(\eta)$] guarantees only $E_{\mathrm{f}} - E_{\mathrm{i}} = O(\eta \ln \eta)$. Therefore, for the purpose of determining $L_{\infty,c}$ through $O(\eta)$, it is sufficient to restrict attention to the early part of the critical orbit, where ΔE and ΔL (specific values) are still $O(\eta \ln \eta)$ at most, and have not yet accumulated $O(1)$ changes. This observation assists us in Appendix B, where we derive an expression for the leading-order full-GSF correction $\delta L_\infty(E_\infty)$.

Our main result in Appendix B is that

$$\delta L_\infty(E_\infty) = \Delta W(\tau_c) + O(\epsilon)O(\eta \ln \eta), \qquad (51)$$

in analogy with the "no dissipation" case, Eq. (49). Here, $\Delta W(\tau_c)$ is the full-GSF integral shown in Eq. (50), evaluated along the orbit from infinity and up to the "end-of-approach" time τ_c, when the orbit settles into a quasicircular motion. Crucially, the contribution to ΔW from the quasicircular part of the orbit is suppressed by a factor of ϵ, so that the precise choice of τ_c does not affect the value of $\Delta W(\tau_c)$ at leading order. This assumes only that $\epsilon \ll |\ln \eta|^{-1}$, so that the error terms in Eq. (51) are negligible compared to $\Delta W(\tau_c) = O(\epsilon^0)O(\eta)$. All we require of τ_c is to be sufficiently late that $|\dot{r}|$ is already very small [specifically, $\dot{r}(\tau_c) = O(\eta)$], but sufficiently early that ΔE is $O(\eta \ln \eta)$ at most. In practice, ΔW may again be evaluated along the critical *geodesic* of energy at infinity E_∞, with the integral in Eq. (51) truncated after, say, 4–5 orbital revolutions. Truncating instead after (e.g.) 10 revolutions should change the result by a negligible amount of only $O(\epsilon)O(\eta)$.

The dissipative piece of $\Delta W(\tau_c)$ [call it ΔW^{diss}, defined by replacing $F_\alpha \to F_\alpha^{\mathrm{diss}}$ in Eq. (51)] may be expressed in terms of radiated quantities. Let $\mathcal{E}_{(\mathrm{apr})}$ and $\mathcal{L}_{(\mathrm{apr})}$ be the total energy and angular momentum in gravitational waves radiated out to infinity and down the black hole during the approach. We shall assume that the balance relation,[4,5]

$$\mathcal{W}_{(\mathrm{apr})} := 2\mathcal{E}_{(\mathrm{apr})} - \mathcal{L}_{(\mathrm{apr})} = -\eta \Delta W^{\mathrm{diss}}(\tau_c) \qquad (52)$$

---

[4]The balance (52) does not follow directly from the theorem of Ref. [27], because the approach part of the critical orbit does not end at infinity. It may be possible to construct a proof by considering a small outward deformation of the orbit (such that the new orbit starts at infinity and scatters back to infinity), then invoking the approximate symmetry about the periapsis, together with the ϵ-suppression of the quasicircular contribution to ΔW^{diss}. We shall not endeavor to provide the details of such a proof here.

[5]One cannot expect to be able to similarly balance $\mathcal{E}_{(\mathrm{apr})}$ and $\mathcal{L}_{(\mathrm{apr})}$ individually, because the dissipative pieces of $\Delta E(\tau_c)$ and $\Delta L(\tau_c)$, unlike $\Delta W^{\mathrm{diss}}(\tau_c)$, *are* sensitive to the choice of τ_c already at leading order. However, such individual balance relations will not be needed in our analysis.

holds at leading order in η and in ϵ. Equations (49) and (51) then lead to

$$\delta L_\infty = \delta L_\infty^{\text{cons}} - \mathcal{W}_{(\text{apr})}/\eta, \qquad (53)$$

where subleading terms have been omitted. This reexpresses δL_∞ as a sum of conservative and radiative contributions, the motivation for which will become clear in the next section.

Finally, let us further write $\mathcal{E}_{(\text{apr})} = \mathcal{E}_{(\text{apr})}^+ + \mathcal{E}_{(\text{apr})}^-$ and similarly for $\mathcal{L}_{(\text{apr})}$ and $\mathcal{W}_{(\text{apr})}$, where hereafter superscripts "+" and "−" denote contributions from fluxes to infinity and down the black hole, respectively. The following argument, based on the first law of black-hole mechanics, suggests that $\mathcal{W}_{(\text{apr})}^-$ must vanish in the limit $\epsilon \to 0$. If we assume the black hole is not overspun during the approach, its horizon's area should increase by an amount δA satisfying

$$\frac{\kappa}{8\pi}\delta A = \mathcal{E}_{(\text{apr})}^- - \Omega_H \mathcal{L}_{(\text{apr})}^-, \qquad (54)$$

where $\kappa = \epsilon/\sqrt{2} + O(\epsilon^2)$ is the horizon's surface gravity, and Ω_H its angular velocity. Since $\Omega_H = \frac{1}{2} + O(\epsilon)$, we identify the right-hand side of (54) as $\mathcal{W}_{(\text{apr})}^-$ at leading order in ϵ. We thus have, in the extremal limit, $\mathcal{W}_{(\text{apr})}^- \simeq \epsilon(c_1 \delta A + c_2 \mathcal{L}_{(\text{apr})}^-)$, with c_1, c_2 certain numerical coefficients. Since δA and $\mathcal{L}_{(\text{apr})}^-$ must remain bounded even in the extremal limit, we conclude that $\mathcal{W}_{(\text{apr})}^-$ vanishes in that limit. Thus, at leading order in ϵ, Eq. (53) becomes

$$\delta L_\infty = \delta L_\infty^{\text{cons}} - \mathcal{W}_{(\text{apr})}^+/\eta, \qquad (55)$$

which now features only outgoing fluxes.

With this we have completed the necessary groundwork for our overspinning analysis, to be presented next.

IV. OVERSPINNING WITH THE FULL SELF-FORCE

A. General form of the censorship condition and reduction to near-critical orbits

Starting with a near-extremal Kerr geometry with $a/M = 1 - \epsilon^2$, consider a particle sent in from infinity with specific energy E_∞ and specific angular momentum L_∞ at $t \to -\infty$. The ADM mass and angular momentum of the spacetime are given in Eqs. (44) and (45) through $O(\eta^2)$. We assume the particle crosses the event horizon[6] at some retarded time u_h, and we let $\mathcal{E}^+$ and $\mathcal{L}^+$ be the total energy and angular momentum radiated to null infinity up until u_h (with $u_h \to \infty$ if the postcapture geometry relaxes

to a Kerr black hole). Then the Bondi mass and angular momentum of spacetime at retarded time u_h are $E_{\text{ADM}} - \mathcal{E}^+$ and $L_{\text{ADM}} - \mathcal{L}^+$, respectively. Overspinning is *avoided* if and only if[7]

$$(E_{\text{ADM}} - \mathcal{E}^+)^2 \geq L_{\text{ADM}} - \mathcal{L}^+. \qquad (56)$$

To rule out the overspinning scenario, we need to show that this inequality holds for all E_∞, L_∞ and for all η, ϵ. Since we work in first-order perturbation theory, we only demand that (56) is satisfied at leading order in η. We also assume $\epsilon \ll 1$ and keep only leading terms in ϵ, but we do not *a priori* restrict the magnitude of ϵ relative to that of η. We shall refer to the inequality (56) as the *censorship condition*.[8]

Substituting from Eqs. (1), (44) and (45), the censorship condition becomes

$$\epsilon^2 + \eta W_\infty + \eta^2(1 + L_\infty E_\infty - E_\infty^2)$$
$$+ (\eta E_\infty - \mathcal{E}^+)^2 - \mathcal{W}^+ \geq 0, \qquad (57)$$

where $W_\infty := 2E_\infty - L_\infty$, $\mathcal{W}^+ := 2\mathcal{E}^+ - \mathcal{L}^+$, and we have omitted subleading terms of $o(\eta^2)$. Note how the various terms here scale with η. The quantities E_∞ and L_∞ (hence also W_∞) are specific values, thus a priori they are $O(\eta^0)$. The radiated energy $\mathcal{E}^+$ is generically $O(\eta^2)$, but may accumulate at $O(\eta)$ for orbits that are fine-tuned to evolve along the global attractor; it is to allow for such orbits that we have kept the terms $2\eta E_\infty \mathcal{E}^+$ and $(\mathcal{E}^+)^2$ in Eq. (57). The quantity $\mathcal{W}^+$ is likewise $O(\eta^2)$ generically and up to $O(\eta)$ with fine-tuning, but, as will be shown below, in the latter case the $O(\eta)$ term is also proportional to ϵ.

Inspecting Eq. (57), we observe that, for all captured orbits that are not sufficiently close to criticality, the term ηW_∞ is $O(\eta)$ and positive, so the censorship condition (57) is trivially satisfied at leading order in η and ϵ. Violation of (57) (hence overspinning) may only be achieved, potentially, if L_∞ is tuned so that $L_\infty = 2E_\infty + O(\epsilon, \eta)$, giving $W_\infty = O(\epsilon, \eta)$. It is therefore sufficient to restrict attention to this class of orbits, to be referred to in what follows as "near critical." Formulated on near-critical orbits, the censorship condition takes the sufficient form

---

[6] More pedantically, we refer here to the crossing of a marginally trapped surface; spacetime need not contain a global horizon.

[7] We do not know, and for our purpose do not need to know, the future evolution of spacetime beyond retarded time u_h in the hypothetical case where (56) is *not* satisfied. The likely scenario involves the formation of a naked singularity and a breakdown of predictability for $u > u_h$ [4]. If (56) *is* satisfied, then, by "no-hair" theorems, geometry should relax to a Kerr black hole.

[8] It may be argued that (56) is guaranteed to hold (with a strong inequality) by virtue of the third law of black-hole mechanics [31], though it is clear that some of the third law's assumptions are not satisfied within our model—see [9] for a discussion. Even if that can be established, it is still of interest to explore the *physical mechanism* that enforces the third law in our setup, which is what our study aims to achieve.

$$\epsilon^2 + \eta W_\infty + \eta^2(1 + E_\infty^2) + (\eta E_\infty - \mathcal{E}^+)^2 - \mathcal{W}^+ \geq 0, \tag{58}$$

where we have dropped $O(\epsilon\eta^2, \eta^3)$ terms. This is required to hold for each member of the reduced two-parameter family $\{E_\infty, L_\infty\}$ with $L_\infty - 2E_\infty = O(\epsilon, \eta)$.

To proceed, we need to make more precise the distinction between fine-tuned orbits and generic near-critical orbits that are not fine-tuned. Referring to Fig. 4, let E_f be the final value of the specific energy with which the particle plunges into the hole, and let $L_{\infty,c}(E_\infty)$ be the perfectly fine-tuned value of L_∞, for which the orbit joins the global attractor and evolves along it to the ISCO. Assuming the universal scaling $L_\infty - L_{\infty,c}(E_\infty) \sim \exp[(E_f - E_\infty)/\eta]$ [30], near-critical orbits as defined above generically have $E_f - E_\infty = O(\eta \ln \eta\epsilon)$ [here we neglect the $O(\eta)$ difference between E_∞ and E_i]. Calibrating L_∞ at higher order in η, ϵ [so that $L_\infty - L_{\infty,c} = O(\eta^n, \epsilon^k)$ with some $n, k > 1$] does not qualitatively change this generic scaling of $E_f - E_\infty$. To achieve $E_f - E_\infty = O(1)$ requires an exponentially accurate tuning, i.e. $L_\infty - L_{\infty,c} \sim \exp(-1/\eta)$. In what follows we use the η-scaling of $E_f - E_\infty$ to distinguish between *generic* and *fine-tuned* members of the near-critical family: The former admit $E_f - E_\infty = O(\eta \ln \eta\epsilon)$, and the latter $E_f - E_\infty = O(1)$. This distinction can also be formulated in terms of the radiated quantities $\mathcal{E}^+$ or $\mathcal{L}^+$:

$$\mathcal{E}^+, \mathcal{L}^+ = O(\eta^2 \ln \eta\epsilon) \quad \text{(generic)}, \tag{59}$$

$$\mathcal{E}^+, \mathcal{L}^+ = O(\eta) \quad \text{(fine-tuned)}. \tag{60}$$

B. Further reduction to critical orbits

The inequality (58) is still a condition on a two-parameter family of orbits. Ignoring fine-tuned orbits for now, it is possible—and beneficial—to reduce it further to a sufficient condition formulated on a one-parameter family. We achieve this by minimizing the left-hand side of Eq. (58) over all near-critical orbits for a given E_∞. We argue that the minimizing orbit is one with L_∞ tuned to its critical value at least through $O(\eta, \epsilon)$, namely

$$L_\infty = 2E_\infty + \epsilon\sqrt{6E_\infty^2 - 2} + \delta L_\infty(E_\infty) + o(\eta, \epsilon), \tag{61}$$

where we have recalled Eq. (16), and $\delta L_\infty(E_\infty)$ is the $O(\eta)$ GSF term derived in the previous section. To see this, note first that $W_\infty = 2E_\infty - L_\infty$ is trivially minimized by $L_\infty = L_{\infty,c}(E_\infty)$, since $L_{\infty,c}(E_\infty)$ maximizes L_∞ (over all captured orbits of a fixed E_∞) by definition of a critical orbit. This means that, to minimize W_∞ through $O(\eta, \epsilon)$ (higher orders are irrelevant in our approximation) it suffices to demand $L_\infty - L_{\infty,c}(E_\infty) = o(\eta, \epsilon)$. Then also note that the two radiative terms on the left-hand side of (58) are insensitive, at relevant order, to variations of L_∞ within the family of nearly critical orbits for a fixed E_∞. For

generic orbits, the term $(\eta E_\infty - \mathcal{E}^+)^2$ is simply $\eta^2 E_\infty^2$ at leading order, recalling Eq. (59). As for the term $-\mathcal{W}^+$, we note that the contribution to that term from the approach part of the orbit, which is already $O(\eta^2)$, is not sensitive, at that order, to $O(\eta, \epsilon)$ variations in L_∞. Meanwhile, the contribution to $\mathcal{W}^+$ from the quasicircular part of the orbit is of $O(\epsilon)O(\eta^2 \ln \eta\epsilon)$ at most (the occurrence of the factor ϵ will be explained below) and hence negligible in Eq. (58), assuming only $\epsilon \ll |\ln \eta|^{-1}$.

Thus, discounting fine-tuned orbits, we find that the entire left-hand side of Eq. (58) is minimized by L_∞ as given in Eq. (61). A new sufficient version of the censorship condition may therefore be written as

$$\epsilon^2 - \eta\epsilon\sqrt{6E_\infty^2 - 2} - \eta\delta L_\infty + \eta^2(1 + E_\infty^2)$$
$$+ (\eta E_\infty - \mathcal{E}^+)^2 - \mathcal{W}^+ \geq 0, \tag{62}$$

which, at the relevant, leading order, is a condition on the one-parameter family of (generic) critical orbits parametrized by E_∞ alone.

It should now be noted that the condition (62) also applies to fine-tuned orbits [whether or not they minimize the left-hand side of (58)], simply because such orbits always satisfy Eq. (61). However, for fine-tuned orbits the condition still involves *two* parameters, conveniently chosen as E_∞ and E_f. Different values of E_f correspond to a fine-tuning of L_∞ at an exponential level. In principle, any value of E_f in the range $E_{\text{isco}} \lesssim E_f \lesssim E_\infty$ may be obtained this way. To rule out overspinning by fine-tuned orbits, the censorship condition (62) must hold for all $\{E_\infty, E_f\}$ with E_f in the above range.

Observe that in Eq. (62) we have

$$\eta\delta L_\infty + \mathcal{W}^+ = \eta\delta L_\infty^{\text{cons}} + \mathcal{W}^+ - \mathcal{W}_{(\text{apr})}^+$$
$$= \eta\delta L_\infty^{\text{cons}} + \mathcal{W}_{(\text{qc})}^+ + \mathcal{W}_{(\text{end})}^+, \tag{63}$$

where in the first line we have recalled Eq. (55), $\mathcal{W}_{(\text{qc})}^+$ is the piece of $\mathcal{W}^+$ coming from the evolution along the quasi-circular part of the orbit, and $\mathcal{W}_{(\text{end})}^+$ is the piece coming from the transition to a final plunge into the black hole and from the plunge itself. It follows that only the *conservative* piece of the shift δL_∞ actually enters our condition:

$$\epsilon^2 - \eta\epsilon\sqrt{6E_\infty^2 - 2} - \eta\delta L_\infty^{\text{cons}} + \eta^2(1 + E_\infty^2)$$
$$+ (\eta E_\infty - \mathcal{E}^+)^2 - \mathcal{W}_{(\text{qc})}^+ - \mathcal{W}_{(\text{end})}^+ \geq 0. \tag{64}$$

In this last form, conservative and dissipative terms of the GSF feature separately. The former are associated with the approach leg of the orbit, and the latter accumulate during the adiabatic evolution along the attractor. In Appendix C we combine results by Ori and Thorne [32], Kesden [33] and Mino and Brink [34], to argue that

the term $\mathcal{W}^+_{(\text{end})}$ is always subdominant and negligible in Eq. (64). We shall therefore omit that term in the rest of our discussion.

In Sec. IV D below we will show that the radiative term $\mathcal{W}^+_{(\text{qc})}$ scales as $O(\epsilon)O[\eta(E_\infty - E_f)]$. This term can thus feature at leading order in Eq. (64) only for fine-tuned orbits, for which $E_\infty - E_f = O(1)$. Likewise, terms involving $\mathcal{E}^+$ feature only for fine-tuned orbits and are negligible otherwise. On the other hand, the conservative term $\eta L^{\text{cons}}_\infty$ is always $O(\eta^2)$, featuring in the censorship condition regardless of fine-tuning. An important consequence is that *dissipative effects of the GSF enter the censorship condition only for fine-tuned orbits*. This seems consistent with suggestions made in earlier analyses [13,33,35] (in which fine-tuning has not been considered).

Below we further simplify the condition (64), and reformulate it explicitly in terms of E_∞ alone (for generic orbits) or E_∞ and E_f alone (for fine-tuned ones), without reference to η and ϵ. We shall consider separately the cases of generic and fine-tuned orbits, starting with the former, simpler case.

C. Censorship condition for generic orbits

As mentioned above (and shown in the next subsection), without fine-tuning the radiative terms $\mathcal{E}^+$ and $\mathcal{W}^+_{(\text{qc})}$ become subdominant in Eq. (64) and drop out of it. The censorship condition then reduces to

$$\epsilon^2 + \eta\epsilon F + \eta^2 H \geq 0, \tag{65}$$

with

$$F := -\sqrt{6E_\infty^2 - 2}, \tag{66}$$

$$H := 1 + 2E_\infty^2 - \delta\check{L}^{\text{cons}}_\infty. \tag{67}$$

Here we have made the η-scaling of $\delta L^{\text{cons}}_\infty$ explicit by introducing the shift per eta

$$\delta\check{L}^{\text{cons}}_\infty := \eta^{-1}\delta L^{\text{cons}}_\infty, \tag{68}$$

which should have a finite (nonzero) limit $\eta \to 0$.

For the overspinning scenario to be ruled out, the inequality (65) must be satisfied for all $\eta, \epsilon > 0$ and all $E_\infty \geq 1$. The condition can be written in the equivalent form $\Phi := \alpha^2 + \alpha F + H \geq 0$, with $\alpha := \epsilon/\eta > 0$. At fixed E_∞, Φ is quadratic in α, with a minimum value $\Phi(\alpha = -F/2) = H - F^2/4$. To guarantee $\Phi \geq 0$ for all E_∞ and all $\alpha > 0$ (hence all $\eta, \epsilon > 0$) we must demand $H \geq F^2/4$; if $H < F^2/4$ for some E_∞, then for that E_∞ there exist η, ϵ values for which overspinning is achieved. In that way, $H \geq F^2/4$ is both sufficient and necessary for overspinning to be avoided. Inserting the values of F and H, the censorship condition takes the simple form

$$\delta\check{L}^{\text{cons}}_\infty \leq \frac{1}{2}(E_\infty^2 + 3). \tag{69}$$

Overspinning is averted (for orbits that are not fine-tuned) if and only if (69) is satisfied for each member of the one-parameter family of critical orbits with $E_\infty \geq 1$, in the limit $\eta, \epsilon \to 0$.

Equation (69) states our final result for generic orbits. As already mentioned, it involves only *conservative* GSF effects, specifically the shift in the critical value of the angular momentum at infinity (at fixed E_∞) due to the conservative piece of the GSF. For easy reference, we give here the explicit formula for $\delta\check{L}^{\text{cons}}_\infty$ in terms of GSF components:

$$\delta\check{L}^{\text{cons}}_\infty(E_\infty) = \lim_{\epsilon \to 0} \frac{1}{\mu^2} \int_{R_\epsilon}^\infty (2MF^{\text{cons}}_t + F^{\text{cons}}_\phi)dr/\dot{r}, \tag{70}$$

where we have recalled Eqs. (49) and (50) (and have restored a factor M for clarity). The integration is carried out along the critical geodesic of specific energy E_∞ on a background with spin $a/M = 1 - \epsilon^2$, starting at the unstable circular orbit of radius $R_\epsilon = R_\epsilon(\epsilon, E_\infty)$ and ending at infinity.

Inspecting Eq. (69), it may seem peculiar that overspinning may be averted even for some *positive* values of $\delta L^{\text{cons}}_\infty$: A positive $\delta L^{\text{cons}}_\infty$ would seem to mean that the GSF *increases* the critical impact parameter, allowing in particles that would otherwise be scattered away. However, we must recall that the shift $\delta L^{\text{cons}}_\infty$ is defined not with respect to the physical, ADM angular momentum, but with respect to the quantity $\hat{L} = \hat{u}_\phi$, which (while convenient to work with in practice) does not have a clear invariant meaning beyond the geodesic approximation. To rewrite (69) in a more physically insightful way, let us, then, recast it in terms of ADM quantities, as follows.

First, let us introduce the specific quantities $E^{\text{p}}_{\text{ADM}}$ and $L^{\text{p}}_{\text{ADM}}$ defined through

$$\mu E^{\text{p}}_{\text{ADM}} := E_{\text{ADM}} - M,$$
$$\mu L^{\text{p}}_{\text{ADM}} := L_{\text{ADM}} - Ma \tag{71}$$

(restoring factors M for clarity), which may be thought of as the particle's contributions to the total ADM energy and angular momentum of the system. Then, denote by $\delta L^{\text{p}}_{\text{ADM}}(E^{\text{p}}_{\text{ADM}})$ the shift, due to the conservative GSF, in the critical value of $L^{\text{p}}_{\text{ADM}}$ for a fixed $E^{\text{p}}_{\text{ADM}}$. A short calculation, based on Eqs. (44) and (45), gives

$$\delta L^{\text{p}}_{\text{ADM}}(E^{\text{p}}_{\text{ADM}}) = \delta L^{\text{cons}}_\infty(E_\infty) - \eta(E_\infty^2 + 1) + O(\eta^2). \tag{72}$$

Thus, in terms of $\delta\check{L}^{\text{p}}_{\text{ADM}} := \eta^{-1}\delta L^{\text{p}}_{\text{ADM}}$, the censorship condition (69) becomes

$$\delta\check{L}^{\mathrm{p}}_{\mathrm{ADM}} \leq \frac{1}{2}(1 - E^2_\infty), \qquad (73)$$

where on the right-hand side E_∞ may be replaced with $E^{\mathrm{p}}_{\mathrm{ADM}}$ at relevant order.

The alternative form (73) is now more intuitive: For unbound orbits ($E_\infty \geq 1$), the GSF averts overspinning if it shifts the critical value of the (ADM-related) angular momentum by a sufficiently negative amount, which depends only on E_∞. In the marginal case of $E_\infty = 1$ (where overspinning is marginally prevented already in the geodesic case), the shift $\delta\check{L}^{\mathrm{p}}_{\mathrm{ADM}}$ need only be nonpositive. We are not aware of any *a priori* argument to suggest that $\delta\check{L}^{\mathrm{p}}_{\mathrm{ADM}}$ must necessarily be nonpositive for all $E_\infty \geq 1$. Verifying this would need to await a numerical calculation. Any counterexample would imply a direct violation of cosmic censorship.

Let us make a few more points about the condition (69). First, in its form (73) it is manifestly gauge invariant (within a class of suitable asymptotically flat gauges) despite the gauge dependence of the local GSF featuring in $\delta\check{L}^{\mathrm{cons}}_\infty$ [Eq. (70)]. The condition involves only quantities that are defined and evaluated at infinite separation, namely the specific energy E_∞ (or $E^{\mathrm{p}}_{\mathrm{ADM}}$) and angular-momentum shift $\delta\check{L}^{\mathrm{p}}_{\mathrm{ADM}}$, each having a clear gauge-invariant physical meaning. The evident invariance of our final condition is reassuring.

Second, as already mentioned, the condition that (69) is satisfied for all $E_\infty \geq 1$ is both sufficient and necessary for overspinning to be avoided within the scenario considered here. It is a *sufficient* condition only in the sense that it guarantees no overspinning occurs for sufficiently small mass ratio η; since we work in the first-order self-force approximation, we cannot make the statement any stronger. Equation (69) describes a *necessary* condition in the sense that its violation for any E_∞ would mean there exist (small) η values for which overspinning occurs.

Finally, the condition (69) involves the single parameter E_∞, and the task of testing whether it is satisfied amounts to evaluating a single function of E_∞, namely $\delta\check{L}^{\mathrm{cons}}_\infty(E_\infty)$. The perturbative parameters themselves, η and ϵ, do not feature in the final condition. This is expected, given our first-order perturbative treatment and the fact that GSF effects (including ADM terms) appear in the overspinning condition already at leading order. It is precisely because of this "order mixing" that one cannot neglect the GSF in considering the overspinning problem, and why there is no sense in which the geodesic limit may be said to provide a useful approximation here.

D. Censorship condition for fine-tuned orbits

It is not *a priori* clear whether fine-tuning favors the overspinning scenario or disfavors it: The answer depends on the details of the radiative evolution along the attractor. Indeed, for fine-tuned orbits the radiative terms $\mathcal{E}^+$ and

$\mathcal{W}^+_{(\mathrm{qc})}$ feature already at leading order in Eq. (64), and cannot be neglected. We may again write Eq. (64) in the form (65), with F and H replaced with, respectively,

$$\bar{F} = -\sqrt{6E^2_\infty - 2} - \check{\mathcal{W}}^+_{(\mathrm{qc})},$$
$$\bar{H} = 1 + E^2_\infty + (E_\infty - \check{\mathcal{E}}^+)^2 - \delta\check{L}^{\mathrm{cons}}_\infty. \qquad (74)$$

Here we have introduced the rescaled quantities

$$\check{\mathcal{E}}^\pm := \eta^{-1}\mathcal{E}^\pm, \qquad \check{\mathcal{W}}^\pm_{(\mathrm{qc})} := (\epsilon\eta)^{-1}\mathcal{W}^\pm_{(\mathrm{qc})}, \qquad (75)$$

which should have finite (nonzero) limits $\epsilon, \eta \to 0$ for fine-tuned orbits [that $\mathcal{W}^\pm_{(\mathrm{qc})} = O(\epsilon)O(\eta)$ will be discussed in the next two paragraphs].

It will prove beneficial to reexpress $\bar{F}$ and $\bar{H}$ in terms of the absorption-related quantities $\check{\mathcal{E}}^-$ and $\check{\mathcal{W}}^-_{(\mathrm{qc})}$, in place of $\check{\mathcal{E}}^+$ and $\check{\mathcal{W}}^+_{(\mathrm{qc})}$. This is easily done for $\bar{H}$, noting $E_\infty - \check{\mathcal{E}}^+ = E_f + \check{\mathcal{E}}^-$ at the relevant, leading order. As for $\bar{F}$, we start by writing $\mathcal{W}^+_{(\mathrm{qc})} = \mathcal{W}_{(\mathrm{qc})} - \mathcal{W}^-_{(\mathrm{qc})}$, where, under the assumption of adiabaticity, the total $\mathcal{W}_{(\mathrm{qc})}$ may be expressed as an integral over the local GSF:

$$\mathcal{W}_{(\mathrm{qc})} = \int^{E_f}_{E_\infty} (2F^{\mathrm{diss}}_t + F^{\mathrm{diss}}_\phi)dE/\dot{E}. \qquad (76)$$

Here we have used Eq. (50), changing the integration variable from τ to specific energy E, and assumed a balance relation as in Eq. (52). We have also neglected the subdominant [$O(\eta^2)$] amount of radiated energy during the approach, replacing the initial energy of the quasicircular motion with E_∞. Then, following the method of Appendix B [cf. Eq. (B4)], we use $u^\alpha F^{\mathrm{diss}}_\alpha = 0$ to obtain

$$F^{\mathrm{diss}}_t + F^{\mathrm{diss}}_\phi/2 = -(u^r/u^t)F^{\mathrm{diss}}_r - \frac{3E}{\sqrt{6E^2 - 2}}\epsilon F^{\mathrm{diss}}_t, \qquad (77)$$

where subdominant terms in ϵ have been omitted. The contribution from the term $\propto F^{\mathrm{diss}}_r$ to the integral in (76) can be evaluated following the same steps as in Appendix B [see the paragraph containing Eq. (B5)], and shown to be of only $O(\epsilon)O(\eta^2)$ (or smaller)—hence negligible. The contribution from the term $\propto F^{\mathrm{diss}}_t$ can be evaluated explicitly upon replacing $F_t = \mu\dot{E}$, giving

$$\mathcal{W}_{(\mathrm{qc})} = -\eta\epsilon\left(\sqrt{6E^2_\infty - 2} - \sqrt{6E^2_f - 2}\right). \qquad (78)$$

Thus, Eqs. (74) are obtained in their alternative form

$$\bar{F} = -\sqrt{6E_f^2 - 2} + \breve{\mathcal{W}}_{(qc)}^-,$$

$$\bar{H} = 1 + E_\infty^2 + (E_f + \breve{\mathcal{E}}^-)^2 - \delta\breve{L}_\infty^{\text{cons}}. \tag{79}$$

We note that Eq. (78) establishes the scaling $\mathcal{W}_{(qc)} = O(\epsilon)O(\eta)$ for fine-tuned orbits. The first-law argument used in the previous section [refer to the discussion around Eq. (54)] can also be used to show $\mathcal{W}_{(qc)}^- = O(\epsilon)O(\eta)$. This then establishes the scaling $\mathcal{W}_{(qc)}^+ = O(\epsilon)O(\eta)$ assumed above.

In both Eqs. (74) and (79), the radiative quantities $\breve{\mathcal{E}}^\pm$ and $\breve{\mathcal{W}}_{(qc)}^\pm$ should be thought of as functions of E_∞ and E_f only. While $\breve{\mathcal{E}}^+$ is necessarily positive, the absorbed energy $\breve{\mathcal{E}}^-$ may be either positive, or—due to superradiance— negative, depending on E_∞ and E_f. Circular equatorial geodesics are superradiant for $\Omega < \Omega_H$, which, in the extremal limit, translates to $E < \frac{2}{\sqrt{3}}$. Thus, $\breve{\mathcal{E}}^-$ is necessarily negative for any $E_f < E_\infty \le \frac{2}{\sqrt{3}}$. The sign (and magnitude) of $\breve{\mathcal{E}}^-$ for other values of E_∞ and E_f, as well as the sign (and magnitude) of $\breve{\mathcal{W}}_{(qc)}^-$, remain to be determined numerically. The quantity $\breve{\mathcal{W}}_{(qc)}^+$, on the other hand, is easily shown to be negative definite. In fact, Eqs. (84) and (89), given below, imply

$$-\breve{\mathcal{W}}_{(qc)}^+ > \breve{\mathcal{E}}^+ > 0. \tag{80}$$

Note this means that $\bar{F}$ in Eq. (74) may change sign, depending on E_∞, E_f.

To proceed, we once again write the condition (65) (for the barred quantities) in the form $\bar{\Phi} := \alpha^2 + \alpha\bar{F} + \bar{H} \ge 0$, with $\alpha = \epsilon/\eta$. Here, however, the sign of $\bar{F}$ is not known *a priori*, which somewhat complicates matters. For $\bar{F} < 0$, $\bar{\Phi}$ has its minimum at $\bar{\Phi}(\alpha = -\bar{F}/2) = \bar{H} - \bar{F}^2/4$, so the condition becomes $\bar{H} \ge \bar{F}^2/4$ as before. However, for $\bar{F} \ge 0$ the condition $\bar{\Phi} \ge 0$ is satisfied trivially for all $\bar{H} \ge 0$, and violated trivially for all $\bar{H} < 0$ (by choosing a sufficiently small α). In that case, therefore, a necessary and sufficient condition for $\bar{\Phi} \ge 0$ to hold for any η, ϵ is $\bar{H} \ge 0$. In summary, we obtain

$$\bar{H} \ge (\min\{\bar{F}/2, 0\})^2 \tag{81}$$

as a necessary and sufficient condition for overspinning to be averted for all η, ϵ. In this condition, $\bar{F}$ and $\bar{H}$ are both functions of the two independent parameters E_∞ and E_f. To rule out overspinning we must require that (81) is satisfied for all $E_\infty > E_f (> E_{\text{isco}})$.

Evaluation of the condition (81) requires knowledge of the radiative quantities $\breve{\mathcal{E}}^\pm$ and $\breve{\mathcal{W}}_{(qc)}^\pm$ (in addition to $\delta\breve{L}_\infty^{\text{cons}}$). To conclude our discussion, we now give convenient

expressions for these two quantities in terms of a single function of one variable, namely the ratio

$$\mathcal{R}(E) := \frac{\dot{\mathcal{E}}^-(E)}{\dot{\mathcal{E}}^+(E)}, \tag{82}$$

where $\dot{\mathcal{E}}^{+/-}(E)$ are the outgoing/incoming fluxes of energy in gravitational waves sourced by a particle on a circular geodesic, evaluated in the extremal limit at a fixed specific energy E. [In deviation from our notation elsewhere, here and in the next two paragraphs an overdot denotes differentiation with respect to (any suitable) coordinate time.] We note $\mathcal{R} < 0$ for $E < \frac{2}{\sqrt{3}}$, the superradiance regime in the extremal limit.

First, we use the specific energy E as a parameter along the global attractor, to write

$$\breve{\mathcal{E}}^+ = \int_{E_\infty}^{E_f} \frac{\dot{\mathcal{E}}^+}{\eta\dot{E}} dE = -\int_{E_\infty}^{E_f} \frac{\dot{\mathcal{E}}^+}{\dot{\mathcal{E}}^+ + \dot{\mathcal{E}}^-} dE, \tag{83}$$

where we assumed the balance relation $\eta\dot{E} = -(\dot{\mathcal{E}}^+ + \dot{\mathcal{E}}^-)$ applies during the adiabatic evolution along the attractor. Thus,

$$\breve{\mathcal{E}}^+(E_\infty, E_f) = -\int_{E_\infty}^{E_f} \frac{dE}{1 + \mathcal{R}(E)}, \tag{84}$$

and, similarly,

$$\breve{\mathcal{E}}^-(E_\infty, E_f) = -\int_{E_\infty}^{E_f} \frac{\mathcal{R}(E)}{1 + \mathcal{R}(E)} dE, \tag{85}$$

which should be evaluated in the extremal limit, $\epsilon \to 0$. Note $\dot{\mathcal{E}}^\pm \to 0$ in the extremal limit [13], while the ratio $\mathcal{R}$ admits a finite, nonzero limit [36,37]. Thus, by writing $\breve{\mathcal{E}}^\pm$ as in Eqs. (84) and (85) we have made it possible for the limit $\epsilon \to 0$ to be taken before the integration, which is advantageous in practice.

As for $\mathcal{W}_{(qc)}^\pm$, we start by writing

$$\dot{\mathcal{W}}_{(qc)}^\pm := 2\dot{\mathcal{E}}^\pm - \dot{\mathcal{L}}^\pm = -\epsilon b(E)\dot{\mathcal{E}}^\pm, \tag{86}$$

where $\dot{\mathcal{L}}^\pm$ are the angular-momentum fluxes corresponding to $\dot{\mathcal{E}}^\pm$, and

$$b(E) := \frac{6E}{\sqrt{6E^2 - 2}}. \tag{87}$$

To derive the second equality in (86), which is valid to leading order in ϵ, we have used the small-ϵ expansion of the orbital angular velocity at fixed E,

$$\Omega = \frac{1}{2} - \frac{1}{4}b(E)\epsilon + O(\epsilon^2), \qquad (88)$$

together with the general relation $\dot{\mathcal{E}}^\pm = \Omega\dot{\mathcal{L}}^\pm$ applicable to the radiation from any circular orbit [38]. Thus, proceeding as with $\check{\mathcal{E}}^+$, we obtain

$$\check{\mathcal{W}}^+_{(qc)} = \lim_{\epsilon \to 0} \int_{E_\infty}^{E_f} \frac{\dot{\mathcal{W}}^+}{\epsilon\eta\dot{E}} dE = \int_{E_\infty}^{E_f} \frac{b(E)}{1 + \mathcal{R}(E)} dE, \quad (89)$$

and, similarly,

$$\check{\mathcal{W}}^-_{(qc)} = \int_{E_\infty}^{E_f} \frac{b(E)\mathcal{R}(E)}{1 + \mathcal{R}(E)} dE. \qquad (90)$$

Equations (84), (85), (89) and (90) express $\check{\mathcal{E}}^\pm$ and $\check{\mathcal{W}}^\pm_{(qc)}$ in terms of the single function $\mathcal{R}(E)$, left to be determined numerically.

In Sec. VI we will assess the numerical task of evaluating our censorship conditions, review the status of relevant existing GSF codes, and comment on how they would need to be modified in order to provide the necessary data. But first, in the next section, we take a short detour to explore an alternative approach to the determination of $\delta\check{L}^{cons}_\infty$, which offers a practical advantage.

V. REFORMULATION IN TERMS OF REDSHIFT VARIABLE

Our final overspinning conditions (69) and (81) feature the critical angular-momentum shift δL^{cons}_∞, whose evaluation, through Eq. (70), requires an integration of the GSF from infinity along critical geodesics. As we discuss in the next section, this step is the main stumbling block when it comes to evaluating the conditions using currently available GSF codes. The integration from infinity comes about, essentially, because of the need to relate the local properties $\hat{E}$ and $\hat{L}$ of the particle just before it falls into the black hole, to ADM properties of spacetime defined at infinity. This would have been unnecessary if we had available explicit formulas for E_{ADM} and L_{ADM} (or for the corresponding Bondi quantities $E_{ADM} - \mathcal{E}^+$ and $L_{ADM} - \mathcal{L}^+$), correct through $O(\eta^2)$, for the configuration of a particle in an unstable circular orbit around a Kerr black hole. Furthermore, given such formulas we would have been able to relax the requirement that the particle is sent in from infinity, and explore the possibility of overspinning with "bound" orbits. (We recall our result from Sec. II that bound *geodesics* cannot overspin; however, in principle, there remains the possibility that GSF effects change this situation.)

By good fortune, suitable formulas have been proposed very recently, in Ref. [16]. The expressions, to be presented below, were obtained using (and in agreement between) two independent frameworks. One is the Hamiltonian approach of Isoyama and collaborators [39], in which the conservative portion of the orbital dynamics is described (through first order in η beyond the geodesic approximation) in terms of geodesic motion in a certain effective smooth spacetime. The other is based on the recently proposed "first law of binary black-hole mechanics" [40,41] (itself a limiting case of the generalized law established in [42]), which relates ADM properties of a helically symmetric binary system of post-Newtonian (PN) particles to the so-called "redshift" of the particles (see below). Neither frameworks is *a priori* guaranteed to correctly describe the strong-field dynamics in the black hole–particle system relevant to us, but some evidence suggests that they might (we return to discuss this point at the end of the section).

The said results, as they are stated in [16], apply to a particle in a circular equatorial orbit (stable or unstable) around a Kerr black hole, ignoring the dissipative piece of the gravitational interaction (or, more precisely, time-symmetrizing the gravitational perturbation, so that space-time admits a global helical symmetry). They express the Bondi[9] energy and angular momentum of that configuration, through $O(\eta^2)$, in terms of Detweiler's redshift variable [28]

$$\hat{z} := (\hat{u}^t)^{-1}, \qquad (91)$$

where u^t is the t component of the four-velocity on the circular orbit, and overhats, recall, denote properties of the GSF-corrected orbit. The usefulness of such relations is in the fact that a computation of $\hat{z}$ requires only GSF information for circular orbits, and there is no need to integrate from infinity. Such information is essentially accessible to existing GSF codes.

Following [16], let us formally expand the redshift $\hat{z}$ in powers of η, in the form

$$\hat{z} = z_0(\Omega) + \eta z_1(\Omega) + O(\eta^2), \qquad (92)$$

where $\Omega(= d\phi/dt)$ is the circular orbit's angular velocity,

$$z_0 = (1 - a\Omega)^{1/2}[1 + a\Omega - 3\Omega^{2/3}(1 - a\Omega)^{1/3}]^{1/2} \qquad (93)$$

is the geodesic limit of $\hat{z}$, and $\eta z_1(\Omega)$ is the $O(\eta)$ GSF correction, defined for a fixed value of Ω. According to Ref. [16], the Bondi energy and angular momentum of the circular-orbit binary are given, through $O(\eta^2)$, by

---

[9]First-law literature [40,41,43] usually alludes to ADM properties, which are defined even in helical symmetry within the PN context in which these works operate. In the context of black-hole perturbation theory, the first-law results should be interpreted as referring to *Bondi* properties. See also [44], where first-law results are formulated directly in terms of Bondi quantities for a black hole–particle system.

$$E_{\rm B}^{\rm sym} = M + \mu E_{\rm B}^{\rm p}, \qquad L_{\rm B}^{\rm sym} = Ma + \mu L_{\rm B}^{\rm p}, \qquad (94)$$

where

$$E_{\rm B}^{\rm p} = \tilde{z} - \Omega \frac{d\tilde{z}}{d\Omega}, \qquad L_{\rm B}^{\rm p} = -\frac{d\tilde{z}}{d\Omega}, \qquad (95)$$

with

$$\tilde{z}(\Omega) = z_0(\Omega) + \frac{1}{2}\eta z_1(\Omega) + O(\eta^2). \qquad (96)$$

The label "sym" is to remind us that these Bondi properties are defined in a time-symmetrized ("half-retarded-plus-half-advanced") spacetime. The function $z_1(\Omega)$ explicitly determines $E_{\rm B}^{\rm sym}$ and $L_{\rm B}^{\rm sym}$ through $O(\eta^2)$.

We are now reaching the crux of our discussion. Consider a critical orbit, subject to the conservative GSF alone (dissipation ignored), which asymptotes to a certain unstable circular orbit at $\tau \to \infty$. Let $E_{\rm B}^{\rm sym}(u)$ and $L_{\rm B}^{\rm sym}(u)$ be the Bondi energy and angular momentum of the corresponding time-symmetrized spacetime, with u a suitable retarded-time coordinate. At $u \to \infty$, these quantities must approach the corresponding Bondi quantities of the asymptotic circular-orbit configuration, as given in Eq. (94). Furthermore,

$$E_{\rm B}^{\rm sym}(u \to \infty) = E_{\rm ADM}, \qquad L_{\rm B}^{\rm sym}(u \to \infty) = L_{\rm ADM}, \qquad (97)$$

where on the right-hand side we have the ADM properties of the physical ("retarded") critical-orbit spacetime. [That this must be the case follows from $E_{\rm B}^{\rm sym}(u \to \infty) = E_{\rm ADM}^{\rm sym} - \mathcal{F}^+ = E_{\rm ADM}^{\rm sym} - \mathcal{F}^- = E_{\rm ADM}$, where $\mathcal{F}^+$ and $\mathcal{F}^-$ are the total energies flowing, respectively, outward at future null infinity and inward at past null infinity, in the time-symmetrized setup where $\mathcal{F}^+ = \mathcal{F}^-$. A similar argument applies to the angular momentum.] As a result, we can write $E_{\rm ADM} = M + \mu E_{\rm ADM}^{\rm p}$ and $L_{\rm ADM} = M + \mu L_{\rm ADM}^{\rm p}$ [as in Eqs. (71)], with

$$E_{\rm ADM}^{\rm p} = \tilde{z} - \Omega \frac{d\tilde{z}}{d\Omega}, \qquad L_{\rm ADM}^{\rm p} = -\frac{d\tilde{z}}{d\Omega}. \qquad (98)$$

These expressions relate the ADM properties of the physical critical-orbit configuration to the redshift of the asymptotic circular orbit when dissipation is ignored.

The conservative GSF shift $\delta L_{\rm ADM}^{\rm p}(E_{\rm ADM}^{\rm p})$ [recall Eq. (72)] may now be obtained simply by considering the $O(\eta)$ piece of $L_{\rm ADM}^{\rm p}$ in Eq. (98), for a fixed $E_{\rm ADM}^{\rm p}$. Equations (98) with (96) immediately give us the $O(\eta)$ piece of $L_{\rm ADM}^{\rm p}$ for a fixed *angular velocity*: $\delta^{(\Omega)}L_{\rm ADM}^{\rm p} = -(\eta/2)dz_1/d\Omega$, where we introduced the operator $\delta^{(X)}$ to denote a linear variation with respect to η at fixed X. To obtain the shift at fixed *energy*, $\delta L_{\rm ADM}^{\rm p} \equiv \delta^{(E)}L_{\rm ADM}^{\rm p}$, we write

$$\delta L_{\rm ADM}^{\rm p} = \delta^{(\Omega)}L_{\rm ADM}^{\rm p} + \frac{dL_{\rm ADM}^{\rm p}}{d\Omega}\delta^{(E)}\Omega$$

$$= \delta^{(\Omega)}L_{\rm ADM}^{\rm p} - \frac{dL_{\rm ADM}^{\rm p}}{dE_{\rm ADM}^{\rm p}}\delta^{(\Omega)}E_{\rm ADM}^{\rm p}$$

$$= \delta^{(\Omega)}L_{\rm ADM}^{\rm p} - \Omega^{-1}\delta^{(\Omega)}E_{\rm ADM}^{\rm p}, \qquad (99)$$

where in the second line we used $\delta^{(E)}\Omega = -(d\Omega/dE_{\rm ADM}^{\rm p})\delta^{(\Omega)}E_{\rm ADM}^{\rm p}$, and in the third line we applied $dL_{\rm ADM}^{\rm p}/dE_{\rm ADM}^{\rm p} = \Omega^{-1}$, which is valid for any circular geodesic (omitting subdominant terms in η). From Eqs. (98) and (96) we find $\delta^{(\Omega)}E_{\rm ADM}^{\rm p} = (\eta/2)[z_1 - \Omega(dz_1/d\Omega)]$, and substituting this with the above result for $\delta^{(\Omega)}L_{\rm ADM}^{\rm p}$, we arrive at the simple expression

$$\delta L_{\rm ADM}^{\rm p} = -\frac{\eta}{2\Omega}z_1. \qquad (100)$$

Note that in the analysis leading to Eq. (100) we have not assumed anything about the spin a of the central black hole, so the result should apply in general (and suggests an interesting new interpretation of z_1 in terms of a shift in the critical value of the angular momentum). In the extremal case, $\Omega = 1/2 + O(\epsilon)$, so at leading order in ϵ we obtain

$$\delta L_{\rm ADM}^{\rm p}(E) = -\eta Z_1(E), \qquad (101)$$

where

$$Z_1(E) := \lim_{\epsilon \to 0} z_1(\Omega(E; \epsilon), \epsilon), \qquad (102)$$

with the limit taken *at fixed energy E*. Here, for clarity, we have made explicit the functional dependence of z_1 and Ω on ϵ, and have parametrized the circular orbits by their geodesic energy E, noting that the difference between E and $E_{\rm ADM}^{\rm p}$ is subdominant in Eq. (101). Indeed, in practice, $Z_1(E)$ may be evaluated by considering a sequence of circular *geodesics* of diminishing ϵ (and a fixed E).

Equation (101) may now be used with Eq. (72) to obtain the sought-for relation

$$\delta \check{L}_\infty^{\rm cons}(E) = E^2 + 1 - Z_1(E), \qquad (103)$$

which may then be used in place of (70) in both conditions (69) and (81). The relation (103) relieves us from the need to restrict attention to particles coming from infinity, which is why we have used in it the argument E in place of E_∞. The energy may now take any value $E > E_{\rm isco}(= 1/\sqrt{3})$, and the conditions (69) and (81) may be evaluated for all corresponding orbits.

We may also use Eq. (101) directly in conjunction with Eq. (73), to write the censorship condition (in the generic case) in the remarkably simple form

$$Z_1(E) \geq \frac{1}{2}(E^2 - 1). \qquad (104)$$

Overspinning is averted if and only if this inequality holds for all $E > E_{\text{isco}}$. The evaluation of the condition (104) requires only redshift information on unstable circular orbits, evaluated at the extremal limit $\epsilon \to 0$ with E held fixed.

It should be emphasized that the applicability of the theoretical framework underpinning Eq. (95) is yet to be rigorously established within our black hole–perturbative context. There is, however, accumulating evidence to suggest it. The first-law framework has been tested exhaustively against results from PN theory in both the Schwarzschild [43] and Kerr [16] cases, and, more importantly for us, it has been shown to precisely reproduce certain rigorous GSF results in the strong-field regime, at least in the Schwarzschild case [43]. Further reassurance is provided, in the Kerr case, by the agreement between the first-law framework and the perturbative Hamiltonian one [16]. Ideally, $\delta \check{L}_\infty^{\text{cons}}$ should be evaluated via both Eq. (103) and the more rigorous GSF formula (70). Indeed, if nothing else, a demonstrated agreement between these two expressions would lend a strong support to the validity of both first-law and Hamiltonian frameworks in the strong-field regime.

VI. ASSESSMENT OF NUMERICAL TASK

A direct evaluation of the censorship condition (69) requires the function $\delta L_\infty^{\text{cons}}(E_\infty)$, which involves the conservative piece of the GSF along critical equatorial geodesics on a Kerr background, in the extremal limit. To explore the case of fine-tuned orbits, via the condition (81), would require, in addition, a calculation of the radiative fluxes (either at infinity or down the event horizon) for unstable circular equatorial geodesics, again in the extremal limit. Finally, to evaluate the overspinning condition in its alternative form (104) demands redshift data for these same unstable circular geodesics. To the best of our knowledge, these numerical tasks are beyond the capability of existing GSF codes—though, perhaps, not by much. We think that custom-built codes to produce the necessary data could be developed through relatively mild adaptations of existing codes. In this section we review relevant numerical methods that are currently available, and discuss how they should be customized.

There now exist two computational frameworks for strong-field GSF calculations in Kerr geometry. One, by Dolan and collaborators [45], tackles the metric perturbation equations in the time-domain (TD), via a direct numerical evolution of the hyperbolic set of linearized Einstein's equations in the Lorenz gauge. A judiciously designed "puncture" scheme is applied to extract the correct regular piece of the metric perturbation, from which the local GSF is calculated along the orbit. The second

framework, due to Shah and collaborators [46,47], is based on Teukolsky's perturbation formalism: The relevant perturbation equations are decomposed into Fourier-harmonic modes and tackled numerically mode by mode, in the frequency domain (FD). The GSF is then reconstructed using a recently formulated mode-sum regularization procedure [48]. Both methods are underpinned by (the same) rigorous theory, and have been tested against each other and against results from PN theory. All applications so far considered the GSF on fixed geodesic orbits, neglecting backreaction on the orbit.

The above two methods are somewhat complementary with respect to the range of problems they can tackle. The FD method does best with bound-orbit configurations, where the perturbation field admits a discrete spectrum. It is not immediately clear how to apply the method to orbits that come from infinity. The TD method, on the other hand, can handle all types of orbits equally well. The special case of *circular* orbits can be tackled by both methods, but much more efficiently in the FD, thanks to the simple spectrum of radiation from such orbits.

Let us consider first the more straightforward of the aforementioned numerical tasks: calculation of energy and angular-momentum fluxes from circular orbits (for the fine-tuning case). This is a standard calculation, most efficiently performed by solving Teukolsky's equation in the FD, either numerically [49] or via the semianalytical method of Mano, Suzuki and Takasugi [50]. Usually calculations focus on *stable* circular orbits as the source of radiation, but exactly the same techniques should be applicable without change to unstable orbits below the ISCO; the stability properties of the orbit are immaterial here. The only potential complication comes from the need to evaluate the fluxes in the extremal limit. Care would need to be taken in correctly identifying an inner "wave zone" in which to impose the boundary conditions, for each finite value $\epsilon \ll 1$. It may prove convenient to work with a rescaled radial coordinate [such as $\bar{r} := (r - R_{\text{eh}})/\epsilon$] in order to better resolve the near-horizon wave dynamics. However, we do not see any issues of principle to hinder such a calculation, and it could be based on any of the existing platforms, such as the one by Shah *et al.* [47].

We next turn to the calculation of $\delta L_\infty^{\text{cons}}(E)$ using the form (103). This requires an evaluation of the redshift function $z_1(\Omega; \epsilon)$ [recall Eq. (102)] on unstable circular orbits, at the limit $\epsilon \to 0$. Calculations of z_1 for *stable* (or otherwise near-ISCO) geodesic orbits have been performed using both FD [16,47] and TD [16] methods, in a nonextremal Kerr geometry. In the Schwarzschild case, such a calculation was performed (in the FD) even for unstable orbits, reaching very near the light ring [51]. The challenge is to extend these calculations to the near-extremal Kerr case in order to take the extremal limit. The issues here are similar to the ones affecting fluxes. In the FD, one would need to carefully set inner boundary conditions, and also

carefully monitor the convergence of the multipole mode sum, particularly at large energy (lessons can be learned from the Schwarzschild analysis of [51]). Some development, tests, and a careful error analysis would be required, but the problem seems to us perfectly tractable.

If one is satisfied with the level of rigor provided by the first-law and Hamiltonian formulations, then no further calculations would be needed: The question of overspinning, for both generic and fine-tuned orbits, can be answered based on numerical data pertaining to circular orbits only. However, to establish a full confidence in the results, a direct evaluation of $\delta L_\infty^{\rm cons}(E)$ via Eq. (70) would be required. Since this involves GSF data on unbound orbits, a TD treatment is preferable. So far, TD calculations of the GSF in Kerr have been limited to circular (and equatorial) orbits in a nonextremal geometry [52]. However, it should not be hard to implement orbits that arrive from infinity, including critical orbits in the extremal case, at least in principle. To achieve this, certain technical details would need to be addressed. In the rest of this section we give an assessment of the challenges involved.

First, to probe the extremal limit, we would need GSF data for a sequence of spacetimes approaching extremality. In the current code [52], accuracy has been observed to degrade with increasing a. In the circular-orbit case the code can handle spins up to $a = 0.7$–$0.8M$, beyond which the loss of accuracy is rapid. The cause of the accuracy degradation is not quite clear yet, but work towards a full diagnosis of the problem, and towards its resolution, is now in progress. It appears likely that the code could be pushed to very high spins with only moderate effort.

Second, one would need to implement plunging (geodesic, critical) orbits. In a TD framework this would entail only a minor coding effort; the basic code architecture should remain intact. Two of the technical details that would need addressing are (1) the handling of the auxiliary world tube (see [45] for details), which would now be moving to track the radial motion of the particle; and (2) the treatment of "junk radiation" (again, see [45] for details), whose problematic effect is expected to be more pronounced than for circular orbits. A method for tackling the latter problem has been developed and successfully implemented in the electromagnetic case, by Zimmerman *et al.* [9]. It remains to test its performance in the gravitational case.

Perhaps the most significant remaining problem is that of the $m = 1$ mode instability—an issue identified and thoroughly analyzed in [45]. In numerical experiments, this particular azimuthal mode of the Lorenz-gauge metric perturbation appeared to develop a linear instability at late time, which so far could not be cured. The seed of the instability appears to be a certain nonradiative, Lorenz-gauge mode, which is perfectly regular on each time slice. Various methods have been tried in attempt to filter that mode out of the numerical solutions,

so far without much success. A simple filter can be applied in the circular-orbit case [45], giving a satisfactory *ad hoc* solution, but the method would not work for noncircular orbits. We are aware of at least two parallel efforts, by two groups, to obtain a more fundamental solution to the problem, and we remain optimistic that the issue will be resolved soon.

The problem of $m = 1$ mode instability is entirely avoided within a third computational framework, now being developed by a collaboration involving one of the authors (L. B.). The new method combines the simplicity of the Teukolsky formulation with the utility and flexibility of the TD approach. It essentially involves a TD implementation of the Teukolsky equation, together with a scheme for constructing the GSF in a certain (non-Lorenz) gauge. When completed, the code will offer a most natural tool for performing the calculation required here.

VII. SUMMARY AND CONCLUDING REMARKS

We studied the scenario in which a massive (spinless) particle is dropped into a nearly extremal Kerr black hole on an equatorial-plane trajectory. For this scenario, we presented a systematic analysis of the censorship condition at first order in the self-force approximation. One of our main goals was to determine what GSF information, precisely, would be needed in order to provide a definitive answer to the question of whether an over-extremal black hole was a possible outcome (within the classical theory). We achieved this by formulating concrete, necessary and sufficient conditions for overspinning to be averted; these are given explicitly in terms of GSF quantities, ready for numerical evaluation.

Along the way, we have established several interesting results:

(i) When the GSF is ignored, overspinning can be achieved in a certain open domain of the parameter space, mapped here precisely for the first time.

(ii) Overspinning is possible (when the GSF is ignored) only with particles thrown in from infinity. For any value of the initial energy at infinity, overspinning can be achieved by choosing the particle's rest mass and angular momentum from within certain open intervals, as prescribed in the last sentence of Sec. II, below Eq. (30).

(iii) In the full-GSF case, a sufficient and necessary censorship condition for generic orbits may be formulated on the one-parameter family of critical geodesics. That condition is sensitive only to the *conservative* piece of the GSF.

(iv) A more general condition may be formulated to encompass also fine-tuned orbits (ones whose parameters are exponentially fine-tuned to produce an adiabatic evolution along the global attractor). That condition involves also the radiative fluxes from unstable circular geodesics.

(v) The conservative GSF effect may be reformulated in terms of the "redshift" variable. This results in an alternative form of the censorship conditions, which involves only perturbative quantities (redshift and radiative fluxes) calculated on unstable circular geodesics.

Our main results are expressed in Eqs. (69), (81) and (104). Equation (69) [with (70)] is the censorship condition for generic orbits. It required as input $\delta\check{L}^{\mathrm{cons}}_{\infty}(E_{\infty})$, the conservative GSF correction to the critical value of the angular momentum at a fixed E_{∞}. The condition is a *sufficient* one in the sense that its validity for all $E_{\infty} \geq 1$ would imply that censorship is protected in our scenario, at least for sufficiently small values of the particle's rest mass (this last caveat is to remind us that our analysis and conditions are formulated within the first-order self-force approximation). Equation (69) is also a *necessary* condition, in the sense that its violation for any value $E_{\infty} \geq 1$ would mean a direct infringement of cosmic censorship.

Equation (81) [with (74) or (79)] is the more general censorship condition that covers also fine-tuned orbits. Its evaluation requires, in addition to $\delta\check{L}^{\mathrm{cons}}_{\infty}$, also the radiative quantities $\check{\mathcal{E}}^{+}$ and $\check{\mathcal{W}}^{+}_{(\mathrm{qc})}$ [or $\check{\mathcal{E}}^{-}$ and $\check{\mathcal{W}}^{-}_{(\mathrm{qc})}$] associated with the adiabatic evolution along the global attractor. Without fine-tuning, $\check{\mathcal{E}}^{\pm}$ and $\check{\mathcal{W}}^{+}_{(\mathrm{qc})}$ vanish at relevant order, and (81) reduces to the generic condition (69). With fine-tuning, the condition (81) should be evaluated on the two-parameter space of initial and final energies, $\{E_{\infty}, E_f\}$. It is both sufficient and necessary, in the same sense as (69).

Finally, Eq. (104) is a reformulation of (69) in terms of the redshift variable $Z_1(E)$, calculated on unstable circular geodesics in the extremal limit (taken with fixed E). The more general condition (81) may also be formulated in terms of the redshift, by substituting for $\delta\check{L}^{\mathrm{cons}}_{\infty}$ using Eq. (103). This alternative form is more readily amenable to numerical evaluation, because it requires only circular-orbit information. An additional advantage is that (104) is applicable to any value $E_{\infty} > E_{\mathrm{isco}}$ of the initial energy, without the restriction that particles have to be sent from infinity. However, the redshift formulation relies on some layers of nonrigorous theory, so a direct evaluation of the conservative GSF effect, via Eq. (70), would be desirable as a check.

We are unable to predict, in advance of a numerical calculation, whether our censorship conditions (69) or (104) hold. We are not familiar with an argument to suggest even the *signs* of the terms $\delta\check{L}^{\mathrm{cons}}_{\infty}$ and Z_1 appearing in these conditions. However, if we *assume* that the generic-orbit conditions (69) [or (104)] are satisfied, then it is possible to conclude that censorship is protected also for fine-tuned orbits [i.e., the inequality (81) holds], assuming only a certain plausible lower bound on the flux ratio $\mathcal{R}(E)$. Specifically, one can show, with the help of Eqs. (84) and (89), that (81) is satisfied for all $E_{\infty} > E_f \geq E_{\mathrm{isco}}$ if

$\mathcal{R}(E) \geq -\frac{1}{3}$. This lower bound lies comfortably below the ISCO value $\mathcal{R}(E_{\mathrm{isco}}) \sim -0.1$ estimated in Ref. [53], and we expect unstable circular orbits to be *less* superradiant than the ISCO [i.e, have $\mathcal{R}(E) > \mathcal{R}(E_{\mathrm{isco}})$] on account of their frequency being larger than the ISCO frequency (recall also $\mathcal{R} > 0$ for $E > \frac{2}{\sqrt{3}}$). If a numerical calculation of $\mathcal{R}(E)$ confirms this expectation, it would mean that fine-tuning disfavors overspinning. In that case, establishing the simple inequality (69) [or (104)] would suffice for ruling out the overspinning scenario.

We remind, finally, that our analysis concentrates on capture orbits that are equatorial and prograde ("corotating"). Ruling out overspinning in this scenario would *not* automatically preclude overspinning by nonequatorial prograde orbits [*retrograde* ("counterrotating") orbits cannot overspin, with or without the self-force]. Such nonequatorial orbits require a separate analysis.

ACKNOWLEDGMENTS

APPENDIX A: ADM ENERGY AND ANGULAR MOMENTUM

This appendix gives a detailed derivation of Eqs. (44) and (45). We consider two relativistic point particles in flat space, representing the black hole–particle system at an infinite separation. We are given the two rest masses, M and μ, and the particle's energy $\mu E_{\infty}(> \mu)$ and angular momentum μL_{∞} in a reference frame attached to the mass M. The mass M has an intrinsic spin Ma perpendicular to the orbital plane, and the mass μ is spinless. Our goal is to obtain the system's total energy and angular momentum in the CoM frame.

First, we note that, in the limit of infinite separation, both black-hole frame and CoM frame are inertial, and they are related via a simple Lorentz boost. Let $x^{\alpha} = (t, x, y, z)$ be a Cartesian frame centered at M, so that the spin of M is aligned with the z direction, and the particle's orbit lies in the x-y plane. Denote the four-momenta of μ and M in the black-hole frame by

$$p^\alpha_{(\mu)} = (\mu E_\infty, p^x_{(\mu)}, p^y_{(\mu)}, 0),$$
$$p^\alpha_{(M)} = (M, 0, 0, 0). \tag{A1}$$

The magnitude of particle's three-momentum satisfies

$$|\mathbf{p}_{(\mu)}|^2 = (p^x_{(\mu)})^2 + (p^y_{(\mu)})^2 = \mu^2(E^2_\infty - 1). \tag{A2}$$

The CoM system $\tilde{x}^\alpha$ is related to x^α via a Lorentz boost Λ^α_β in the x–y plane. Let $\tilde{p}^\alpha_{(\mu)} = \Lambda^\alpha_\beta p^\beta_{(\mu)}$ and $\tilde{p}^\alpha_{(M)} = \Lambda^\alpha_\beta p^\beta_{(M)}$ denote the 4-momenta of μ and M in $\tilde{x}^\alpha$. The CoM condition,

$$\tilde{\mathbf{p}}_{(M)} + \tilde{\mathbf{p}}_{(\mu)} = 0, \tag{A3}$$

yields two nontrivial equations for the two boost parameters $\beta^x = v^x/c$ and $\beta^y = v^y/c$ (where v^x and v^y are components of the boost velocity). One finds $\beta^i = p^i_{(\mu)}(M + \mu E_\infty)^{-1}$ for $i = x, y$. Thus, using (A2),

$$\beta = [(\beta^x)^2 + (\beta^y)^2]^{1/2} = \frac{\eta(E^2_\infty - 1)^{1/2}}{1 + \eta E_\infty}, \tag{A4}$$

where $\eta = \mu/M$.

Now that we have at hand the boost $\Lambda^\alpha_\beta(\beta)$, the relativistic energies of μ and M in the CoM system are obtained as $\tilde{p}^0_{(\mu)} = \Lambda^0_\beta p^\beta_{(\mu)}$ and $\tilde{p}^0_{(M)} = \Lambda^0_\beta p^\beta_{(M)}$. The sum of these two energies is the total, ADM energy of the system. A short calculation gives

$$E_{\mathrm{ADM}} = \tilde{p}^0_{(\mu)} + \tilde{p}^0_{(M)} = M(1 + 2\eta E_\infty + \eta^2)^{1/2}. \tag{A5}$$

This result is valid for any mass ratio η. For $\eta \ll 1$, an expansion in η through $O(\eta)$ gives Eq. (44).

To obtain the total angular momentum in the CoM system, we need first to relate the particle's CoM position $\tilde{\mathbf{x}}_{(\mu)}$ and black hole's CoM position $\tilde{\mathbf{x}}_{(M)}$ to their separation $\mathbf{x}$ in the black-hole frame. This is achieved by solving the CoM condition

$$\tilde{\mathbf{x}}_{(\mu)} \tilde{p}^0_{(\mu)} + \tilde{\mathbf{x}}_{(M)} \tilde{p}^0_{(M)} = 0, \tag{A6}$$

simultaneously with $|\tilde{\mathbf{x}}_{(\mu)}| + |\tilde{\mathbf{x}}_{(M)}| = |\tilde{\mathbf{x}}|$, where $\tilde{x}^i = \Lambda^i_\beta x^\beta$ $(i = x, y)$ is the separation in the CoM frame. The particle's CoM orbital angular momentum is then given by $\tilde{L}_{(\mu)} = \tilde{x}_{(\mu)} \tilde{p}^y_{(\mu)} - \tilde{y}_{(\mu)} \tilde{p}^x_{(\mu)}$, and similarly for the mass M. The total, ADM angular momentum of the system (with respect to the CoM) is the sum of $\tilde{L}_{(\mu)}$, $\tilde{L}_{(M)}$ and the spin angular momentum. A short calculation, using Eqs. (A2) and (A4) and the relation $\mu L_\infty = x p^y_{(\mu)} - y p^x_{(\mu)}$, gives

$$L_{\mathrm{ADM}} = M(a + \mu L_\infty/E_{\mathrm{ADM}}). \tag{A7}$$

This is valid for any η. Substituting for E_{ADM} and expanding through $O(\eta^2)$ produces Eq. (45).

APPENDIX B: SHIFT IN THE CRITICAL VALUE OF THE ANGULAR MOMENTUM DUE TO THE FULL SELF-FORCE

In this appendix we derive Eq. (51), which describes the GSF-induced shift δL_∞ in the critical value of the angular momentum at infinity for a given energy at infinity, at leading order in η and in ϵ. In Sec. III D 1 we obtained an expression for δL_∞ [Eq. (49)] under the (nonphysical) simplifying assumption that the GSF has no dissipative piece. Here we restore dissipation and calculate the shift δL_∞ caused by the full GSF.

In our analysis we only require the leading order [$O(\eta)$] of δL_∞. As explained in the text, for that purpose it is sufficient to ignore fine-tuning and assume the changes in (specific) energy and angular momentum along the critical orbit are $O(\eta \ln \eta)$ at most.

Consider, then, a critical orbit parametrized by $E_\infty(\geq 1)$, and an arbitrary moment $\tau = \tau_0$ after the orbit had settled into quasicircular motion, but before $\Delta E(\tau_0)$ and $\Delta L(\tau_0)$ have accumulated $O(1)$ changes [specifically, assume $\Delta E(\tau_0), \Delta L(\tau_0) = O(\eta \ln \eta)$ at most]. Evaluate the full-GSF radial equation of motion (35) at time τ_0, subject to the near-circularity condition $dr/d\tau = O(\eta)$, substituting for $\hat{E}(\tau_0)$ and $\hat{L}(\tau_0)$ from Eq. (37), and replacing L_∞ with $L_c(E_\infty) + \delta L_\infty(E_\infty)$. At leading order in η and in ϵ one obtains the solution

$$\delta L_\infty(E_\infty) = \Delta W(\tau_0), \tag{B1}$$

where ΔW is the GSF integral defined in Eq. (50). This result can only make sense if (i) the expression on the right-hand side is in fact independent of τ_0 at leading order, and (ii) the quantity $\Delta W(\tau_0) = 2\Delta E(\tau_0) - \Delta L(\tau_0)$ remains of $O(\eta)$ even for τ_0 large enough that the individual terms $\Delta E(\tau_0)$ and $\Delta L(\tau_0)$ are already of $O(\eta \ln \eta)$. We now argue that both conditions are satisfied.

To make the argument, let us split $\Delta W(\tau_0)$ into an "approach" piece $\Delta W(\tau_c)$, and a "quasicircular" piece

$$\Delta W(\tau_c, \tau_0) := \Delta W(\tau_0) - \Delta W(\tau_c)$$
$$= -\eta^{-1} \int_{\tau_c}^{\tau_0} (2F_t + F_\phi) d\tau. \tag{B2}$$

Here the end-of-approach time τ_c can be taken to be any moment $\tau_c < \tau_0$ after the orbit had settled into quasicircular motion, in the sense that $\dot{r}(\tau_c) = O(\eta)$ at most. We will now show that

$$\Delta W(\tau_c, \tau_0) = O(\epsilon)O(\eta \ln \eta) \tag{B3}$$

at most, for any choice of τ_0 and of τ_c. Assuming $\epsilon|\ln \eta| \ll 1$, this would mean that $\Delta W(\tau_0)$ is dominated

by its approach piece $\Delta W(\tau_c)$, so that (i) $\Delta W(\tau_0)$ does not depend on τ_0 at leading order, and (ii) $\Delta W(\tau_0) \simeq \Delta W(\tau_c) = O(\eta)$ as argued above.

To establish the scaling in (B3), use the orthogonality relation $\hat{u}^\alpha F_\alpha = 0$ to write $F_t + \Omega F_\phi = -(u^r/u^t)F_r$, where $\Omega = u^\phi/u^t$, and the replacement $\hat{u}^\alpha \to u^\alpha$ does not affect the expression at leading order in η. We wish to evaluate this relation for $\tau_c \leq \tau \leq \tau_0$, when the orbit is quasicircular. At leading order in η, $\Omega = \Omega(E; \epsilon)$ is then equal to the angular velocity of an unstable circular geodesic of specific energy $E = \hat{E}(\tau)$. At fixed energy, the angular velocity admits the small-ϵ expansion $\Omega = \frac{1}{2} - \frac{1}{4}b(E)\epsilon + O(\epsilon^2)$ with $b = 6E(6E^2 - 2)^{-1/2}$ [see Eqs. (87) and (88)]. Thus, omitting terms that are subdominant in ϵ, the integrand in Eq. (B2) is $2F_t + F_\phi = -2(u^r/u^t)F_r + \frac{1}{2}b\epsilon F_\phi$, or, equivalently,

$$2F_t + F_\phi = -2(u^r/u^t)F_r - b\epsilon F_t. \tag{B4}$$

Let us denote the contributions to $\Delta W(\tau_0, \tau_c)$ from the first and second terms on the right-hand side of (B4) by $\Delta W_{(r)}$ and $\Delta W_{(t)}$, respectively. In what follows we consider each of the two contributions in turn.

Start with $\Delta W_{(r)}$, given by

$$\Delta W_{(r)} = (2/\eta) \int_{r(\tau_c)}^{r(\tau_0)} (F_r/u^t) dr. \tag{B5}$$

Note $r(\tau_0) - r(\tau_c) = O(\epsilon)$, since both radii belong to unstable circular geodesics. From Eq. (43) we recall that for $|\dot{r}| \ll 1$ the radial component F_r is dominated by its conservative piece, $F_r^{cons} \simeq F_{0r}(r)$, which is approximately constant within the integration domain and may therefore (as we only keep track of the leading term in ϵ) be pulled out of the integral. A simple calculation gives $1/u^t \propto \epsilon$ (at fixed E), and this factor can likewise be taken out of the integral. We thus obtain the scaling $\Delta W_{(r)} \sim \epsilon^2 F_{0r}/\eta$, where F_{0r} is evaluated, e.g., at $r = r(\tau_c)$, the end of approach. It remains to determine the ϵ-scaling of F_{0r}. Numerical evidence, to be presented elsewhere [36], suggests the scaling $F_{0r} \sim \epsilon^{-1}$. This is consistent with what one would obtain by assuming that the GSF components in a normalized coordinate basis are finite: $F_{0r} = F_{0\hat{r}}(g^{rr})^{-1/2} \sim \epsilon^{-1}$ (or smaller), assuming the normalized component $F_{0\hat{r}}$ is finite and noting $g^{rr} \sim \epsilon^2$. Assuming, therefore, that F_{0r} does not diverge faster than $\sim \epsilon^{-1}$, and recalling $F_{0r} \propto \eta^2$ as usual, we arrive at

$$\Delta W_{(r)} = O(\epsilon)O(\eta) \tag{B6}$$

(or smaller).

Next, consider the contribution $\Delta W_{(t)}$:

$$\Delta W_{(t)} = (\epsilon/\eta) \int_{\tau_0}^{\tau_c} b(E)F_t d\tau = -\epsilon \int_{E(\tau_c)}^{E(\tau_0)} b(E) dE$$
$$= -\epsilon \sqrt{6E^2 - 2}\Big|_{E(\tau_c)}^{E(\tau_0)}. \tag{B7}$$

In the second equality we have used $F_t = -\mu dE/d\tau$, and in the third we have substituted for $b(E)$ and integrated explicitly. Since the energy difference $E(\tau_0) - E(\tau_c)$ is at most of $O(\eta \ln \eta)$, we conclude that

$$\Delta W_{(t)} = O(\epsilon)O(\eta \ln \eta) \tag{B8}$$

(or smaller).

The combination of Eqs. (B6) and (B8) leads to the scaling stated in Eq. (B3). The upshot is that the contribution to $\Delta W(\tau_0)$ from the quasicircular part, $\Delta W(\tau_c, \tau_0)$, is negligible compared to the contribution from the approach part, $\Delta W(\tau_c) = O(\epsilon^0)O(\eta)$ (assuming $\epsilon|\ln \eta| \ll 1$). In other words, the GSF integral $\Delta W(\tau_0)$ in Eq. (B1) may be truncated at the end-of-approach time τ_c, with the latter taken to be any instance after the orbit had settled into quasicircular motion [but before the specific energy has accumulated $O(1)$ changes]. This establishes Eq. (51) in the main text.

APPENDIX C: RADIATION FROM TRANSITION TO PLUNGE AND FINAL PLUNGE

In this appendix we argue that the term $\mathcal{W}^+_{(end)}$ in Eq. (64) is subdominant for $\eta, \epsilon \ll 1$ and may therefore be dropped within our leading-order analysis. Recall $\mathcal{W}^+_{(end)} = 2\mathcal{E}^+_{(end)} - \mathcal{L}^+_{(end)}$, where $\mathcal{E}^+_{(end)}$ and $\mathcal{L}^+_{(end)}$ are the energy and angular momentum radiated out to infinity during the transition from adiabatic inspiral along the attractor to a final plunge into the black hole, and during the plunge itself.

Critical orbits may transit into plunge in one of two ways: If the orbit is perfectly fine-tuned, the transition will occur around the location of the ISCO, and will then be similar to the transition at the end of a physical adiabatic inspiral (on a *stable* orbit) that has already been studied in detail [32,33,35,54]. With any less than perfect fine-tuning, the particle will slide off the peak of the effective potential and into the black hole before the ISCO is reached (cf. Fig. 4). It is reasonable to expect the former scenario (transition through the ISCO) to yield the maximal radiation output, because (i) orbits linger much longer around the ISCO location, where the potential is very flat, than they do around the peak of the potential, and (ii) the remaining distance to the horizon is maximal when the transition is at the ISCO. Below we start by looking at the "worse case" scenario of transition through the ISCO, and argue that, even in that case, $\mathcal{W}^+_{(end)}$ is negligible in Eq. (64). We then also examine the more generic scenario of a transition from

an unstable orbit, and, as expected, arrive at a similar conclusion.

1. Plunge from the ISCO

For a transition through the ISCO, we use results by Ori and Thorne [32], who studied radiation from the transition regime at the end of a quasicircular inspiral. An inspection of their analysis reveals that the main results are insensitive to whether the particle arrives at the transition regime along stable or unstable orbits. In particular, their Eq. (3.26) for the deficits in (specific) energy and angular moment over the entire transition should hold in either case. We write it here in the form

$$(\Delta L)_{\text{trans}} = -\mathcal{A}(a)\eta^{4/5},$$
$$(\Delta E)_{\text{trans}} = -\Omega_{\text{isco}}\mathcal{A}(a)\eta^{4/5}, \qquad (\text{C1})$$

where $\mathcal{A}(a)$ is a certain (positive) function of the spin a only, given explicitly in [32]. These expressions hold for any a, at leading order in $\eta \ll 1$. The factor $\eta^{4/5}$ arises, essentially, from the fact that the transiting orbit spends an amount of proper time $\propto \eta^{-1/5}$ whirling around the ISCO location on a nearly circular orbit [cf. Eq. (3.20) of [32]].

The function $\mathcal{A}(a)$ involves a certain a-dependent dimensionless factor, denoted $\dot{\mathcal{E}}$ in [32] (not to be confused with our $\mathcal{E}^{\pm}$), which describes corrections to the leading-order quadrupole emission formula for circular orbits, and is to be determined numerically. Kesden [33] used numerical data by Hughes (from a code presented in [55]) to estimate, in the near-extremal case, $\dot{\mathcal{E}} \propto \epsilon^{2/3}$, a scaling previously suggested by Chrzanowski [56]. We have confirmed this scaling using much more accurate numerical results by M. van de Meent [37], to be presented elsewhere [36]. We also note that the scaling $\dot{\mathcal{E}} \propto \epsilon^{2/3}$ follows simply from a regularity assumption, namely that $dE/d\tau$ must remain bounded (and nonzero) even in the limit $\epsilon \to 0$: Noting that the Ori-Thorne function $\dot{\mathcal{E}}$ is defined with respect to coordinate time t (not proper time τ), and that $(d\tau/dt)_{\text{isco}} \propto \epsilon^{2/3}$, we obtain $\dot{\mathcal{E}} \propto (dE/d\tau)(d\tau/dt)_{\text{isco}} \propto \epsilon^{2/3}$. Assuming this scaling, and expanding the remaining a dependence of $\mathcal{A}(a)$ in ϵ, we obtain, at leading order in ϵ,

$$(\Delta L)_{\text{trans}} = -a_0 \epsilon^{-2/15}\eta^{4/5},$$
$$(\Delta E)_{\text{trans}} = -a_0 \Omega_{\text{isco}}\epsilon^{-2/15}\eta^{4/5}, \qquad (\text{C2})$$

with some (positive) numerical coefficient a_0 whose explicit value will not be needed here.

References [33,35] discuss the reason for the non-physical divergence of $(\Delta E)_{\text{trans}}$ and $(\Delta L)_{\text{trans}}$ when $\epsilon \to 0$ is taken with a fixed η, but this will not concern us here. We are instead interested in the combination $(\Delta W)_{\text{trans}} := 2(\Delta E)_{\text{trans}} - (\Delta L)_{\text{trans}}$, which, noting $\Omega_{\text{isco}} = \frac{1}{2} + O(\epsilon^{2/3})$ and keeping only the leading term, reads

$$(\Delta W)_{\text{trans}} = b_0 \epsilon^{8/15}\eta^{4/5}, \qquad (\text{C3})$$

with some (positive) numerical coefficient b_0. Assuming the usual balance between the local dissipative GSF and the flux of energy and angular momentum in gravitational waves emitted during the transition, we have $\mathcal{W}^{+}_{\text{trans}} = -\eta(\Delta W)_{\text{trans}}$ and thus

$$\mathcal{W}^{+}_{\text{trans}} = -b_0 \epsilon^{8/15}\eta^{9/5}. \qquad (\text{C4})$$

Now examine the magnitude of $\mathcal{W}^{+}_{\text{trans}}$ compared to that of other terms in the censorship condition (64). If $\epsilon > \eta$, we have $\epsilon^{8/15}\eta^{9/5} < \epsilon^{7/3}$, which, for $\epsilon \ll 1$, is much smaller than the ϵ^2 term in Eq. (64). If instead we have $\epsilon \leq \eta$, then $\epsilon^{8/15}\eta^{9/5} \leq \eta^{7/3}$, which, for $\eta \ll 1$, is much smaller than the $O(\eta^2)$ terms in that equation. The conclusion is that the contribution to $\mathcal{W}^{+}_{(\text{end})}$ from the transition regime, $\mathcal{W}^{+}_{\text{trans}}$, is always subdominant in Eq. (64) for $\eta, \epsilon \ll 1$.

It remains to assess the contribution to $\mathcal{W}^{+}_{\text{plng}}$ from the final plunge into the hole. In general (when the black hole is not near extremal) one identifies a final stage, extending smoothly from the transition regime, where radiation reaction is negligible and the orbit plunges into the black hole on a nearly geodesic trajectory [32,34]. The picture may change a little in the near-extremal case, because the radial velocity remains small, $|\dot{r}| \propto \epsilon \ll 1$, all the way to the horizon. This can mean that the conditions that define the transition regime never quite break down before the horizon is reached. In other words, there is a possibility that the particle crosses the horizon while still in the transition regime.[10]

That possibility can be assessed using Eq. (3.20) of [32], according to which the radial extent $(\Delta r)_{\text{trans}}$ of the transition regime, in the near-extremal case, is $\propto \epsilon^{4/15}\eta^{2/5}$, with a coefficient of order unity. This should be compared with the radial distance from the ISCO to the horizon, $\Delta r \simeq (2\epsilon)^{2/3}$ [recall Eqs. (10) and (13)]. It follows that for ϵ/η smaller than a number of order unity, the transition regime extends to the horizon. In such cases, the amplitude of $\mathcal{W}^{+}_{\text{trans}}$ in Eq. (C4) serves as an upper bound for the amplitude of $\mathcal{W}^{+}_{(\text{end})}$, and it follows immediately that the entire term $\mathcal{W}^{+}_{(\text{end})}$ is negligible in Eq. (64).

Let us then consider the case where the transition ends before the horizon is reached, so that a plunge regime is identifiable. In the plunge regime, by definition, the motion is very nearly geodesic, and the near-horizon analysis of Mino and Brink [34] should be applicable. Reference [34] obtained (among other things) an analytic expression for the energy output from the final plunge, by analyzing solutions to Teukolsky's equation in the near-horizon,

---

[10]This possibility was studied in some detail in Ref. [33]; see, in particular, Fig. 5 therein, in which "δ" is equivalent to our ϵ^2.

low-frequency, quadrupole approximations. They find (in our notation),[11,12]

$$\mathcal{E}^+_{\text{plung}} \propto \eta^2 \epsilon^5 (r_{\text{init}} - r_{\text{eh}})(E - \Omega_H L)^{-2} \qquad (C5)$$

at leading order in ϵ, where r_{init} is the radius at the start of the plunge, and r_{eh} is the horizon's radius (denoted R_{eh} in the main text). For a plunge following a transition at the ISCO, $r_{\text{init}} - r_{\text{eh}} = O(\epsilon^{2/3})$ and $E - \Omega_H L = O(\epsilon)$, giving $\mathcal{E}^+_{\text{plung}} = O(\eta^2 \epsilon^{11/3})$. Thus $\mathcal{E}^+_{\text{plung}}$ is strongly suppressed at small ϵ. Since the motion is nearly circular even during the plunge, we have $\mathcal{L}^+_{\text{plung}} \simeq \Omega^{-1} \mathcal{E}^+_{\text{plung}} \simeq 2 \mathcal{E}^+_{\text{plung}}$ and the radiated angular momentum is similarly suppressed. The combination $\mathcal{W}^+_{\text{plung}} := 2 \mathcal{E}^+_{\text{plung}} - \mathcal{L}^+_{\text{plung}}$ is even more strongly suppressed at small ϵ, and clearly contributes negligibly to $\mathcal{W}^+_{\text{(end)}}$.

In conclusion, we have found that, for a transition through the ISCO, $\mathcal{W}^+_{\text{(end)}} = \mathcal{W}^+_{\text{trans}} + \mathcal{W}^+_{\text{plunge}} \simeq \mathcal{W}^+_{\text{trans}}$ is always negligible in Eq. (64) for $\eta, \epsilon \ll 1$. That radiation from the transition to plunge should have a negligible effect on the conditions for overspinning was previously suggested by Kesden [33] and Harada and Kimura [35].

2. Plunge from an unstable circular orbit

This scenario is rather different from—and much simpler than—a transition through the ISCO. As the orbit is perturbed away from unstable equilibrium, its subsequent evolution is almost immediately controlled by the "geosedic" radial force (proportional to the derivative of the effective potential), and backreaction corrections become negligible. Let us state this point more precisely. Suppose $r = r_{\text{end}}$ is the radius at which the particle leaves the attractor (for concreteness, this may be chosen as the radius of the last turning point along the attractor). Note that the radial acceleration due to the geodesic potential is $\propto (r - r_{\text{end}})$, with a coefficient of order unity [since the

second derivative of the effective potential at r_{end} is of $O(\epsilon^0)$]. Since $r_{\text{end}} - r_{\text{eh}} = O(\epsilon)$, the magnitude of the geodesic radial acceleration is of $O(\epsilon)$ throughout much of the plunge. This should be compared with the magnitude of the radial self-acceleration, which is of $O(\eta \epsilon)$.[13] We may *define* the onset of plunge as the point where the geodesic acceleration takes over from the GSF in controlling the motion; this happens near a radius $r = r_{\text{plng}}$ satisfying $r_{\text{plng}} - r_{\text{end}} = O(\eta \epsilon)$. For $r \lesssim r_{\text{plng}}$, the motion is governed by the geodesic equation of motion (3) to a good approximation.

We wish to bound the magnitude of $\mathcal{W}^+_{\text{(end)}}$ sufficiently well to show that it contributes negligibly in Eq. (64). Let us first consider the contribution to $\mathcal{W}^+_{\text{(end)}}$ from the preplunge orbital segment $r_{\text{plng}} \leq r \leq r_{\text{end}}$. The proper-time interval along this segment is $\Delta \tau \sim (r_{\text{end}} - r_{\text{plng}})/\dot{r} = O(\eta^0 \epsilon^0)$ (at most), since $\dot{r} = O(\eta \epsilon)$ along the attractor, where the evolution is driven by radiation reaction. Hence, the change experienced by the specific energy and angular momentum along this segment is $\Delta E \simeq \Omega \Delta L = O(\eta)$. Recalling $\Omega = 1/2 + O(\epsilon)$, this gives $\Delta W = O(\epsilon \eta)$ and thus a contribution of $O(\epsilon \eta^2)$ to $\mathcal{W}^+_{\text{(end)}}$, negligible compared to the $O(\eta^2)$ terms in Eq. (64).

Next consider the contribution to $\mathcal{W}^+_{\text{(end)}}$ from the plunge segment $r_{\text{eh}} \leq r < r_{\text{plng}}$. Using the geodesic equation of motion (3), one finds that the proper-time interval along that segment is $\propto \ln[(r_{\text{end}} - r_{\text{eh}})/(r_{\text{end}} - r_{\text{plng}})] \sim \ln \eta$. The corresponding change in specific energy is $\Delta E = O(\eta \ln \eta)$. Since the radial velocity remains small, $\dot{r} = O(\epsilon)$, throughout the entire plunge, we have $\Delta L \simeq \Omega^{-1} \Delta E$, giving $\Delta W = O(\epsilon \eta \log \eta)$ and a contribution of $O(\epsilon \eta^2 \log \eta)$ to $\mathcal{W}^+_{\text{(end)}}$. Once again, this is negligible compared to the $O(\eta^2)$ terms in Eq. (64), assuming only $\epsilon \ll 1/|\ln \eta|$.

We conclude that, whether the plunge from the attractor occurs at the ISCO or earlier, the term $\mathcal{W}^+_{\text{(end)}}$ in Eq. (64) is always subdominant and negligible for $\epsilon, \eta \ll 1$. Within our approximation, the energy the particle carries with it as it crosses the horizon is the energy with which it has left the attractor.

---

[11]See Eq. (4.3) of [34], noting their κ is $\sqrt{2}\epsilon$ in our notation. The discussion following that equation seems to ignore the ϵ dependence implicit in the factors $\exp[-\kappa(T - t_0)/r_+] \sim (r_{\text{init}} - r_+)/r_+$ and $(E - \Omega_H L)$.

[12]It is not clear to us whether the low-frequency and quadrupole approximations, introduced in [34] to enable analytic calculation, are justifiable for near-horizon orbits in the near-extremal case (where the dimensionless angular velocity is $\sim 1/2$). We prefer to regard the form of Eq. (C5) as indicative only, but this should suffice for our purpose.

[13]This can be seen in one of two ways. First, by noting that the shift in the radial location of the unstable equilibrium due to the GSF is of $O(\eta \epsilon)$ [see the discussion below Eq. (48)], and, second, by using the regularity argument presented below Eq. (B5) to show that F^r must be $\propto \epsilon$ (in addition to being $\propto \eta^2$).

[1] R. Penrose, Rivista del Nuovo Cimento, Numero Speziale **1,** 252 (1969).

[2] M. W. Choptuik, Phys. Rev. Lett. **70,** 9 (1993).

[3] R. Wald, arXiv:gr-qc/9710068.

[4] R. Wald, Ann. Phys. (N.Y.) **82,** 548 (1974).

[5] V. E. Hubeny, Phys. Rev. D **59,** 064013 (1999).

[6] T. Jacobson and T. P. Sotiriou, Phys. Rev. Lett. **103,** 141101 (2009).

[7] A. Saa and R. Santarelli, Phys. Rev. D **84,** 027501 (2011).

[8] S. Isoyama, N. Sago, and T. Tanaka, Phys. Rev. D **84,** 124024 (2011).

[9] P. Zimmerman, I. Vega, E. Poisson, and R. Haas, Phys. Rev. D **87,** 041501 (2013).

[10] P. Zimmerman and E. Poisson, Phys. Rev. D **90,** 084030 (2014).

[11] T. M. Linz, J. L. Friedman, and A. G. Wiseman, Phys. Rev. D **90,** 084031 (2014).

[12] E. Barausse, V. Cardoso, and G. Khanna, Phys. Rev. Lett. **105,** 261102 (2010).

[13] E. Barausse, V. Cardoso, and G. Khanna, Phys. Rev. D **84,** 104006 (2011).

[14] T. Jacobson, Classical Quantum Gravity **28,** 187001 (2011).

[15] S. Hod, Phys. Rev. D **66,** 024016 (2002).

[16] S. Isoyama, L. Barack, S. R. Dolan, A. Le Tiec, H. Nakano, A. G. Shah, T. Tanaka, and N. Warburton, Phys. Rev. Lett. **113,** 161101 (2014).

[17] J. Levin and G. Perez-Giz, Phys. Rev. D **79,** 124013 (2009).

[18] K. Glampedakis and D. Kennefick, Phys. Rev. D **66,** 044002 (2002).

[19] J. M. Bardeen, W. H. Press, and S. A. Teukolsky, Astrophys. J. **178,** 347 (1972).

[20] S. E. Gralla and R. M. Wald, Classical Quantum Gravity **25,** 205009 (2008).

[21] A. Pound, Phys. Rev. D **81,** 024023 (2010).

[22] A. I. Harte, Classical Quantum Gravity **29,** 055012 (2012).

[23] E. Poisson, A. Pound, and I. Vega, Living Rev. Relativity **14,** 1 (2011).

[24] A. I. Harte, arXiv:1405.5077.

[25] L. Barack, Classical Quantum Gravity **26,** 213001 (2009).

[26] L. Barack and A. Ori, Phys. Rev. D **64,** 124003 (2001).

[27] T. C. Quinn and R. M. Wald, Phys. Rev. D **60,** 064009 (1999).

[28] S. Detweiler, Phys. Rev. D **77,** 124026 (2008).

[29] E. E. Flanagan, S. A. Hughes, and U. Ruangsri, Phys. Rev. D **89,** 084028 (2014).

[30] C. Gundlach, S. Akcay, L. Barack, and A. Nagar, Phys. Rev. D **86,** 084022 (2012).

[31] W. Israel, Phys. Rev. Lett. **57,** 397 (1986).

[32] A. Ori and K. S. Thorne, Phys. Rev. D **62,** 124022 (2000).

[33] M. Kesden, Phys. Rev. D **83,** 104011 (2011).

[34] Y. Mino and J. Brink, Phys. Rev. D **78,** 124015 (2008).

[35] T. Harada and M. Kimura, Phys. Rev. D **84,** 124032 (2011).

[36] M. Colleoni, A. Shah, M. van de Meent, and L. Barack (to be published).

[37] M. van de Meent (private communication).

[38] D. Kennefick, Phys. Rev. D **58,** 064012 (1998).

[39] S. Isoyama, R. Fujita, A. Le Tiec, H. Nakano, N. Sago, and T. Tanaka (to be published).

[40] A. Le Tiec, L. Blanchet, and B. F. Whiting, Phys. Rev. D **85,** 064039 (2012).

[41] L. Blanchet, A. Buonanno, and A. Le Tiec, Phys. Rev. D **87,** 024030 (2013).

[42] J. L. Friedman, K. Uryū, and M. Shibata, Phys. Rev. D **65,** 064035 (2002).

[43] A. Le Tiec, E. Barausse, and A. Buonanno, Phys. Rev. Lett. **108,** 131103 (2012).

[44] S. E. Gralla and A. Le Tiec, Phys. Rev. D **88,** 044021 (2013).

[45] S. R. Dolan and L. Barack, Phys. Rev. D **87,** 084066 (2013).

[46] T. S. Keidl, A. G. Shah, J. L. Friedman, D.-H. Kim, and L. R. Price, Phys. Rev. D **82,** 124012 (2010).

[47] A. G. Shah, J. L. Friedman, and T. S. Keidl, Phys. Rev. D **86,** 084059 (2012).

[48] A. Pound, C. Merlin, and L. Barack, Phys. Rev. D **89,** 024009 (2014).

[49] S. A. Hughes, Phys. Rev. D **61,** 084004 (2000).

[50] S. Mano, H. Suzuki, and E. Takasugi, Prog. Theor. Phys. **95,** 1079 (1996).

[51] S. Akcay, L. Barack, T. Damour, and N. Sago, Phys. Rev. D **86,** 104041 (2012).

[52] S. Dolan and L. Barack (to be published).

[53] S. J. Kapadia, D. Kennefick, and K. Glampedakis, Phys. Rev. D **87,** 044050 (2013).

[54] A. Buonanno and T. Damour, Phys. Rev. D **62,** 064015 (2000).

[55] S. A. Hughes, Phys. Rev. D **61,** 084004 (2000).

[56] P. L. Chrzanowski, Phys. Rev. D **13,** 806 (1976).