

# Two-timescale evolution of extreme-mass-ratio inspirals: waveform generation scheme for quasicircular orbits in Schwarzschild spacetime

Jeremy Miller<sup>1,2</sup> and Adam Pound<sup>3</sup>

<sup>1</sup>*Department of Physics, National Institute of Oceanography,  
Israel Oceanographic and Limnological Research, Haifa, Israel*

<sup>2</sup>*Jerusalem College of Technology, 21 Havaad Haleumi Street, Jerusalem 9372115, Israel*

<sup>3</sup>*School of Mathematical Sciences and STAG Research Centre,  
University of Southampton, Southampton, SO17 1BJ, United Kingdom*

(Dated: January 24, 2021)

Extreme-mass-ratio inspirals, in which a stellar-mass compact object spirals into a supermassive black hole in a galactic core, are expected to be key sources for LISA. Modelling these systems with sufficient accuracy for LISA science requires going to second (or *post-adiabatic*) order in gravitational self-force theory. Here we present a practical two-timescale framework for achieving this and generating post-adiabatic waveforms. The framework comprises a set of frequency-domain field equations that apply on the fast, orbital timescale, together with a set of ordinary differential equations that determine the evolution on the slow, inspiral timescale. Our analysis is restricted to the special case of quasicircular orbits around a Schwarzschild black hole, but its general structure carries over to the realistic case of generic (inclined and eccentric) orbits in Kerr spacetime. In our restricted context, we also develop a tool that will be useful in all cases: a formulation of the frequency-domain field equations using hyperboloidal slicing, which significantly improves the behavior of the sources near the boundaries. We give special attention to the slow evolution of the central black hole, examining its impact on both the two-timescale evolution and the earlier self-consistent evolution scheme.

## I. INTRODUCTION AND SUMMARY

Four years after the first direct detection of gravitational waves [1], the LIGO-Virgo collaboration now announces new detections on a regular basis [2–4]. To date, all the signals have originated from compact binary inspirals, involving either black holes or neutron stars spiralling toward each other and eventually merging. Observations of these systems have provided a wealth of information about the population of black holes in the universe [5], the equation of state of neutron stars [6], and the validity of general relativity in the strong-field regime [7, 8].

However, these binaries have all occupied a restricted region of the parameter space, in which the two objects are of roughly equal size and their orbits are approximately quasicircular. When the space-based detector LISA is launched, one of its key sources will be a very different class of binaries called extreme-mass-ratio inspirals (EMRIs), comprising a stellar-mass compact object of mass  $\mu$  slowly spiraling into a black hole of mass  $M \sim 10^5\text{--}10^7 M_\odot$  [9]. Because the inspiral is very slow, an EMRI can lie in the LISA band for the full duration of the mission. The small object can execute hundreds of thousands of intricate orbits in that time, generating a high-resolution map of the massive black hole’s spacetime. This map, as represented by precise measurements of the black hole’s multipole moments, for example, is encoded in the emitted gravitational radiation, along with other detailed information about the strong-field dynamics [10].

To extract this information from a detected EMRI waveform using matched filtering, we require a model that maintains phase coherence over potentially  $\sim 10^5$

wave cycles. *Gravitational self-force theory* currently provides the only viable route to meeting this stringent accuracy goal [11]. Broadly speaking, the gravitational self-force describes a gravitating object’s deviation from test-body motion due to the object’s own gravitational field. In the context of an EMRI, the small object of mass  $\mu$  slightly perturbs the spacetime of the large black hole of mass  $M$ , such that the total metric takes the form  $\mathbf{g}_{\alpha\beta} = g_{\alpha\beta} + \varepsilon h_{\alpha\beta}^1 + \varepsilon^2 h_{\alpha\beta}^2 + \mathcal{O}(\varepsilon^3)$ , where  $g_{\alpha\beta}$  is the Kerr metric of the large black hole and  $\varepsilon$  is a formal counting parameter that counts powers of the small mass ratio  $\mu/M$ . The perturbations  $h_{\alpha\beta}^n$  accelerate the small object away from geodesic motion in  $g_{\alpha\beta}$ , driving the object’s slow inspiral, and we interpret this to be the effect of a self-force. If we represent the object’s trajectory with a worldline  $z^\mu$ , its covariant acceleration in  $g_{\alpha\beta}$  becomes

$$\frac{D^2 z^\alpha}{d\tau^2} = \varepsilon f_1^\alpha + \varepsilon^2 f_2^\alpha + \mathcal{O}(\varepsilon^3), \quad (1)$$

where the proper time  $\tau$  and covariant derivative  $\frac{D}{d\tau} = \frac{dz^\alpha}{d\tau} \nabla_\alpha$  are defined with respect to  $g_{\alpha\beta}$ . The forces (per unit mass)  $f_n^\alpha$  include both the gravitational self-force, due to  $h_{\alpha\beta}^n$ , and finite-size effects, due to the object’s spin (in  $f_1^\alpha$ ), quadrupole moments (in  $f_2^\alpha$ ), and higher moments (in  $f_{n>2}^\alpha$ ).

### A. EMRI models: requirements and status

It has been stressed for some time [12] that to accurately model EMRIs, one must include the second-order terms in the equation of motion (1). This follows from a simple scaling argument. Take  $\mathcal{E}$  to be the energy

of the small object and  $\dot{\mathcal{E}}$  to be the gravitational-wave flux of energy out of the system. Given that  $\mathcal{E} \sim \mu$  and  $\dot{\mathcal{E}} \sim (\varepsilon h_{\alpha\beta}^1)^2$ , the inspiral will take place over the *radiation-reaction time*  $t_{rr} = \mathcal{E}/\dot{\mathcal{E}} \sim M/\varepsilon$ . On this time scale, the second-order force  $f_2^\alpha$  causes a cumulative shift in  $z^\alpha$  of order

$$\delta z^\alpha \sim \varepsilon^2 f_2^\alpha t_{rr}^2 \sim \varepsilon^0. \quad (2)$$

Since matched filtering will require errors in orbital phase to be much less than 1 radian, this suggests that the contribution of  $f_2^\alpha$  cannot be neglected. However, the same reasoning shows that the effect of  $f_3^\alpha$  can be safely neglected. Therefore, it is both necessary and sufficient to include second-order effects.

This scaling argument was made more precise by Hinderer and Flanagan [13], who showed that on the radiation-reaction time scale, the phase of the gravitational waveform has an expansion of the form

$$\varphi = \frac{1}{\varepsilon} [\varphi_0(\varepsilon t) + \varepsilon \varphi_1(\varepsilon t) + \mathcal{O}(\varepsilon^2)]. \quad (3)$$

The leading term in this expansion,  $\frac{1}{\varepsilon}\varphi_0$ , is said to be of *adiabatic* order. Computing it requires only a certain time average of the dissipative piece of  $f_1^\alpha$ ,  $f_{1,\text{diss}}^\alpha$ . The first subleading term,  $\varphi_1$ , is said to be of *first post-adiabatic* order. Computing it requires the complete  $f_1^\alpha$  (including the conservative piece,  $f_{1,\text{cons}}^\alpha$ ) and a certain time average of the dissipative piece of  $f_2^\alpha$ ,  $f_{2,\text{diss}}^\alpha$ .<sup>1</sup> The *second* post-adiabatic correction can be neglected, as it only contributes  $\sim \varepsilon$  to the accumulated phase.

In recent decades, there has been a significant effort to develop and implement adiabatic and post-adiabatic EMRI models, reviewed in Ref. [15]. Practical methods of calculating only the necessary input for adiabatic evolution have been formulated and implemented [16–19]. Due to the high-dimensional parameter space, significant work remains to actually generate adiabatic waveforms using this input [20], but it is expected that a template bank based on such waveforms will suffice to detect many (or even most) EMRI signals, though not to perform high-precision parameter estimation. In the absence of such adiabatic templates, existing “kludge models” may even suffice for detection [21].

For post-adiabatic modeling, the full first-order self-force can now also be calculated along generic bound orbits in Kerr spacetime [22], and there is ongoing work to incorporate first-order effects of the small object’s spin [23–25]. However, calculations of  $f_2^\alpha$  are far less mature. The basic formalism of self-force theory at second order, including the fundamental analytical ingredients, was derived by one of us in Ref. [26] (see also [27–31]).

Since then, there has been steady progress in developing this formalism into a practical numerical scheme [32–38]. In Ref. [39], we reported the first implementation of that scheme, resulting in a calculation of the gravitational binding energy of quasicircular EMRIs around Schwarzschild black holes. But there remain many challenges in generating post-adiabatic waveforms, particularly in the astrophysically relevant case of generic orbits around Kerr black holes.

## B. Orbital evolution and the two-timescale approximation

One way that second-order calculations are more complex than first-order ones is that they must incorporate the system’s evolution *ab initio*. At first order there is a sense in which one can delay the choice of evolution scheme. On short enough time scales, the orbit can be approximated as a geodesic. The metric perturbation  $h_{\alpha\beta}^1$ , and the self-force  $f_1^\alpha$ , can then be calculated as if generated by a point mass moving on that geodesic, with the freedom to later choose how to utilize those results to drive the evolution. At second order, this is no longer true: the evolution of the system acts as a source for the second-order field,  $h_{\alpha\beta}^2$ , meaning one must choose an evolution scheme before one can even write down the second-order field equation.

The simplest approach to evolution, used in derivations by Gralla and Wald [29, 40], is to consider perturbative corrections to the trajectory, as in  $z^\alpha(\tau, \varepsilon) = z_0^\alpha(\tau) + \varepsilon z_1^\alpha(\tau) + \mathcal{O}(\varepsilon^2)$ . In this approach, the mass  $\mu$  moving on the geodesic  $z_0^\alpha$  creates the metric perturbation  $h_{\alpha\beta}^1$ , which drives the correction to the trajectory,  $z_1^\alpha$ .  $z_1^\alpha$  then contributes to the next-order perturbation  $h_{\alpha\beta}^2$ , which contributes to the next-order correction to the trajectory,  $z_2^\alpha$ , and so on. This is conceptually simple, but it is obvious from the outset that it can only be accurate on short time scales: Over the course of an inspiral,  $\mu$  will move far from  $z_0^\alpha$ , such that the “small correction”  $\varepsilon z_1^\alpha$  will grow large with time. This growth in  $z_1^\alpha$  will create a commensurate growth in  $h_{\alpha\beta}^2$ . When the corrections become comparable to the leading terms, the expansion will no longer be valid.

Another approach is the *self-consistent* approximation, which treats  $z^\alpha$  nonperturbatively and makes each  $h_{\alpha\beta}^n$  a functional of that nonperturbative trajectory [31, 41]. The trajectory and metric perturbation are then to be determined together as a coupled system. This approximation accurately accounts for the long-term evolution of the trajectory. It has also been used in many of the foundational derivations in self-force theory, and it will be our starting point in this paper. However, it has only been concretely implemented in a scalar toy model [42], and because it uses a trajectory that is always evolving, it abandons the usual advantages of having approximately geodesic motion at first order.

The alternative that has generally been used in prac-

<sup>1</sup> This description is somewhat altered by the existence of transient resonances, but the main conclusions are unchanged [14].

tice is a method of *osculating geodesics* [23, 43–48]. In this approach, at each instant  $\tau$  along the accelerated worldline, one computes the self-force as if, for its entire past history, the mass  $\mu$  had been moving on the geodesic that is tangential to the worldline at that instant  $\tau$ . The orbit then effectively evolves smoothly from one geodesic to the next. In principle this method can be carried to second order (and beyond), as sketched in Ref. [49]. However, it does not make maximal use of the properties of an EMRI—specifically, the system’s near-periodicity.

The second-order calculation reported in Ref. [39] was instead based on another alternative: a *two-timescale expansion* [50] (or multiscale expansion) of the Einstein field equations. This expansion, previously utilized within self-force analyses in Refs. [13, 35, 51–55] and within post-Newtonian analyses in Refs. [56–59], is tailored to the properties of an EMRI. An EMRI has two disparate time scales: the orbital period  $\sim M$  and the much longer radiation-reaction time  $\sim M/\varepsilon$ . On the orbital time scale, the system is triperiodic, with frequencies  $\Omega_r$ ,  $\Omega_\theta$ , and  $\Omega_\phi$  associated with the radial, polar, and azimuthal motions. This leads to a leading-order metric perturbation with a discrete frequency spectrum,

$$h_{\alpha\beta}^1 = \sum_{mpq} h_{\alpha\beta}^{1,\omega_{mpq}}(r, \theta, \phi) e^{-i(p\Omega_r + q\Omega_\theta + m\Omega_\phi)t}, \quad (4)$$

where the frequency label refers to  $\omega_{mpq} = p\Omega_r + q\Omega_\theta + m\Omega_\phi$ , and  $(t, r, \theta, \phi)$  denote the Boyer-Lindquist coordinates of the background Kerr geometry. On the radiation-reaction time scale  $\sim M/\varepsilon$ , the system’s amplitudes and frequencies slowly evolve. To efficiently and accurately capture the behavior on both time scales, a two-timescale expansion introduces multiple time variables: *fast times*  $\varphi_r = \int \Omega_r dt$ ,  $\varphi_\theta = \int \Omega_\theta dt$ , and  $\varphi_\phi = \int \Omega_\phi dt$ , which vary on the orbital time scale; and a *slow time*  $\tilde{t} \sim \varepsilon t$ , which varies on the radiation-reaction time. The analog of Eq. (4) is then

$$h_{\alpha\beta}^n = \sum_{mpq} \tilde{h}_{\alpha\beta}^{n,\omega_{mpq}}(\tilde{t}, r, \theta, \phi) e^{-i(p\varphi_r + q\varphi_\theta + m\varphi_\phi)}. \quad (5)$$

On short time scales, the amplitudes  $\tilde{h}_{\alpha\beta}^{n,\omega_{mpq}}$  are approximately constant, while the phases are approximately  $\varphi_\alpha \sim \Omega_\alpha t$ , such that the two-timescale expansion reduces to the ordinary Fourier expansion (4) at leading order. However, the form (5) remains approximately triperiodic for *all* orders  $n$  over the entire inspiral, while an expansion of the form (4), which freezes the frequencies and amplitudes, can only be accurate at leading order and for short intervals of time.

We will see below that the ansatz (5) splits the Einstein equations into two distinct sets: frequency-domain equations that govern the amplitudes  $\tilde{h}_{\alpha\beta}^{n,\omega_{mpq}}$  at each fixed value of  $\tilde{t}$ , and evolution equations that determine the amplitudes and frequencies as functions of  $\tilde{t}$ . At first order, the equations for  $\tilde{h}_{\alpha\beta}^{1,\omega_{mpq}}$  are identical to those for the ordinary Fourier coefficients  $h_{\alpha\beta}^{1,\omega_{mpq}}$ , meaning a two-

timescale computation can naturally build on existing frequency-domain codes [22, 60, 61].

### C. Overview of this paper and outline of wave-generation framework

Although two-timescale expansions have been utilized in the past to explore features of EMRIs, they have largely centred on the expansion of the equation of motion (1). The work of Hinderer and Flanagan [13] has been particularly influential in this regard, having provided a complete treatment of that expansion, and having led to Eq. (3). However, because the mass  $\mu$ ’s trajectory  $z^\alpha$  is coupled to the metric perturbation, an expansion of the equation of motion is by itself incomplete; it must be combined with an expansion of the Einstein field equations.

Here we focus on that coupled problem, building on our previous work on a scalar toy model [35]. We restrict our analysis to the simplest case of quasicircular orbits in Schwarzschild spacetime. The orbital dynamics in this scenario is comparatively trivial, but the field equations have most of the essential features of the full problem. In this restricted context, we present the explicit form of the expanded Einstein equations and the framework they provide for generating post-adiabatic waveforms. A forthcoming series of papers [62–65] will extend our analysis to generic orbits.<sup>2</sup>

In the remainder of this section, we provide a complete outline of the paper and of the wave-generation framework. The body of the paper then fills in the technical details for interested readers.

We begin in Sec. II with a review of self-force theory in the self-consistent framework, keeping the discussion general enough to describe inspirals into a Kerr black hole. Through second order in  $\varepsilon$ , the coupled field equations and equation of motion take the form

$$E_{\alpha\beta}[\bar{h}^{1\mathcal{R}}] = -E_{\alpha\beta}[\bar{h}^{1\mathcal{P}}], \quad (6)$$

$$E_{\alpha\beta}[\bar{h}^{2\mathcal{R}}] = 2\delta^2 G_{\alpha\beta} - E_{\alpha\beta}[\bar{h}^{2\mathcal{P}}], \quad (7)$$

$$\frac{D^2 z^\alpha}{d\tau^2} = \varepsilon f_1^\alpha[h^{1\mathcal{R}}] + \varepsilon^2 f_2^\alpha[h^{1\mathcal{R}}, h^{2\mathcal{R}}], \quad (8)$$

given below in Eqs. (42) and (40)–(41) with (29), (33), (38), and (39). Here  $E_{\alpha\beta}$  is the linearized Einstein tensor in the Lorenz gauge (up to a factor of  $-1/2$ ), a bar denotes a trace-reversed field  $\bar{h}_{\alpha\beta}^n := h_{\alpha\beta}^n - \frac{1}{2}g_{\alpha\beta}g^{\mu\nu}h_{\mu\nu}^n$ , and  $\delta^2 G_{\alpha\beta}$  is the piece of the full Einstein tensor that is quadratic in  $h_{\alpha\beta}^1$ . Rather than solving for the physical fields  $h_{\alpha\beta}^n$ , we solve for the *residual* fields  $h_{\alpha\beta}^{n\mathcal{R}} := h_{\alpha\beta}^n - h_{\alpha\beta}^{n\mathcal{P}}$ , where  $h_{\alpha\beta}^{n\mathcal{P}}$  are analytically known *punctures*,

<sup>2</sup> Some months after this paper was submitted for publication, one of us provided an overview of the method for generic orbits [38]. Future papers will provide the complete details.

which diverge on  $z^\alpha$  but guarantee that the full fields  $h_{\alpha\beta}^n = h_{\alpha\beta}^{n\mathcal{P}} + h_{\alpha\beta}^{n\mathcal{R}}$  agree with the physical fields outside the small compact object. At first order, this is equivalent to approximating the small object as a point mass  $\mu$  on  $z^\alpha$ .

Typically,  $h_{\alpha\beta}^1$  is taken to *only* include the linear field of the mass  $\mu$ . However, because the large black hole absorbs gravitational radiation, its mass and spin slowly change with time. To accurately account for this, we modify previous descriptions of the self-consistent approximation such that the solution to Eq. (6) includes perturbations proportional to  $\delta M$  and  $\delta S$ , the small, evolving corrections to the black hole's mass and spin. (The background metric, on the other hand, remains stationary.)

In Sec. III, we review the two-timescale method, and in Sec. IV B we apply it to the equation of motion (8), specializing to quasicircular orbits in a Schwarzschild background in the process. For a quasicircular orbit, rather than three orbital frequencies and associated phases, there is only one frequency,  $\Omega(t, \varepsilon) := \frac{d\phi_p}{dt}$ , associated with the azimuthal angle  $\phi_p$  of the particle's orbit. (Here and below we adopt standard Schwarzschild coordinates.) Both  $\Omega$  and the orbital radius  $r_p$  are slowly evolving functions of  $\tilde{t} = \varepsilon t$ , given by the expansions

$$\Omega = \Omega_0(\tilde{t}) + \varepsilon \Omega_1(\tilde{t}) + \mathcal{O}(\varepsilon^2), \quad (9)$$

$$r_p = r_0(\tilde{t}) + \varepsilon r_1(\tilde{t}) + \mathcal{O}(\varepsilon^2). \quad (10)$$

The terms  $r_n$  satisfy ordinary differential equations with the schematic form

$$\frac{dr_0}{d\tilde{t}} = \dot{r}_0(r_0, f_{1,\text{diss}}^\alpha), \quad (11)$$

$$\frac{dr_1}{d\tilde{t}} = \dot{r}_1(r_0, r_1, f_{2,\text{diss}}^\alpha), \quad (12)$$

given explicitly by Eqs. (103) and (106), and the terms  $\Omega_n$  are determined by relations

$$\Omega_0 = \Omega_0(r_0), \quad (13)$$

$$\Omega_1 = \Omega_1(r_0, r_1, f_{1,\text{cons}}^\alpha), \quad (14)$$

given in Eqs. (102) and (104). Here and throughout this paper, an overdot denotes a derivative with respect to slow time.

In Sec. V, we turn to the two-timescale expansion of the field equations (6) and (7). In place of the general expansion (5), we assume an expansion of the form<sup>3</sup>

$$\bar{h}_{\alpha\beta}^n = \sum_{i\ell m} \frac{a_{i\ell}}{r} R_{i\ell m}^n(\tilde{s}, r) e^{-im\phi_p} Y_{\alpha\beta}^{i\ell m}, \quad (15)$$

<sup>3</sup> In actuality, order- $\varepsilon$  corrections appear on the right-hand side of this expansion. Here we have implicitly absorbed those corrections into the higher- $n$  mode amplitudes.

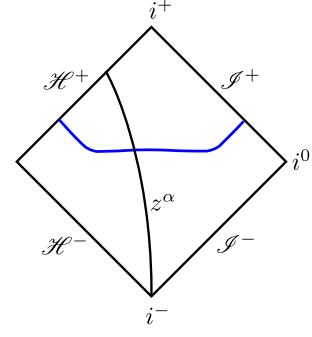


FIG. 1. Penrose diagram of Schwarzschild spacetime illustrating a slice (blue curve) of constant hyperboloidal time  $s = t - k(r^*)$ .  $s$  transitions from advanced time  $v = t + r^*$  near the future horizon, to Schwarzschild time  $t$  in a region including the particle's worldline  $z^\alpha$ , to retarded time  $u = t - r^*$  near future null infinity.

and the analogs for  $\bar{h}_{\alpha\beta}^{n\mathcal{R}}$  and  $\bar{h}_{\alpha\beta}^{n\mathcal{P}}$ . Here in addition to the multiscale expansion, we have performed an expansion in tensor spherical harmonics;  $i = 1, \dots, 10$  label the 10 linearly independent harmonics  $Y_{\alpha\beta}^{i\ell m}$ , given explicitly in Appendix B, and  $a_{i\ell}$  is a convenient numerical factor. More importantly, we have also exploited the freedom that arises in extending the slow and fast time variables away from the worldline. Rather than simply using  $\tilde{t} = \varepsilon t$ , we have introduced a slow time  $\tilde{s} := \varepsilon s(t, r)$ . This allows us to naturally account for retardation: the slow evolution of the orbit does not propagate out from  $z^\alpha$  instantaneously along slices of constant  $t$ , but instead along null curves. In our analysis we hence consider a hyperboloidal slow time  $s$  that is equal to  $t$  in a neighbourhood of the particle but becomes null as  $r \rightarrow 2M$  or  $r \rightarrow \infty$ , as illustrated in Fig. 1. Analogously, as our fast time, rather than  $\phi_p(t, \varepsilon) = \int^t \Omega(z, \varepsilon) dz$ , we use  $\phi_p(s, \varepsilon) = \int^s \Omega(z, \varepsilon) dz$ .

When we substitute the expansion (133) into Eqs. (6) and (7), derivatives with respect to  $t$  and  $r$  act on the slow- and fast-time dependence, as shown in Eqs. (134) and (135). In this counting, a derivative with respect to slow time is suppressed by a factor of  $\varepsilon$ . Factoring out the spherical harmonics and fast-time phase factors, we obtain a sequence of frequency-domain equations for the mode amplitudes  $R_{i\ell m}^n$ :

$$E_{ij\ell m}^0 R_{j\ell m}^{1\mathcal{R}} = -E_{ij\ell m}^0 R_{j\ell m}^{1\mathcal{P}}, \quad (16)$$

$$E_{ij\ell m}^0 R_{j\ell m}^{2\mathcal{R}} = 2\delta^2 G_{i\ell m}^0 - E_{ij\ell m}^0 R_{j\ell m}^{2\mathcal{P}} - E_{ij\ell m}^1 R_{j\ell m}^1, \quad (17)$$

where  $E_{ij\ell m}^0$  and  $E_{ij\ell m}^1$  are given in Eq. (144), and the repeated basis label  $j$  is summed over.  $E_{ij\ell m}^0$  is a radial operator in which  $\partial_t$  has been replaced with  $-i\omega_m$ , where  $\omega_m := m\Omega_0$ .  $E_{ij\ell m}^1$  contains the first subleading effect of the time derivatives in  $E_{\alpha\beta}$ ; it is linear in  $\partial_{\tilde{s}} - im\Omega_1$ . The left-hand side of Eqs. (16) and (17) have precisely the same form as if the metric perturbations had been expanded in hyperboloidal-time Fourier modes  $e^{-i\omega_m s}$ . Such use of hyperboloidal slicing in the frequency do-

main has been considered before in, e.g., Refs. [66, 67], though not directly for the linearized Einstein equation. In the context of our two-timescale expansion, this slicing dramatically improves the behavior of the source term  $E_{i\ell m}^1 R_{j\ell m}^1$  as  $r \rightarrow 2M$  or  $r \rightarrow \infty$ .

In Sec. VI we derive our waveform-generation scheme. The solutions to Eqs. (16) and (17) (for the full fields  $R_{i\ell m}^{\mathcal{R}} + R_{i\ell m}^{\mathcal{P}}$ ) have the schematic form

$$R_{i\ell m}^1 = \hat{R}_{i\ell m}^1(r_0, r) + \bar{x}_{i\ell m}(M_1, S_1, r), \quad (18)$$

$$R_{i\ell m}^2 = \hat{R}_{i\ell m}^2(r_0, r_1, M_1, S_1, r) + \bar{x}_{i\ell m}(M_2, S_2, r), \quad (19)$$

where  $\hat{R}_{i\ell m}^1 = R_{i\ell m}^{\text{PP}}$  is the usual linear perturbation due to a point particle on a circular orbit, and  $M_n$  and  $S_n$  are the coefficients in

$$\delta M = \varepsilon M_1(\tilde{s}) + \varepsilon^2 M_2(\tilde{s}) + \mathcal{O}(\varepsilon^3), \quad (20)$$

$$\delta S = \varepsilon S_1(\tilde{s}) + \varepsilon^2 S_2(\tilde{s}) + \mathcal{O}(\varepsilon^3). \quad (21)$$

All the slow-time dependence is encoded in the dependence on  $r_n(\tilde{s})$ ,  $M_n(\tilde{s})$ , and  $S_n(\tilde{s})$ .  $\bar{x}_{i\ell m}(M_n, S_n, r)$  is the same function for all  $n$ , just with different arguments. It represents a linear perturbation toward a slowly evolving Kerr metric with mass and spin parameters  $M + \varepsilon^n M_n(\tilde{s})$  and  $\varepsilon^n S_n(\tilde{s})$ ; this has no fast-time dependence, containing only  $\ell = 0$  and  $\ell = 1, m = 0$  contributions.

Once one has obtained the solutions (18) and (19), one can calculate the self-forces appearing in Eqs. (11)–(14). These forces depend only on the mode amplitudes  $R_{i\ell m}^n$ ; because  $Y_{\alpha\beta}^{i\ell m} e^{-im\phi_p} \propto e^{im(\phi - \phi_p)}$ , the phases drop out of the metric perturbation on the worldline, where  $\phi = \phi_p$ , and therefore out of the force. The forces then have the schematic forms  $f_{1,\text{diss}}^\alpha = f_{1,\text{diss}}^\alpha(r_0)$ ,  $f_{1,\text{cons}}^\alpha = f_{1,\text{cons}}^\alpha(r_0, M_1, S_1)$ , and  $f_{2,\text{diss}}^\alpha = f_{2,\text{diss}}^\alpha(r_0, r_1, M_1, S_1)$ . Substituting these dependences into Eqs. (11)–(14), and noting that  $\tilde{s} = \tilde{t}$  on the orbit, we see that the orbital evolution equations take the form

$$\Omega_0 = \Omega_0(r_0), \quad \frac{dr_0}{d\tilde{s}} = \dot{r}_0(r_0), \quad (22)$$

$$\Omega_1 = \Omega_1(r_0, r_1, M_1, S_1), \quad \frac{dr_1}{d\tilde{s}} = \dot{r}_1(r_0, r_1, M_1, S_1). \quad (23)$$

We also have that

$$\frac{dM_1}{d\tilde{s}} = \dot{E}_H(r_0), \quad \frac{dS_1}{d\tilde{s}} = \dot{L}_H(r_0), \quad (24)$$

where  $\dot{E}_H$  and  $\dot{L}_H$  are the leading-order gravitational-wave fluxes of energy and angular momentum into the black hole; these depend only on the first-order mode amplitudes  $R_{i\ell m}^{\text{PP}}$  and therefore only on  $r_0$ . In Sec. VII, we re-derive Eq. (24) from first principles, directly from our two-timescale field equations; at the same time, we re-derive the standard balance laws relating the loss of orbital energy and angular momentum,  $\dot{\mathcal{E}}_0(r_0)$  and  $\dot{\mathcal{L}}_0(r_0)$ , to the total fluxes out of the system,  $\dot{E}_H(r_0) + \dot{E}_\infty(r_0)$  and  $\dot{L}_H(r_0) + \dot{L}_\infty(r_0)$ .

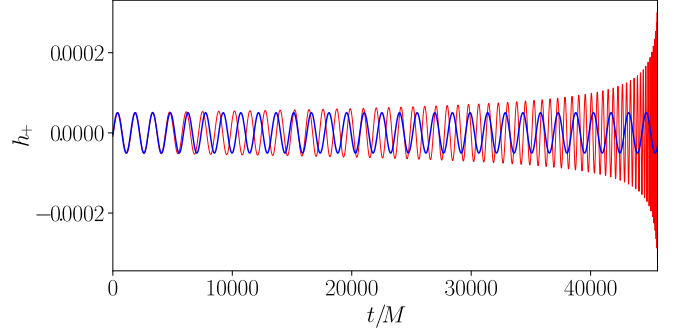


FIG. 2. The  $\ell = 2$ ,  $m = \pm 2$  mode of the ‘+’ polarization of the waveform at infinity in units of  $\mu$ . The thin red curve shows an adiabatic waveform,  $h_+ = \lim_{r \rightarrow \infty} \sum a_{i\ell} R_{i\ell m}^1(\varepsilon u, r) r^{-2} Y_{\theta\theta}^{i\ell m} e^{-im \int_0^u \Omega_0(\varepsilon z) dz}$ , where the sum is over  $i = 7, 10$ ,  $\ell = 2$ ,  $m = \pm 2$ . We have used an exaggerated mass ratio  $\varepsilon = 0.1$  to make the evolution clearly visible. For comparison, the thick blue curve shows the waveform produced by a point mass on a circular geodesic with the same initial frequency as the adiabatic orbit,  $h_+ = \lim_{r \rightarrow \infty} \sum a_{i\ell} R_{i\ell m}^1(0, r) r^{-2} Y_{\theta\theta}^{i\ell m} e^{-im \Omega_0(0)u}$ .

Equation (22) represents an adiabatic evolution scheme. One can use  $\frac{dr_0}{d\tilde{s}} = \dot{r}_0(r_0)$  to evolve the leading-order orbital radius  $r_0(\tilde{s})$ , recover the leading-order frequency  $\Omega_0(r_0)$ , and from it recover the adiabatic-order phase  $\phi_p(s, \varepsilon) = \int^s \Omega_0(\varepsilon z) dz$ . Since one has already computed the amplitudes  $R_{i\ell m}^{\text{PP}}(r_0, r)$  as input for  $\dot{r}_0(r_0)$ , one then has the waveform  $\sum R_{i\ell m}^{\text{PP}} e^{-im \int \Omega_0 ds} Y_{\alpha\beta}^{i\ell m}$  as an output (since it only evolves slowly,  $\bar{x}_{i\ell m}$  is not required in the leading-order waveform). Fig. 2 shows an adiabatic waveform generated with this scheme, using the computational methods described in Ref. [68].

Equations (22)–(24) together represent a post-adiabatic evolution scheme. In this case one has four parameters to evolve:  $r_0$ ,  $r_1$ ,  $M_1$ , and  $S_1$ . From their evolution, one extracts the frequency evolution  $\Omega_0(r_0) + \varepsilon \Omega_1(r_0, r_1, M_1, S_1)$ ; from the frequency, the post-adiabatic phase evolution  $\phi_p(s, \varepsilon) = \int^s [\Omega_0(\varepsilon z) + \varepsilon \Omega_1(\varepsilon z)] dz$ ; and from the phase, the waveform. Explicitly, the two polarizations of the leading-order waveform are given by [69]

$$h_+ = \lim_{r \rightarrow \infty} \sum \frac{a_{i\ell}}{r^2} R_{i\ell m}^{\text{PP}}(\tilde{u}, r) e^{-im\phi_p(u, \varepsilon)} Y_{\theta\theta}^{i\ell m}, \quad (25)$$

$$h_\times = \lim_{r \rightarrow \infty} \sum \frac{a_{i\ell}}{r^2 \sin^2 \theta} R_{i\ell m}^{\text{PP}}(\tilde{u}, r) e^{-im\phi_p(u, \varepsilon)} Y_{\theta\phi}^{i\ell m}, \quad (26)$$

where the sums run over  $i = 7, 10$ ,  $\ell \geq 2$ ,  $m \neq 0$ . Because the waveform’s amplitude is not required to be highly accurate for matched filtering, it is unlikely that we would need to include the second-order mode amplitudes in this waveform.

Figure 3 shows the full sequence of steps required to generate adiabatic and post-adiabatic waveforms. The heart of this evolution scheme lies in solving the field equations (16) and (17) at fixed values of slow time—

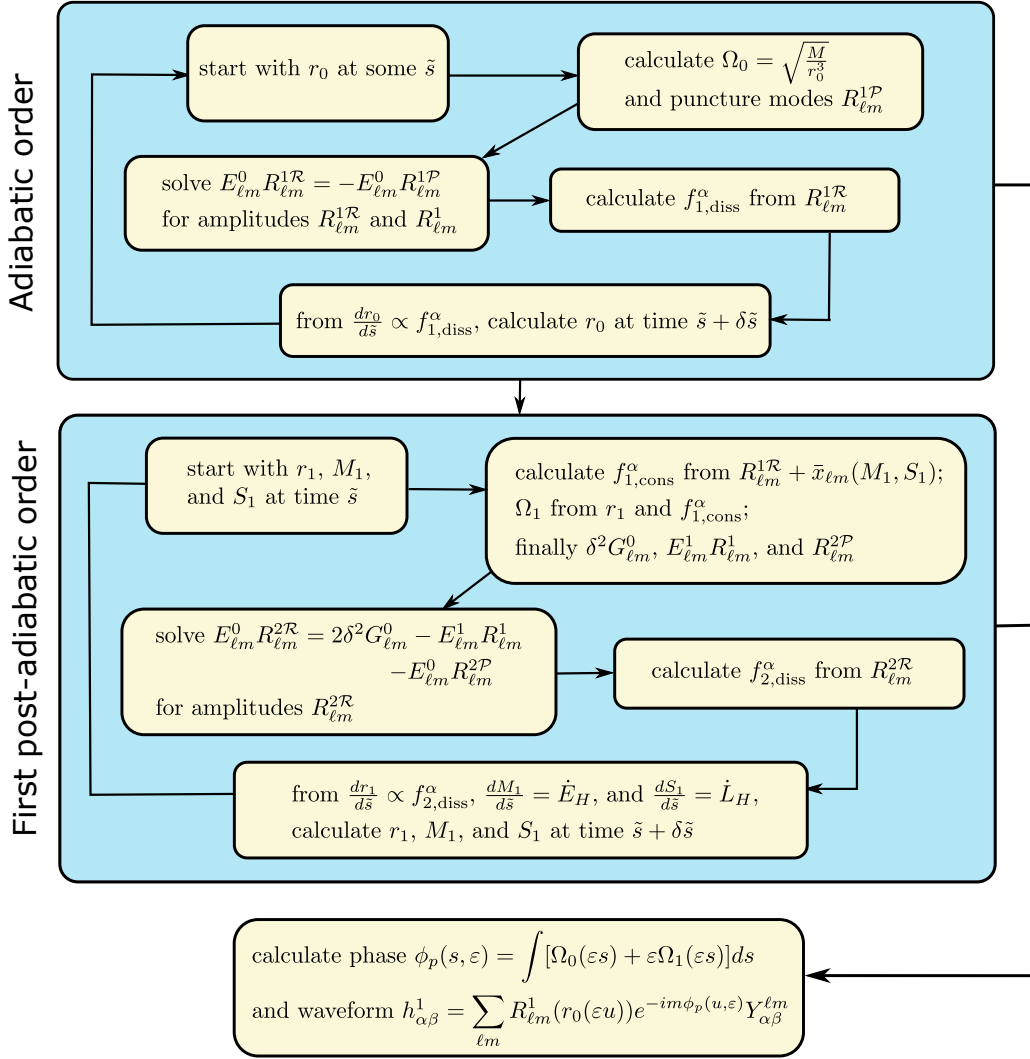


FIG. 3. Our procedure for computing the gravitational waveform to adiabatic and first post-adiabatic orders. The computation moves along a sequence of orbital radii, described as functions of slow time  $\tilde{s} = \varepsilon s$  by  $r_p = r_0(\tilde{s}) + \varepsilon r_1(\tilde{s}) + \mathcal{O}(\varepsilon^2)$ . At each value of slow time, one solves (discrete) frequency-domain field equations for the metric-perturbation mode amplitudes. From the mode amplitudes, one computes the self-force, which determines the evolution to the next value of slow time as well as the post-adiabatic correction  $\Omega_1$  to the orbital frequency. Once the mode amplitudes and frequency evolution are known, one can construct the full time-domain waveform. For simplicity we suppress details such as the tensor-harmonic basis labels and the numerical factor  $a_{i\ell}$ .

i.e., for given points in the  $(r_0, r_1, M_1, S_1)$  parameter space. The evolution then progresses through that parameter space. One could alter this evolution scheme in various ways. For example, one can instead work in the  $(\Omega_0, \Omega_1, M_1, S_1)$  parameter space by simply rearranging Eqs. (13)–(14) to obtain  $r_0 = r_0(\Omega_1)$  and  $r_1 = r_1(\Omega_0, \Omega_1, f_1^r)$ . More significantly, one could adjust the evolution scheme to generate the waveform as a function of the full, nonperturbative frequency  $\Omega$ , reducing  $\tilde{s}$  to an auxiliary variable. This approach, which we describe in Appendix A, hews slightly closer to our starting point in the self-consistent expansion, as it treats more of the particle’s trajectory nonperturbatively. It also provides a convenient way to associate a local state

of the binary to an asymptotic waveform: the two are naturally identified when their frequencies are the same.

But the core of the scheme, solving first- and second-order frequency-domain field equations, is largely independent of which of these approaches is taken. In all cases, the calculations at each point in the parameter space require (i) a practical method of solving the field equations, subject to given boundary conditions, (ii) a computation of the source terms in the field equations, and (iii) a specification of physical boundary conditions. We will present these remaining requirements in a sequence of followup papers [64, 65, 70–72].

In this paper we use a mostly positive metric signature,  $(-, +, +, +)$ , and geometrical units with  $G = c = 1$ .

Indices are raised and lowered with the background metric  $g_{\alpha\beta}$ , and  $\nabla$  and a semicolon both denote the covariant derivative compatible with  $g_{\alpha\beta}$ .  $(t, r, \theta, \phi)$  denote Schwarzschild coordinates, and  $\theta^A = (\theta, \phi)$ . Unless otherwise stated,  $g_{\alpha\beta}$  denotes the Schwarzschild metric,  $\text{diag}(-f, f^{-1}, r^2, r^2 \sin^2 \theta)$ , where  $f := 1 - 2M/r$ .

## II. SELF-FORCE THEORY IN THE SELF-CONSISTENT FRAMEWORK

In this section we review our starting point: self-force theory through second order in  $\varepsilon$ . We follow the self-consistent formalism of Refs. [41, 49], which provides the most direct line to a two-timescale approximation. However, as alluded to in the introduction, we extend previous descriptions to accurately incorporate the evolution of the central black hole. Sec. IIA reviews the standard description in the literature, and Sec. IIB discusses what prevents that description from admitting a two-timescale expansion. Sec. IIC describes how to overcome the problem using an evolving background space-time, and Sec. IID simplifies that formulation by working with small, slowly evolving perturbations rather than an evolving background. The end result, Eqs. (72)–(73), is a self-consistent formalism that can be straightforwardly expanded in two-timescale form.

### A. Review of previous descriptions

The self-consistent formalism begins by expanding  $\mathbf{g}_{\alpha\beta}$  in the limit  $\varepsilon \rightarrow 0$  while holding  $z^\mu(t, \varepsilon)$  fixed:

$$\mathbf{g}_{\alpha\beta} = g_{\alpha\beta}(x^\mu) + \sum_{n \geq 1} \varepsilon^n h_{\alpha\beta}^n(x^\mu; z^\mu). \quad (27)$$

Here  $g_{\alpha\beta}(x^\mu)$  is the background metric (e.g., the metric of the central Kerr black hole in an EMRI),  $x^\mu$  are a set of background coordinates (e.g., Boyer-Lindquist coordinates on the Kerr background), and  $h_{\alpha\beta}^n(x^\mu; z^\mu)$  is the  $n$ th-order perturbation due to the small object.<sup>4</sup> The expansion (27) is taken to be accurate except in a small region around the small object, called the *body zone* or inner region, where the object itself is the dominant source of gravity. In the body zone, a second approximation is used, called a scaled or *inner expansion*. The two are linked using the method of matched asymptotic expansions. We find  $h_{\alpha\beta}^n$  by solving the vacuum field equations *outside* the small object, subject to the *matching condition* that near the body zone, the metric (27) suitably agrees with, or “matches”, the inner expansion.

Because  $z^\mu$  depends on  $\varepsilon$ , the perturbations  $h_{\alpha\beta}^n(x^\mu; z^\mu)$  do as well. This means we need a special

formulation of the field equations. To see this, substitute the metric (27) into the vacuum Einstein equations  $G_{\alpha\beta}[\mathbf{g}] = 0$ , and move nonlinear terms to the right-hand side, to obtain

$$G_{\alpha\beta}[g] + \varepsilon \delta G_{\alpha\beta}[h^1] + \varepsilon^2 \delta G_{\alpha\beta}[h^2] = -\varepsilon^2 \delta^2 G_{\alpha\beta}[h^1] + \mathcal{O}(\varepsilon^3), \quad (28)$$

where  $\delta G_{\alpha\beta}$  is the linearized Einstein tensor, and  $\delta^2 G_{\alpha\beta}$  is the piece of  $G_{\alpha\beta}[g + h]$  that is quadratic in  $h_{\alpha\beta}$ . Explicitly, in vacuum,

$$\delta^2 G_{\alpha\beta} = \delta^2 R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} g^{\mu\nu} \delta^2 R_{\mu\nu}, \quad (29)$$

where

$$\begin{aligned} \delta^2 R_{\alpha\beta}[h] = & \frac{1}{2} \bar{h}^{\mu\nu}{}_{;\mu} (2h_{\nu(\alpha;\beta)} - h_{\alpha\beta;\mu}) + \frac{1}{4} h^{\mu\nu}{}_{;\alpha} h_{\mu\nu;\beta} \\ & + \frac{1}{2} h^\mu{}_\beta{}^{;\nu} h_{\mu\alpha;\nu} - \frac{1}{2} h^\mu{}_\beta{}^{;\nu} h_{\nu\alpha;\mu} \\ & - \frac{1}{2} h^{\mu\nu} (2h_{\mu(\alpha;\beta)\nu} - h_{\alpha\beta;\mu\nu} - h_{\mu\nu;\alpha\beta}). \end{aligned} \quad (30)$$

Because of the  $\varepsilon$  dependence in  $h_{\alpha\beta}^n(x^\mu; z^\mu)$ , we cannot naively equate coefficients of powers of  $\varepsilon$  in Eq. (28). By virtue of the Bianchi identity, the Einstein equations constrain any material degrees of freedom in the system, and equating coefficients of powers of  $\varepsilon$  in Eq. (28) forces the  $z^\mu$  in  $h_{\alpha\beta}^1(x^\mu; z^\mu)$  to be an  $\varepsilon$ -independent geodesic  $z_0^\mu$ . The  $\varepsilon$  dependence of the object’s motion then appears in the higher-order perturbations  $h_{\alpha\beta}^{n>0}$  in the form of deviation vectors  $z_n^\mu$ , which describe the small object’s deviation away from  $z_0^\mu$ . In effect, the perturbative Einstein equations force the trajectory to be expanded in the form  $z^\mu(\tau, \varepsilon) = z_0^\mu(\tau) + \varepsilon z_1^\mu(\tau) + \dots$ , with  $h_{\mu\nu}^1$  only depending on  $z_0^\mu$ ,  $h_{\mu\nu}^2$  depending on  $z_0^\mu$  and  $z_1^\mu$ , etc. This is the Gralla-Wald treatment of the motion, which (as described in the introduction) becomes inaccurate over an inspiral because the corrections  $z_{n>0}^\mu$  grow large with time.

We avoid this problem by casting the field equations in a *relaxed* form that does not constrain  $z^\mu$ . To achieve that, we impose the Lorenz gauge condition

$$\nabla^\beta (\varepsilon \bar{h}_{\alpha\beta}^1 + \varepsilon^2 \bar{h}_{\alpha\beta}^2) = \mathcal{O}(\varepsilon^3). \quad (31)$$

(Refs. [30] and [73] discuss more general choices of gauge to achieve the same end.) This condition puts Eq. (28) in the form of a weakly nonlinear wave equation,

$$\varepsilon E_{\alpha\beta}[\bar{h}^1] + \varepsilon^2 E_{\alpha\beta}[\bar{h}^2] = 2\varepsilon^2 \delta^2 G_{\alpha\beta}[h^1] + \mathcal{O}(\varepsilon^3), \quad (32)$$

where

$$E_{\alpha\beta}[\bar{h}] := \nabla^\mu \nabla_\mu \bar{h}_{\alpha\beta} + 2R^\mu{}_\alpha{}^\nu{}_\beta \bar{h}_{\mu\nu}. \quad (33)$$

Equation (32) is the desired relaxed equation; it can be solved for arbitrary  $z^\mu$ . We can now equate coefficients

<sup>4</sup>  $\log \varepsilon$  terms also generically appear in this expansion [30]. For visual simplicity, we hide these inside the coefficients  $h_{\alpha\beta}^n$ .



of powers of  $\varepsilon$  to obtain a solution that is valid for *all*  $z^\mu$ , yielding  $G_{\alpha\beta}[g] = 0$  for the background metric and

$$E_{\alpha\beta}[\bar{h}^1] = 0, \quad (34)$$

$$E_{\alpha\beta}[\bar{h}^2] = 2\delta^2 G_{\alpha\beta}[h^1] \quad (35)$$

for the perturbations.

The solutions to Eqs. (34)–(35) are required to satisfy the matching condition. We can enforce that condition using a puncture scheme, which amounts to replacing the small object with a singular puncture in the space-time geometry. The matching condition dictates that in a neighbourhood of  $z^\mu$ , the metric perturbations are required to satisfy

$$h_{\alpha\beta}^1 = h_{\alpha\beta}^{1\mathcal{P}} + h_{\alpha\beta}^{1\mathcal{R}}, \quad (36)$$

$$h_{\alpha\beta}^2 = h_{\alpha\beta}^{2\mathcal{P}} + h_{\alpha\beta}^{2\mathcal{R}}, \quad (37)$$

where  $h_{\alpha\beta}^{n\mathcal{P}}$  are the puncture fields, which diverge on  $z^\mu$ , and  $h_{\alpha\beta}^{n\mathcal{R}}$  are the residual fields, which satisfy certain regularity conditions on  $z^\mu$ . The puncture fields capture the dominant physical behavior of the metric near the small object, and they encode the object's multipole moments. If we define  $z^\mu$  to be the object's center of mass, then for a generic compact object with spin  $s^\mu$ , the punctures have the schematic structure

$$h_{\alpha\beta}^{1\mathcal{P}} \sim \frac{\mu}{|x^\mu - z^\mu|} + \mathcal{O}(|x^\mu - z^\mu|^0), \quad (38)$$

$$h_{\alpha\beta}^{2\mathcal{P}} \sim \frac{\mu^2 + s^\mu}{|x^\mu - z^\mu|^2} + \frac{\mu h_{\alpha\beta}^{1\mathcal{R}}}{|x^\mu - z^\mu|} + \mathcal{O}(|x^\mu - z^\mu|^0), \quad (39)$$

given explicitly in covariant form in Ref. [32] (with the exception of the  $s^\mu$  term, given in local coordinates centered on  $z^\mu$  in [30]). With the punctures known, we can move them to the right-hand side of the field equations and solve for the residual fields. This can be done in two equivalent ways: with a worldtube [74] or with a window function [75]. Here for simplicity we adopt the window method, in which one makes the puncture go to zero (smoothly or sharply) outside some region  $\Gamma$  around  $z^\mu$ . The field equations then take the form

$$E_{\alpha\beta}[\bar{h}^{1\mathcal{R}}] = -E_{\alpha\beta}[\bar{h}^{1\mathcal{P}}], \quad (40)$$

$$E_{\alpha\beta}[\bar{h}^{2\mathcal{R}}] = 2\delta^2 G_{\alpha\beta}[h^1] - E_{\alpha\beta}[\bar{h}^{2\mathcal{P}}]. \quad (41)$$

Outside  $\Gamma$ , the residual field  $\bar{h}_{\alpha\beta}^{n\mathcal{R}}$  reduces to the physical field  $\bar{h}_{\alpha\beta}^n$ .

A solution to the wave equations (40)–(41) only becomes a solution to the non-relaxed Einstein equations if it also satisfies the gauge constraint (31). By imposing this condition on the fields near  $z^\mu$ , one determines that for a nonspinning, approximately spherical object, the object's center-of-mass trajectory is governed by the equation of motion [26, 31]

$$\frac{D^2 z^\mu}{d\tau^2} = -\frac{1}{2} P^{\mu\nu} (g_\nu{}^\rho - h_\nu{}^{\mathcal{R}\rho}) (2h_{\beta\rho;\alpha}^{\mathcal{R}} - h_{\alpha\beta;\rho}^{\mathcal{R}}) u^\alpha u^\beta + \mathcal{O}(\varepsilon^3), \quad (42)$$

where  $P^{\mu\nu} := g^{\mu\nu} + u^\mu u^\nu$  and  $h_{\alpha\beta}^{\mathcal{R}} = \varepsilon h_{\alpha\beta}^{1\mathcal{R}} + \varepsilon^2 h_{\alpha\beta}^{2\mathcal{R}}$ . The punctures then move on this trajectory. The gauge condition also determines  $d\mu/d\tau = \mathcal{O}(\varepsilon^3)$  and  $Ds^\mu/d\tau = \mathcal{O}(\varepsilon^3)$ . Note that here and throughout this paper,  $\tau$  denotes proper time in  $g_{\mu\nu}$ , and  $u^\mu := dz^\mu/d\tau$  denotes the four-velocity normalized in  $g_{\mu\nu}$ .

The coupled set of equations (40)–(42) represent the self-consistent evolution scheme through second order. By sidestepping a Taylor expansion of  $z^\mu$ , this scheme avoids the associated, secularly growing errors that would occur in an ordinary, Gralla-Wald-type perturbative approximation. In the next section, we will describe the remaining secular errors that nevertheless *do* arise in this scheme.

However, before doing so we remark on how the puncture formulation of self-force theory, used in the description above, relates to a point-particle description. In the field equations (40),  $E_{\alpha\beta}[\bar{h}^{1\mathcal{P}}]$  is defined off  $z^\mu$  as an ordinary function and defined on  $z^\mu$  by taking the limit from off  $z^\mu$ . This makes the source terms regular at  $z^\mu$ , allowing us to enforce regularity of  $\bar{h}_{\alpha\beta}^{1\mathcal{R}}$ . However, if we move the puncture  $\bar{h}_{\alpha\beta}^{1\mathcal{P}}$  back to the left-hand side and treat the derivatives in  $E_{\alpha\beta}[\bar{h}^{1\mathcal{P}}]$  distributionally, then we find that the total field  $\bar{h}_{\alpha\beta}^1 = \bar{h}_{\alpha\beta}^{1\mathcal{P}} + \bar{h}_{\alpha\beta}^{1\mathcal{R}}$  is precisely the field of a point mass, replacing Eq. (40) with

$$E_{\alpha\beta}[\bar{h}^1] = -16\pi T_{\alpha\beta}^1, \quad (43)$$

where

$$T_{\alpha\beta}^1 = \mu \int u_\alpha u_\beta \frac{\delta^4 [x^\mu - z^\mu(\tau)]}{\sqrt{-\det g}} d\tau \quad (44)$$

is the stress-energy tensor of a point mass  $\mu$  moving on  $z^\mu$  in the background  $g_{\mu\nu}$ . Therefore, at first order the puncture scheme is equivalent to approximating the small object as a point particle.

At second order in a generic gauge, the strongly singular behavior of  $\delta^2 G_{\alpha\beta}[h^1]$  prevents us from straightforwardly writing a unique distributional source for the total field  $\bar{h}_{\alpha\beta}^2 = \bar{h}_{\alpha\beta}^{2\mathcal{P}} + \bar{h}_{\alpha\beta}^{2\mathcal{R}}$  [49] (see also Sec. 1.2.5 of Ref. [37]). However, forthcoming work [76] will show that there exists a canonical distributional interpretation of  $\delta^2 G_{\alpha\beta}[h^1]$ , under which the field equation takes the form

$$E_{\alpha\beta}[\bar{h}^2] = -16\pi T_{\alpha\beta}^2 + 2\delta^2 G_{\alpha\beta}[h^1] \quad (45)$$

with

$$T_{\alpha\beta}^2 = -\frac{\mu}{2} \int_\gamma u_\alpha u_\beta Q^{\rho\sigma} h_{\rho\sigma}^{1\mathcal{R}} \frac{\delta^4 [x^\mu - z^\mu(\tau)]}{\sqrt{-\det g}} d\tau. \quad (46)$$

Here  $Q^{\rho\sigma} := g^{\rho\sigma} - u^\rho u^\sigma$ . The sum  $T_{\alpha\beta} = \varepsilon T_{\alpha\beta}^1 + \varepsilon^2 T_{\alpha\beta}^2 + \mathcal{O}(\varepsilon^3)$  can also be written as what we will call the *De-tweiler stress-energy*

$$T_{\alpha\beta} = \int \mu \tilde{u}_\alpha \tilde{u}_\beta \frac{\delta^4 [x^\mu - z^\mu(\tilde{\tau})]}{\sqrt{-\det \tilde{g}}} d\tilde{\tau}, \quad (47)$$



where  $\tilde{u}_\alpha := \tilde{g}_{\alpha\beta} \frac{dz^\beta}{d\tilde{\tau}}$  is normalized in the effective metric  $\tilde{g}_{\alpha\beta} = g_{\alpha\beta} + h_{\alpha\beta}^{\mathcal{R}}$ , and  $\tilde{\tau}$  is proper time in that metric. This stress-energy, introduced in Ref. [28], is that of a point mass  $\mu$  in  $\tilde{g}_{\alpha\beta}$ . Ref. [76]'s proof of its validity in Eq. (45) is based on the class of *highly regular* gauges derived in Ref. [31], and there is considerable subtlety in meaningfully extending it outside that class. We will not explore these subtleties here, focusing instead on the more well-developed puncture formulation of the second-order problem. However, we will also make some use of this point-particle formulation.

If the object has a spin  $s^\alpha$ , then the spin term in the puncture (39) is equivalent to an additional second-order stress-energy

$$T_{\alpha\beta}^{2(\text{spin})} = \int u_{(\alpha} s_{\beta)}^\gamma \nabla_\gamma \frac{\delta^4 [x^\mu - z^\mu(\tau)]}{\sqrt{-\det g}} d\tau \quad (48)$$

where  $s_{\alpha\beta} := \epsilon_{\mu\alpha\beta\gamma} u^\mu s^\gamma$  and  $\epsilon_{\mu\alpha\beta\gamma}$  is the Levi-Civita tensor; see, e.g., Ref. [30]. We will not include this source in our two-timescale analysis, but doing so should be relatively straightforward [25].

## B. Inaccuracies due to the black hole's evolution

While accounting carefully for the evolution of  $z^\mu$ , our description above has ignored the evolution of another component of the system: the central black hole. We know that radiation falls into the black hole, causing it to evolve. The black hole's mass, for example, changes at an average rate [77, 78]  $\langle \frac{dM_{BH}}{dt} \rangle \sim \varepsilon^2 M^2 \partial h^1 \partial h^1$ . Over a radiation-reaction time  $t \sim M/\varepsilon$ , this accumulates to a change  $\delta M \sim \varepsilon M$ . Hence, on the radiation-reaction time scale, we expect the black hole mass to behave as  $M_{BH} \sim M + \delta M(\varepsilon t)$ , where  $M$  is a constant, zeroth-order mass; the correction  $\delta M$  remains  $\sim \varepsilon M$  over the inspiral, and its rate of change is small as well, such that  $dM_{BH}/dt \sim \varepsilon^2$ . Here and below we focus on the mass, but analogous statements apply to the black hole's spin, which is expected to behave as  $S_{BH} \sim aM + \delta S(\varepsilon t)$  with  $\delta S \sim \varepsilon M^2$ .

These changes in the black hole must manifest themselves in the metric outside the black hole. In a two-timescale expansion, we expect that  $h_{\mu\nu}^1$  will contain slowly evolving terms proportional to  $\delta \dot{M}$  and  $\delta S$ . This was the ansatz of Hinderer and Flanagan [13], and at least in the case of a Schwarzschild background, the ansatz is borne out by a complete analysis of the Einstein field equations, sketched in Ref. [39] and detailed later in this paper.

On the other hand, if we were to perform an ordinary Taylor expansion of the metric, then the evolving mass  $M_{BH} \sim M + \delta M(\varepsilon t)$  would become  $M_{BH} \sim M + \delta M(0) + \varepsilon \frac{d\delta M}{dt}(0)t$ . What was a slowly evolving first-order perturbation would hence become a linearly growing second-order perturbation. This approximation clearly goes bad on the radiation-reaction time, when

$\varepsilon t \sim M$  (although it remains accurate much longer than a Gralla-Wald-type expansion, which fails on the dephasing time  $t \sim M/\sqrt{\varepsilon}$  due to quadratic growth in  $z_1^\mu$ ).

The question, then, is how these effects appear in the self-consistent formalism. Do they evolve dynamically on a long time scale, as they should in an accurate description, or do they grow linearly with time, at a forever-fixed rate, as they would in a strict Taylor series?

Although we cannot answer this with complete certainty, we can provide strong evidence that the self-consistent evolution produces the less accurate, linear-growth description of the black hole. Consider prescribing initial data on a spatial slice  $\Sigma$  and then evolving into the future of  $\Sigma$  using the self-consistent equations (43), (45), and (42) [or (40)–(42)]. Given a retarded Green's function  $G_{\alpha\beta\alpha'\beta'}$  for the operator  $E_{\alpha\beta}$ , we can write the solution to this Cauchy-type problem as [30]

$$\bar{h}_{\alpha\beta}^1 = \int_\Sigma \left( \bar{h}_{\alpha'\beta'}^1 \nabla_{\gamma'} G_{\alpha\beta}^{\alpha'\beta'} - G_{\alpha\beta}^{\alpha'\beta'} \nabla_{\gamma'} \bar{h}_{\alpha'\beta'}^1 \right) d\Sigma^{\gamma'} - 16\pi \int_V G_{\alpha\beta}^{\alpha'\beta'} T_{\alpha'\beta'}^1 dV', \quad (49)$$

$$\bar{h}_{\alpha\beta}^2 = \int_\Sigma \left( \bar{h}_{\alpha'\beta'}^2 \nabla_{\gamma'} G_{\alpha\beta}^{\alpha'\beta'} - G_{\alpha\beta}^{\alpha'\beta'} \nabla_{\gamma'} \bar{h}_{\alpha'\beta'}^2 \right) d\Sigma^{\gamma'} + \int_V G_{\alpha\beta}^{\alpha'\beta'} (2\delta^2 G_{\alpha'\beta'} - 16\pi T_{\alpha'\beta'}^2) dV', \quad (50)$$

where  $V$  is the future Cauchy development of  $\Sigma$  and primes denote quantities at the integration point. This form splits the solution into two pieces: the integral over  $V$ , which is a particular solution to the wave equation (43) or (45) with zero initial data; and the integral over  $\Sigma$ , which is a homogeneous solution to the wave equation with the prescribed initial data. The key fact is the following: the linearly growing perturbation proportional to  $\delta \dot{M}|_\Sigma$  is a *homogeneous* solution to the wave equation, and for physical initial data, it always arises from the first integral in Eq. (50).

To illustrate this, we take the background spacetime to be Schwarzschild and  $\Sigma$  to be the surface  $t = 0$ . The initial data comprises the metric perturbation and its  $t$  derivative at  $t = 0$ . If the data describes a slice of time in a physical inspiral, then it must include a linear metric perturbation induced by  $\delta M(0)$ . If  $\delta \dot{M}$  were constant in time, then that perturbation could be written as  $2g_{\alpha\beta} \delta \dot{M}(0)$ , a homogeneous Lorenz-gauge perturbation (reviewed in Appendix D). But we must also include a nonzero  $t$  derivative  $\delta \dot{M}(0)$  in the data. Imposing the Lorenz gauge condition at  $t = 0$  then demands that we add a  $t$ - $r$  component to the initial field, such that  $h_{\alpha\beta}^{\delta M}(0) = 2g_{\alpha\beta} \delta M(0) + \varepsilon p_{\alpha\beta}$ , where  $p_{\alpha\beta} = 2\delta \dot{M}(0)(r - 3M)f^{-1}\delta_{(\alpha}^t \delta_{\beta)}^r$ . This form holds true even if we choose initial data with  $\delta M(0) = 0$ , because we do not have the freedom to set  $\delta \dot{M}(0)$  to zero; this is shown explicitly in Sec. VII. Finally, with  $h_{\alpha\beta}^{\delta M}(0)$  as initial data, it is straightforward to show that the unique homogeneous solution to  $E_{\alpha\beta}[\bar{h}] = 0$  is  $h_{\alpha\beta}^{\delta M}(t) = h_{\alpha\beta}^{\delta M}(0) + 2\varepsilon g_{\alpha\beta} \delta \dot{M}(0)t$ .

Hence, unless the volume integral in Eq. (50) somehow miraculously cancels this homogeneous perturbation, the presence of a time-varying mass perturbation in the initial data automatically induces a linearly growing perturbation  $h_{\alpha\beta}^{\delta M}(t)$ .

The induced perturbation  $h_{\alpha\beta}^{\delta M}(t)$  is a physical one. It satisfies the Lorenz gauge condition in addition to the wave equation, and it correctly encodes the evolution of the black hole's mass. On time scales much shorter than the radiation-reaction time, it is perfectly suitable. But on the time scale of an inspiral, it becomes inaccurate. Moreover, it clearly does not admit a two-timescale expansion, meaning we cannot take the existing self-consistent expansion as our starting point for the two-timescale approximation. So we must (slightly) reformulate the self-consistent expansion.

### C. Self-consistent approximation on an evolving background

At its core, the self-consistent approximation is based on a division of the system into its *field* degrees of freedom and its *mechanical* degrees of freedom. The field degrees of freedom are contained in the metric  $g_{\alpha\beta}$ , and the mechanical degrees of freedom are in the trajectory  $z^\mu$ . The piece of the Einstein field equations that leaves  $z^\mu$  unconstrained can be solved perturbatively by expanding  $g_{\alpha\beta}$  in a suitable perturbative series. The piece of the field equations that constrains  $z^\mu$ , on the other hand, can be used to obtain an approximate equation of motion for  $z^\mu$  *without* expanding  $z^\mu$  in a perturbative series.

However, our analysis above makes clear that we are putting different mechanical degrees of freedom on unequal footing: while we derive and actively enforce a uniformly accurate evolution equation for the nonperturbative  $z^\mu$ , we leave the evolution of the black hole parameters to be determined passively, and we end up with an approximation that is not uniformly accurate over long time scales. To obtain an accurate evolution, we must treat the black hole parameters in the same way we have treated  $z^\mu$ .

This motivates a generalization of the self-consistent framework. Rather than dividing the system into the metric and the trajectory, we divide it into the metric and the complete list of mechanical degrees of freedom: the effective positions and full set of multipole moments of both objects. We write the effective position of the large black hole as  $Z^\mu$ , and its set of multipole moments collectively as  $\hat{M}$ . Analogously, we write the effective position of the small object as  $z^\mu$ , and its multipole moments collectively as  $\hat{m}$ . The full set of mechanical degrees of freedom is  $\hat{P} = \{Z^\mu, \hat{M}, z^\mu, \hat{m}\}$ .

We now split the metric into a background plus a perturbation,

$$g_{\alpha\beta} = \hat{g}_{\alpha\beta}(x^\mu; Z, \hat{M}) + \hat{h}_{\alpha\beta}(x^\mu, \varepsilon; \hat{P}). \quad (51)$$

Here the background metric  $\hat{g}_{\alpha\beta}(x^\mu; Z, \hat{M})$  describes an evolving black hole spacetime, and  $\hat{h}_{\alpha\beta}(x^\mu, \varepsilon; \hat{P})$  is the correction due to both the small object and the background's evolution. We imagine prescribing some concrete form for  $\hat{g}_{\alpha\beta}(x^\mu; Z, \hat{M})$  in terms of moments with an unspecified time dependence. For concreteness, we take it to be a Kerr metric with mass  $M_{\text{BH}}$  and spin  $S_{\text{BH}}$  treated as arbitrary (order- $\mu^0$ ) functions of some time parameter. All its higher moments are then determined by the Kerr relationship  $M_\ell + iS_\ell = M_{\text{BH}}(iS_{\text{BH}}/M_{\text{BH}})^\ell$ , where  $M_\ell$  denotes the  $\ell$ th mass moment and  $S_\ell$  the  $\ell$ th current moment. This choice  $\hat{g}_{\alpha\beta}$  sets  $Z^\mu$  to be at the “origin” of the Boyer-Lindquist coordinates, and the black hole's spin direction to be along the  $\theta = 0$  axis. In reality the large black hole's position and spin direction will be slightly perturbed by the small object's gravity, but these corrections are gauge perturbations (removable by a time-dependent translation and rotation of the coordinates) that we can incorporate into  $\hat{h}_{\alpha\beta}$ . We also specialize the small object to be compact, with a radius  $\rho \sim \mu$ , implying that its  $\ell$ th moments scale as  $\mu\rho^\ell \sim \mu^{\ell+1}$ .

For simplicity, we write the field equations formally with a stress-energy tensor rather than a puncture, as

$$G_{\alpha\beta}[\hat{g}] + \delta G_{\alpha\beta}[\hat{h}; \hat{g}] + S_{\alpha\beta}[\hat{h}; \hat{g}] = 8\pi T_{\alpha\beta} \quad (52)$$

where  $T_{\alpha\beta} = T_{\alpha\beta}[z, \hat{m}, \hat{g} + \hat{h}^{\mathcal{R}}]$  is a point stress-energy associated with the small object; through order  $\varepsilon^2$ , it will be the sum of (47) plus (48) with  $\tilde{g}_{\alpha\beta}$  replaced by  $\hat{g}_{\alpha\beta} + \hat{h}_{\alpha\beta}^{\mathcal{R}}$ .  $S_{\alpha\beta}[\hat{h}; \hat{g}]$  denotes the sum of all the terms in  $G_{\alpha\beta}[\hat{g} + \hat{h}]$  that are nonlinear in  $\hat{h}_{\alpha\beta}$ . As mentioned above, a field equation of this form can be made well defined at least through order  $\varepsilon^2$  [76].

Due to the black hole's evolution,  $\hat{g}_{\alpha\beta}$  is not precisely a vacuum metric, and we move its Einstein tensor to the right-hand side of the field equations, treating it as a source for the metric perturbation. We also once again impose the Lorenz gauge condition, defining the trace reversal and divergence with respect to  $\hat{g}_{\alpha\beta}$ , such that the field equations become

$$\hat{E}_{\alpha\beta}[\hat{h}] = 2G_{\alpha\beta}[\hat{g}] + 2S_{\alpha\beta}[\hat{h}; \hat{g}] - 16\pi T_{\alpha\beta}. \quad (53)$$

Here  $\hat{E}_{\alpha\beta}$  is given by Eq. (33) with the covariant derivatives and Riemann tensor associated with  $\hat{g}_{\alpha\beta}$ . We write Eq. (53) more compactly as

$$\hat{E}_{\alpha\beta}[\hat{h}] = \hat{S}_{\alpha\beta}[Z, \hat{M}, z, \hat{m}, \hat{h}]. \quad (54)$$

In the true evolution,  $\hat{g}_{\alpha\beta}$  is only *slowly* varying, meaning it is almost a vacuum metric. In that case, every term in the source  $\hat{S}_{\alpha\beta}$  is small. We have suggestively written the source as a function of  $(Z, \hat{M})$  instead of  $\hat{g}_{\alpha\beta}$ , assuming the background metric is a prescribed function of  $(Z, \hat{M})$ .

We can construct a solution to Eq. (54) iteratively, introducing one additional multipole moment at each iteration and solving

$$\hat{E}_{\alpha\beta}[\hat{h}^{(0)}] = \hat{S}_{\alpha\beta}[Z, \hat{M}, z, \hat{m} = 0, \hat{h} = 0] \quad (55)$$

for the zeroth iteration and

$$\hat{E}_{\alpha\beta}[\hat{h}^{(n)}] = \hat{S}_{\alpha\beta}[Z, \hat{M}, z, \hat{m}^{(n-1)}, \hat{h}^{(n-1)}] \quad (56)$$

for subsequent iterations. Here  $\hat{m}^{(n-1)}$  stands for the set of all of the small object's multipole moments up to its  $(n-1)$ th moment. Adopting the retarded solution at every iteration, we obtain

$$\hat{h}_{\alpha\beta}^{(0)} = 2 \int \hat{G}_{\alpha\beta}{}^{\alpha'\beta'} G_{\alpha'\beta'}[\hat{g}] dV', \quad (57)$$

$$\begin{aligned} \hat{h}_{\alpha\beta}^{(1)} = 2 \int \hat{G}_{\alpha\beta}{}^{\alpha'\beta'} \{ & G_{\alpha'\beta'}[\hat{g}] + S_{\alpha'\beta'}[\hat{h}^{(0)}; \hat{g}] \\ & - 8\pi T_{\alpha'\beta'}[z, \mu, \hat{g} + \hat{h}^{(0)}] \} dV', \end{aligned} \quad (58)$$

$$\begin{aligned} \hat{h}_{\alpha\beta}^{(2)} = 2 \int \hat{G}_{\alpha\beta}{}^{\alpha'\beta'} \{ & G_{\alpha'\beta'}[\hat{g}] + S_{\alpha'\beta'}[\hat{h}^{(1)}; \hat{g}] \\ & - 8\pi T_{\alpha'\beta'}[z, \mu, s, \hat{g} + \hat{h}^{(1)}] \} dV', \end{aligned} \quad (59)$$

where  $\hat{G}_{\alpha\beta}{}^{\alpha'\beta'}$  is the retarded Green's function associated with  $\hat{E}_{\alpha\beta}$ . In this way,  $\hat{h}_{\alpha\beta}^{(0)}$  is fully determined by  $Z^\mu$  and  $\hat{M}$ ;  $\hat{h}_{\alpha\beta}^{(1)}$  by  $Z^\mu$ ,  $\hat{M}$ ,  $z^\mu$ , and  $\mu$ ; and  $\hat{h}_{\alpha\beta}^{(2)}$  by  $Z^\mu$ ,  $\hat{M}$ ,  $z^\mu$ ,  $\mu$ , and the spin  $s^\mu$ . In the limit  $n \rightarrow \infty$ ,  $\hat{h}_{\alpha\beta}^{(n)}$  should converge to a solution to Eq. (54).

The iterative solution can be found (at least formally) for any behavior of the mechanical degrees of freedom  $\hat{P}$ . The true behavior of  $\hat{P}$  is then found by enforcing the Lorenz gauge condition, ensuring that the solution satisfies the full Einstein equation and not just the relaxed one. Here we will only conjecture that one can use the gauge condition to systematically derive uniform-in-time evolution equations for  $\hat{M}$ , which would complement the existing equations for  $z^\alpha$  and  $s^\alpha$ . It will become clear from the two-timescale analysis in Sec. VII that if these self-consistent evolution equations are to admit a two-timescale solution, they must be compatible with the expected laws: the black hole's mass and spin must evolve according to the gravitational-wave flux down the horizon.

From this iterative solution, if desired, one can extract an expansion analogous to (27),

$$\hat{h}_{\alpha\beta}^{(n)}(x^\mu, \varepsilon; \hat{P}) = \sum_{p=0}^n \varepsilon^p \hat{h}_{\alpha\beta}^{(n,p)}(x^\mu; \hat{P}), \quad (60)$$

where the terms in the expansion are obtained by reading off explicit factors of  $\varepsilon$  that come from the scaling of the multipole moments. For example,

$$\begin{aligned} \hat{h}_{\alpha\beta}^{(1,0)}(x^\mu; \hat{P}) = 2 \int \hat{G}_{\alpha\beta}{}^{\alpha'\beta'} \{ & G_{\alpha'\beta'}[\hat{g}] \\ & + S_{\alpha'\beta'}[\hat{h}^{(0)}; \hat{g}] \} dV', \end{aligned} \quad (61)$$

$$\hat{h}_{\alpha\beta}^{(1,1)}(x^\mu; \hat{P}) = -16\pi \int \hat{G}_{\alpha\beta}{}^{\alpha'\beta'} T_{\alpha\beta}^1[z, \hat{g} + \hat{h}^{(0)}] dV', \quad (62)$$

where  $T_{\alpha\beta}^1[z, \hat{g} + \hat{h}^{(0)}]$  is given by Eq. (44) with the four-velocity, proper time, and metric determinant all defined

with respect to  $\hat{g}_{\alpha\beta} + \hat{h}_{\alpha\beta}^{(0)}$ . Note that in this power counting,  $\hat{h}_{\alpha\beta}^{(0)}$  is formally of order 1, and the source  $S_{\alpha\beta}[\hat{h}^{(0)}; \hat{g}]$  can be infinitely nonlinear in  $\hat{h}_{\alpha\beta}^{(0)}$ . However, once we substitute the true time dependence of the multipole moments, as determined from the gauge condition,  $\hat{h}_{\alpha\beta}^{(0)}$  will be small because  $G_{\alpha\beta}[\hat{g}]$  will be small. This means that in practice, one would be able to truncate  $S_{\alpha\beta}[\hat{h}^{(0)}; \hat{g}]$  at a finite power of  $\hat{h}_{\alpha\beta}^{(0)}$ .

We will not belabour this construction, as we will move onto a simpler one in the next section.

#### D. Self-consistent approximation with evolving mass and spin perturbations

The problem becomes more tractable if we note that over the course of an inspiral, the background mass and spin change only by an amount of order  $\varepsilon$ . We can therefore divide  $\hat{M} = (M_{\text{BH}}, S_{\text{BH}})$  into constant, zeroth-order parameters  $P = (M, S)$  and small, evolving corrections  $\delta P = (\delta M, \delta S)$ .

Given this division of the black hole mechanical degrees of freedom, we may re-expand the metric at fixed  $\{P, p\}$ , where  $p = \{z^\mu, \hat{m}, \delta P\}$  is the complete list of perturbative degrees of freedom. Equation (51) then becomes

$$\mathbf{g}_{\alpha\beta} = g_{\alpha\beta}(x^\mu; P) + h_{\alpha\beta}(x^\mu, \varepsilon; P, p), \quad (63)$$

where the background  $g_{\alpha\beta}$  is now an exact Kerr metric with parameters  $P$ , and the perturbation is  $h_{\alpha\beta} = \hat{h}_{\alpha\beta}(x^\mu, \varepsilon; P, p) + \delta g_{\alpha\beta}(x^\mu, \varepsilon; P, \delta P)$  with  $\delta g_{\alpha\beta} := \hat{g}_{\alpha\beta} - g_{\alpha\beta}$ .  $h_{\alpha\beta}$  contains all the information about the small mechanical parameters. Both pieces of the perturbation can be expanded with fixed mechanical parameters, as

$$\hat{h}_{\alpha\beta}(x^\mu, \varepsilon; P, p) = \sum_{n>0} \varepsilon^n \hat{h}_{\alpha\beta}^n(x^\mu; P, p), \quad (64)$$

$$\delta g_{\alpha\beta}(x^\mu, \varepsilon; P, p) = \sum_{n>0} \varepsilon^n g_{\alpha\beta}^n(x^\mu; P, \delta P), \quad (65)$$

where

$$g_{\alpha\beta}^n := \frac{1}{n!} \left[ \frac{d^n}{d\varepsilon^n} \hat{g}_{\alpha\beta}(M + \varepsilon \delta M, S + \varepsilon \delta S) \right]_{\varepsilon=0} \quad (66)$$

(holding  $\delta M$  and  $\delta S$  fixed while evaluating the derivatives). Or in total,

$$h_{\alpha\beta}(x^\mu, \varepsilon; P, \hat{p}) = \sum_{n>0} \varepsilon^n h_{\alpha\beta}^n(x^\mu; P, p). \quad (67)$$

At the first few orders, this reads

$$\begin{aligned} \mathbf{g}_{\alpha\beta} = g_{\alpha\beta}(x; P) + \varepsilon h_{\alpha\beta}^1(x; P, \delta P, z, \mu) \\ + \varepsilon^2 h_{\alpha\beta}^2(x; P, \delta P, z, \mu, s) + \mathcal{O}(\varepsilon^3). \end{aligned} \quad (68)$$

Just as in the case of an evolving background, we wish to treat the evolving  $\delta P$  as a source in the field equations.

If we impose the Lorenz gauge condition (31) on  $h_{\alpha\beta}$  and move  $\delta g_{\alpha\beta}$  to the right-hand side of the field equations, we obtain the relaxed Einstein equation

$$\varepsilon E_{\alpha\beta}[\hat{h}^1] + \varepsilon^2 E_{\alpha\beta}[\hat{h}^2] = 2\varepsilon^2 \delta^2 G_{\alpha\beta}[h^1] - \varepsilon E_{\alpha\beta}[\bar{g}^1] - \varepsilon^2 E_{\alpha\beta}[\bar{g}^2] + \mathcal{O}(\varepsilon^3). \quad (69)$$

Demanding that  $\hat{h}_{\alpha\beta}^n$  satisfies these field equations for arbitrary  $\hat{p}$ , we have

$$E_{\alpha\beta}[\hat{h}^1] = -16\pi T_{\alpha\beta}^1 - E_{\alpha\beta}[\bar{g}^1], \quad (70)$$

$$E_{\alpha\beta}[\hat{h}^2] = -16\pi T_{\alpha\beta}^2 + 2\delta^2 G_{\alpha\beta}[h^1] - E_{\alpha\beta}[\bar{g}^2]. \quad (71)$$

Or in a puncture scheme,

$$E_{\alpha\beta}[\hat{h}^{1\mathcal{R}}] = -E_{\alpha\beta}[\bar{g}^1] - E_{\alpha\beta}[\hat{h}^{1\mathcal{P}}], \quad (72)$$

$$E_{\alpha\beta}[\hat{h}^{2\mathcal{R}}] = 2\delta^2 G_{\alpha\beta}[h^1] - E_{\alpha\beta}[\bar{g}^2] - E_{\alpha\beta}[\hat{h}^{2\mathcal{P}}]. \quad (73)$$

We seek the retarded solutions to these equations for arbitrary  $\delta M$  and  $\delta S$ .

Evolution equations for  $z^\mu$ ,  $\mu$ ,  $s^\mu$ ,  $\delta M$ , and  $\delta S$  are then to be determined from the gauge condition (31). The local analysis near the small object, which led to Eq. (42),  $d\mu/d\tau = \mathcal{O}(\varepsilon^3)$ , and  $Ds^\mu/d\tau = \mathcal{O}(\varepsilon^3)$  in prior work, goes through essentially unchanged. A similar analysis near the horizon of the central black hole should yield (uniform-in-time) evolution equations for  $\delta M$  and  $\delta S$ . However, as in the case of an evolving background, we note that in order to admit a two-timescale solution, those evolution equations must take the expected form in terms of fluxes of energy and angular momentum down the horizon. We therefore put aside any further analysis of this modified scheme, satisfying ourselves that it is now, at least in principle, compatible with a two-timescale expansion. In Sec. VII we will then use the two-timescale scheme to derive the standard flux-balance equations for  $\delta M$  and  $\delta S$ .

At this stage, astute readers will have noticed that if we simply move  $g_{\mu\nu}^n$  back to the left-hand side of the field equations (72) and (73), then this new self-consistent framework recovers exactly the equations in the usual self-consistent scheme, (40) and (41). What, then, has changed? In the old scheme, we had no way of enforcing that  $h_{\mu\nu}^1$  includes a dynamically evolving  $g_{\mu\nu}^1$ . Our method of solving the field equations implicitly treated  $g_{\mu\nu}^1$  as the Taylor series  $g_{\mu\nu}^1|_{t=0} + \varepsilon t \frac{dg_{\mu\nu}^1}{dt}|_{t=0} + \mathcal{O}(\varepsilon^2)$ , such that the evolution of  $g_{\mu\nu}^1$  covertly slipped into  $h_{\mu\nu}^2$  in the form of a linearly growing perturbation. In the altered scheme, the perturbations  $g_{\mu\nu}^n$  instead play the role of punctures. Just as the usual punctures on  $z^\mu$  enforce the correct physical behaviour near the small object, the punctures  $g_{\mu\nu}^n$  enforce the correct long-term behavior. More concretely, if we refer back to the discussion in Sec. IIB, we see that the initial data which seeded the unwanted linear growth in  $h_{\mu\nu}^2$  will not be included in the initial data for  $\hat{h}_{\mu\nu}^2$ ; everything linear in  $\delta M$  and its derivatives will already be accounted for in  $h_{\mu\nu}^1$ .

However, for simplicity in later sections, we *will* move  $g_{\mu\nu}^n$  back to the left-hand side of the field equations, as in Eqs. (40)–(41). The reason this will pose no problem is that the two-timescale framework will automatically enforce the correct long-term behavior of the solution. In that context, the quantity  $x_{\mu\nu}$  that appears in the two-timescale solution will implicitly replace  $g_{\mu\nu}^1$ .

Before proceeding to the multiscale expansion, we make one comment. Rather than starting from the self-consistent framework, one could instead build a two-timescale approximation starting from the Gralla-Wald formalism [29, 40]. As discussed above, this approach cannot remain accurate on the radiation-reaction time. However, one could still utilize its results to inform a two-timescale expansion by performing a short-time re-expansion of the two-timescale variables around each fixed value of slow time and matching the results, term by term, to the Gralla-Wald variables.<sup>5</sup>

### III. OVERVIEW OF TWO-TIMESCALE EXPANSIONS

Before performing the two-timescale expansion of our system of equations, we begin with a review of the method.

Consider a differential equation

$$D\psi(t, x^a, \varepsilon) = S(t, x^a, \varepsilon), \quad (74)$$

where  $D$  is a linear differential operator,  $x^a$  are a set of  $n$  spatial coordinates, and  $\varepsilon$  is a small parameter. Suppose we know that the source  $S$  and solution  $\psi$  are characterized by a very nearly discrete frequency spectrum, but with a slowly evolving set of discrete frequencies  $\omega_k = k\omega \sim \varepsilon^0$ . For simplicity, assume the timescale over which  $\omega$  varies is  $\sim 1/\varepsilon$ , such that  $\omega = \omega(\varepsilon t)$  [or more generally,  $\omega = \omega(\varepsilon t, \varepsilon)$ ]. The system is then characterized by two timescales: a short time scale  $1/\omega \sim \varepsilon^0$ , and a long time scale  $1/\varepsilon$ . We can characterize the system's dependence on these time scales by introducing two new time variables: a *fast time*  $\varphi(t, \varepsilon) \sim \varepsilon^0 t$  given by  $\frac{d\varphi}{dt} = \omega$ ; and a *slow time*  $\tilde{t}(t, \varepsilon) = \varepsilon t$ .  $\tilde{t}$  changes appreciably only over the long time scale  $t \sim 1/\varepsilon$ , while  $\varphi$  changes on the time scale  $\sim 1/\omega$ .

Things very quickly go wrong if we attempt to solve Eq. (74) by assuming regular asymptotic expansions  $\psi(t, x^a, \varepsilon) = \sum_{n=0}^{\infty} \varepsilon^n \psi_n(t, x^a)$  and  $S(t, \varepsilon) = \sum_{n=0}^{\infty} \varepsilon^n S_n(t, x^a)$ . The exact, slowly evolving solution will contain oscillatory functions of the fast time, of the

<sup>5</sup> This would be made (slightly) more difficult again if Gralla's explicit second-order results in Ref. [29] were used, because Gralla's gauge appears to exhibit growth in time that would not arise even in a short-time expansion of the two-timescale metric; this extra growth can be inferred from the combination of Gralla's Eqs. (80) and (83) with his Eq. (B4)–(B6). The matching procedure would hence also require a gauge transformation.

form  $e^{ik\varphi(t,\varepsilon)}$ . If we expand this in a regular power series, using

$$\begin{aligned}\varphi &= \int_0^t \omega(\varepsilon t) dt \\ &= \int_0^t [\omega(0) + \varepsilon t \dot{\omega}(0) + \mathcal{O}(\varepsilon^2)] dt \\ &= \omega(0)t + \frac{1}{2}\varepsilon \dot{\omega}(0)t^2 + \mathcal{O}(\varepsilon^2),\end{aligned}\quad (75)$$

we obtain

$$e^{ik\varphi(t,\varepsilon)} = e^{ik\omega(0)t} \left[ 1 + \frac{ik}{2}\varepsilon \dot{\omega}(0)t^2 + \mathcal{O}(\varepsilon^2) \right]. \quad (76)$$

What was a sinusoid with a slowly varying frequency and constant amplitude has become a sinusoid with a constant frequency and growing amplitude. After a time  $t \sim 1/\varepsilon$ , the sinusoid's frequency  $\omega(0)$  will differ by order  $\varepsilon^0$  from the true frequency  $\omega(\varepsilon t)$ . After that same time, the sinusoid's amplitude will be  $\sim 1/\varepsilon \gg 1$ , rendering the expansion entirely invalid. For a similar reason, the  $\omega_k = 0$  piece of the solution also fails. A slowly varying solution of the form  $\psi_0(\varepsilon t)$  becomes  $\psi_0(0) + \varepsilon \dot{\psi}_0(0)t + \mathcal{O}(\varepsilon^2)$ . After a time  $t \sim 1/\varepsilon$ , every term in the expansion becomes of order  $\varepsilon^0$ . More dramatically, the expansion of  $\varphi$  in Eq. (75) breaks down on the shorter *dephasing time*  $t \sim 1/\sqrt{\varepsilon}$ ; this is especially disastrous in gravitational-wave data analysis, where accurate phasing is paramount.

The multiscale approximation method eliminates these errors by treating  $\tilde{t}$  and  $\varphi$  as independent variables, transforming the  $(n+1)$ -dimensional problem (74) into a more tractable  $(n+2)$ -dimensional problem.

First, we assume that  $\psi(t, x^a, \varepsilon)$  can be written as  $\psi(t, x^a, \varepsilon) = \tilde{\psi}(\varepsilon t, \varphi(t, \varepsilon), x^a, \varepsilon)$ ,<sup>6</sup> where the function  $\tilde{\psi}(\tilde{t}, x^a, \varphi, \varepsilon)$  has a regular asymptotic expansion at fixed  $\tilde{t}$ ,  $\varphi$ , and  $x^a$ ,

$$\tilde{\psi}(\tilde{t}, \varphi, x^a, \varepsilon) = \sum_{n=0}^{\infty} \varepsilon^n \tilde{\psi}_n(\tilde{t}, \varphi, x^a). \quad (77)$$

Each coefficient  $\psi_n$  is further assumed to be (i) a periodic function of  $\varphi$  and (ii) a bounded function of  $\tilde{t}$  (at fixed  $x^a$ ). Analogously, we write  $S(t, x^a, \varepsilon) = \tilde{S}(\varepsilon t, \varphi(t, \varepsilon), x^a, \varepsilon)$ , where  $\tilde{S}$  is the asymptotic series

$$\tilde{S}(\tilde{t}, \varphi, x^a, \varepsilon) = \sum_{n=0}^{\infty} \varepsilon^n \tilde{S}_n(\tilde{t}, \varphi, x^a). \quad (78)$$

<sup>6</sup> We can relax this to the weaker assumption that  $\psi(t, x^a, \varepsilon)$  is uniformly approximated, on an interval of time  $\sim 1/\varepsilon$ , by the asymptotic series (77). That is, if we define  $\tilde{\psi}_N(\tilde{t}, x^a, \varphi, \varepsilon) = \sum_{n=0}^N \varepsilon^n \tilde{\psi}_n(\tilde{t}, \varphi, x^a)$ , then  $\lim_{\varepsilon \rightarrow 0} \frac{\sup |\psi(t, x^a, \varepsilon) - \tilde{\psi}_N(\varepsilon t, \varphi(t, \varepsilon), x^a, \varepsilon)|}{\varepsilon^N} = 0$  for all integers  $N \geq 0$ , where the supremum is taken over  $t \in [T_1, T_2/\varepsilon]$  for fixed  $x^a$  and some fixed  $T_{1,2} \in \mathbb{R}$ .

These expansions can be compared to the self-consistent approximation, in which the expansions were carried out while holding fixed the mechanical functions rather than slow and fast times.

When we substitute  $\psi = \tilde{\psi}$  into Eq. (74), we can evaluate time derivatives using the chain rule,

$$\begin{aligned}\frac{d\psi}{dt} &= \frac{d\varphi}{dt} \frac{\partial \tilde{\psi}}{\partial \varphi} + \frac{d\tilde{t}}{dt} \frac{\partial \tilde{\psi}}{\partial \tilde{t}} \\ &= \omega \frac{\partial \tilde{\psi}}{\partial \varphi} + \varepsilon \frac{\partial \tilde{\psi}}{\partial \tilde{t}}.\end{aligned}\quad (79)$$

The operator  $D$  then becomes a differential operator  $\tilde{D}$  on the  $(n+2)$ -dimensional manifold charted by  $(\tilde{t}, \varphi, x^a)$ , and Eq. (74) becomes

$$\tilde{D}\tilde{\psi} = \tilde{S}. \quad (80)$$

This is completely equivalent to Eq. (74) when evaluated at  $\tilde{t} = \varepsilon t$  and  $\varphi = \varphi(t, \varepsilon)$ . But we now consider it as a partial differential equation rather than an ordinary one, with  $\tilde{t}$  and  $\varphi$  as independent variables, and with periodicity in  $\varphi$  providing an additional boundary condition. Thus, we effectively define  $\tilde{\psi}(\tilde{t}, \varphi, x^a, \varepsilon)$  to be the periodic-in- $\varphi$  solution to Eq. (80), to any desired order in  $\varepsilon$ . By construction, when evaluated at  $\tilde{t} = \varepsilon t$  and  $\varphi = \varphi(t, \varepsilon)$ ,  $\tilde{\psi}$  will then also be a solution to the original equation (74), to any desired order in  $\varepsilon$ , with the desired property of approximate periodicity.

To solve (80), we now fully expand it in powers of  $\varepsilon$  at fixed  $\tilde{t}$  and  $\varphi$ . In the case that  $\omega$  depends on both  $\tilde{t}$  and  $\varepsilon$ , this requires that we assume that it has an asymptotic expansion

$$\omega(\tilde{t}, \varepsilon) = \sum_{n=0}^{\infty} \varepsilon^n \omega_n(\tilde{t}). \quad (81)$$

Given the expansion (81), we can write Eq. (79) as

$$\frac{d\psi}{dt} = \omega_0 \partial_\varphi \tilde{\psi} + \varepsilon (\partial_{\tilde{t}} + \omega_1 \partial_\varphi) \tilde{\psi} + \mathcal{O}(\varepsilon^2). \quad (82)$$

Due to their periodicity, the coefficients can be expanded in a discrete Fourier series

$$\tilde{\psi}_n(\tilde{t}, \varphi, x^a) = \sum_{k=-\infty}^{+\infty} \tilde{\psi}_{nk}(\tilde{t}, x^a) e^{-ik\varphi}. \quad (83)$$

Analogously,

$$\tilde{S}_n(\tilde{t}, \varphi, x^a) = \sum_{k=-\infty}^{+\infty} \tilde{S}_{nk}(\tilde{t}, x^a) e^{-ik\varphi}. \quad (84)$$

This means the  $\varphi$  derivatives can be explicitly evaluated, giving us

$$\frac{d\psi_n}{dt} = \sum_k \left[ -ik\omega_0 \psi_{nk} + \varepsilon \tilde{\partial}_{\tilde{t}} \psi_{nk} + \mathcal{O}(\varepsilon^2) \right] e^{-ik\varphi}, \quad (85)$$

where  $\tilde{\partial}_{\tilde{t}} := \partial_{\tilde{t}} - ik\omega_1$ .

This treatment of derivatives leads to the operator  $\tilde{D}$  becoming  $\tilde{D}_{0,k} + \varepsilon \tilde{D}_{1,k} + \mathcal{O}(\varepsilon^2)$ , and the original differential equation (74) becoming

$$\sum_{n,j,k} \varepsilon^{n+j} [\tilde{D}_{j,k} \tilde{\psi}_{n,k}(\tilde{t}, x^a)] e^{-ik\varphi} = \sum_{n,k} \varepsilon^n \tilde{S}_{n,k}(\tilde{t}, x^a) e^{-ik\varphi}. \quad (86)$$

If we restrict this equation to the submanifold  $\tilde{t} = \varepsilon t$  and  $\varphi = \varphi(t, \varepsilon)$ , then we are not guaranteed that we can equate coefficients of explicit powers of  $\varepsilon$  or Fourier coefficients in this equation. However, since we define  $\tilde{\psi}$  to be the solution to Eq. (80) for *all*  $\tilde{t}$  and  $\varphi$ , we are given that guarantee. Hence, we have

$$\sum_{j+n'=n} \tilde{D}_{j,k} \tilde{\psi}_{n',k}(\tilde{t}, x^a) = \tilde{S}_{n,k}(\tilde{t}, x^a). \quad (87)$$

This is a set of partial differential equations on the new manifold charted by  $(\tilde{t}, x^a)$ . Solving these equations then determines the amplitudes' evolution with slow time. Restricting back to the submanifold defined by  $\tilde{t} = \varepsilon t$  and  $\varphi = \varphi(t, \varepsilon)$ , we obtain the desired quasiperiodic solution

$$\psi(t, x^a, \varepsilon) = \sum_{n,k} \varepsilon^n \tilde{\psi}_{n,k}(\varepsilon t, x^a) e^{-ik \int^t \omega(\varepsilon s, \varepsilon) ds}. \quad (88)$$

The description in this section is of the textbook method of multiple scales. In our particular problem, an alternative approach is to perform an expansion at fixed  $\omega$  instead of fixed  $\tilde{t}$ . The end result is then an expansion of the form  $\psi(t, x^a, \varepsilon) = \sum_{n,k} \varepsilon^n \tilde{\psi}_{n,k}(\omega(\varepsilon t, \varepsilon), x^a) e^{-ik \int^t \omega(\varepsilon s) ds}$ . This is one step less removed from the self-consistent method, as it treats the system's parameters nonperturbatively. We explore this approach in Appendix A.

Ultimately, the validity of expansions such as these should be established a posteriori by showing that (i) all the equations (87) can be solved, and (ii) the coefficients  $\tilde{\psi}_{n,k}$  are uniformly order 1 on any interval  $t \in [T_1, T_2/\varepsilon]$ ; i.e.,  $\lim_{\varepsilon \rightarrow 0} \sup |\tilde{\psi}_{n,k}(\varepsilon t)| < \infty$ . (ii) rules out behavior like  $\tilde{\psi}_{n,k} = 1/\tilde{t}$ , which would be very large when  $t \ll 1/\varepsilon$ .

#### IV. TWO-TIMESCALE EXPANSION OF THE ORBITAL DYNAMICS

We now apply the two-timescale prescription to the worldline  $z^\alpha$  and its equation of motion (42). Our presentation closely follows Sec. IVA of Ref. [35].

##### A. Parametrization of quasicircular orbits

Without loss of generality, we place the orbit in the equatorial plane, parametrizing it as

$$z^\mu(t, \varepsilon) = (t, r_p(t, \varepsilon), \pi/2, \phi_p(t, \varepsilon)). \quad (89)$$

To specialize to quasicircular orbits, we assume that the orbital radius  $r_p$  and orbital frequency  $\Omega := \frac{d\phi_p}{dt}$  evolve slowly, on the radiation-reaction timescale  $\Omega/\dot{\Omega} \sim 1/\varepsilon$ , with no oscillations on the orbital timescale  $1/\Omega \sim \varepsilon^0$ . More explicitly, we assume that they can be written as slowly varying functions

$$r_p(t, \varepsilon) = r_0(\varepsilon t) + \varepsilon r_1(\varepsilon t) + \mathcal{O}(\varepsilon^2), \quad (90)$$

$$\Omega(t, \varepsilon) = \Omega_0(\varepsilon t) + \varepsilon \Omega_1(\varepsilon t) + \mathcal{O}(\varepsilon^2). \quad (91)$$

Here we have introduced the natural slow time variable

$$\tilde{t}(t, \varepsilon) = \varepsilon t. \quad (92)$$

Our natural fast time variable is the orbital phase,

$$\phi_p(t, \varepsilon) = \int_0^t \Omega(s, \varepsilon) ds + \phi_p(0, \varepsilon). \quad (93)$$

As one might expect, for quasicircular inspirals the orbital dynamics is independent of this fast time; we will show this in later sections. However, the metric of the full, evolving system will have a periodic dependence on this phase, and it will play a critical role in the two-timescale expansion of the field equations.

##### B. Equation of motion

We now substitute our expansion of the worldline into the equation of motion (42). This will lead to a sequence of equations for each  $r_n$  and  $\Omega_n$ .

Written in terms of the non-affine parameter  $t$ , Eq. (42) reads

$$\frac{d^2 z^\mu}{dt^2} + U^{-1} \frac{dU}{dt} \frac{dz^\mu}{dt} + \Gamma_{\beta\gamma}^\mu \frac{dz^\beta}{dt} \frac{dz^\gamma}{dt} = U^{-2} f^\mu, \quad (94)$$

where  $U := u^t$ , and where  $f^\mu$  is the self force per unit mass, given by the right-hand side of Eq. (42).

To expand Eq. (94), we will require an expansion of  $U$ . The normalization condition  $u^\mu u^\nu g_{\mu\nu} = U^2 \frac{dz^\mu}{dt} \frac{dz^\nu}{dt} g_{\mu\nu} = -1$  gives us  $U^{-2} = -g_{\mu\nu} \frac{dz^\mu}{dt} \frac{dz^\nu}{dt}$ . Using  $\frac{dz^\mu}{dt} = (1, dr_p/dt, 0, \Omega)$ , Eqs. (90) and (91), and  $d/dt = \varepsilon d/\tilde{d}t$ , we obtain the expansion

$$U^{-2} = 1 - 3M/r_0 - 2\varepsilon r_0^2 \Omega_0 \Omega_1 + \mathcal{O}(\varepsilon^2). \quad (95)$$

This then yields

$$U = U_0 + \varepsilon U_1 + \mathcal{O}(\varepsilon^2), \quad (96)$$

where

$$U_0 = (1 - 3M/r_0)^{-1/2}, \quad (97)$$

$$U_1 = r_0^2 \Omega_0 \Omega_1 (1 - 3M/r_0)^{-3/2}. \quad (98)$$

We also require an expansion of  $f^\mu$ . As we mentioned above, the orbital dynamics is independent of the fast time  $\phi_p$ . We can hence write

$$f^\mu(t, \varepsilon) = \varepsilon \tilde{f}_1^\mu(\varepsilon t) + \varepsilon^2 \tilde{f}_2^\mu(\varepsilon t) + \mathcal{O}(\varepsilon^3). \quad (99)$$

The explicit terms in this expansion are given by Eq. (203) below.  $\tilde{f}_1^\mu(\tilde{t})$  is the standard first-order self-force on a circular orbit of radius  $r_0(\tilde{t})$ , but with the inclusion of terms proportional to  $M_1(\tilde{t})$  and  $S_1(\tilde{t})$ . Similarly,  $\tilde{f}_2^\mu(\tilde{t})$  is specified by  $r_{n \leq 1}(\tilde{t})$ ,  $M_{n \leq 2}(\tilde{t})$ , and  $S_{n \leq 2}(\tilde{t})$ .

As mentioned in the introduction, dissipative and conservative pieces of the force contribute to the long-term phase evolution at different orders. In the quasicircular case, the division into dissipative and conservative pieces is straightforward. The dissipative piece is the one that is antisymmetric under the time reversal  $(t, \phi) \rightarrow (-t, -\phi)$  (at fixed  $\tilde{t}$ ):

$$\tilde{f}_{n,\text{diss}}^\mu = (\tilde{f}_n^t, 0, 0, \tilde{f}_n^\phi). \quad (100)$$

Analogously, the conservative piece is the one that is symmetric under that time reversal:

$$\tilde{f}_{n,\text{cons}}^\mu = (0, \tilde{f}_n^r, 0, 0). \quad (101)$$

Finally, we substitute the expansions of  $z^\mu$ ,  $U$ , and  $f^\mu$  into the equation of motion (94). The result is a sequence of equations for  $\Omega_n$  and  $r_n$ , which take the form of (i) algebraic equations for  $\Omega_n$  in terms of  $r_n$  and (ii) differential equations for the slow time evolution of  $r_n(\tilde{t})$ .

At order  $\varepsilon^0$  the only nontrivial piece of Eq. (94) is the  $r$  component, from which we derive the relation

$$\Omega_0 = \sqrt{M/r_0^3}. \quad (102)$$

This is just the standard expression for the orbital frequency of a circular geodesic orbit of radius  $r_0$  in Schwarzschild spacetime.

At linear order in  $\varepsilon$ , from the  $t$  component of Eq. (94) we derive that

$$\frac{dr_0}{d\tilde{t}} = \frac{2(r_0 - 3M)^2(r_0 - 2M)}{M(r_0 - 6M)} \tilde{f}_1^t(\tilde{t}). \quad (103)$$

This tells us that at leading order, the slow evolution of the radius is completely determined by  $\tilde{f}_{1,\text{diss}}^\mu$ . We see that the evolution diverges when the orbit reaches the innermost stable circular orbit (ISCO),  $r_0 = 6M$ , signalling the breakdown of our quasicircular ansatz. Physically, this corresponds to the fact that as the small object approaches the ISCO, it transitions into a plunging orbit. The transition occurs over a new slow timescale [79], and a complete treatment of the inspiral would utilize an alternative approximation in the transition region. However, the transition to plunge should have a minimal impact on LISA data analysis, since it adds relatively little to the integrated SNR. So we expect that for the purpose of matched filtering, we can simply abort the waveform model at some time prior to the transition.

Next, still at linear order in  $\varepsilon$ , from the  $r$  component of Eq. (94) we find that

$$\Omega_1 = -\frac{1}{2f_0 r_0 \Omega_0} \left( U_0^{-2} \tilde{f}_1^r + \frac{3M r_1 f_0}{r_0^3} \right), \quad (104)$$

where  $f_0 := 1 - 2M/r_0$ . This equation tells us how the conservative piece of  $\tilde{f}_1^\mu$  affects the orbital frequency at a given orbital radius.

Now moving to quadratic order in  $\varepsilon$ , from the  $t$  component of Eq. (94), we obtain an equation for the slow evolution of  $r_1$  and  $\Omega_1$ , given by

$$\begin{aligned} \frac{2M}{r_0^2 f_0} \frac{dr_1}{d\tilde{t}} + r_0^2 U_0^2 \Omega_0 \frac{d\Omega_1}{d\tilde{t}} \\ = U_0^{-2} \tilde{f}_2^t + \frac{8(r_0 - 3M)^2(r_0 - M)}{r_0^2(r_0 - 2M)(r_0 - 6M)} r_1 \tilde{f}_1^t \\ - \left( \frac{r_0^3(r_0 - 9M)f_0}{M(r_0 - 6M)} + 2r_0^2 \right) \Omega_0 \Omega_1 \tilde{f}_1^t. \end{aligned} \quad (105)$$

Substituting our result (104) for  $\Omega_1$  then yields an equation for  $r_1(\tilde{t})$  alone:

$$\begin{aligned} \frac{dr_1}{d\tilde{t}} = \frac{2r_0 f_0 (r_0 - 3M)^2}{M(r_0 - 6M)} \tilde{f}_2^t + \frac{r_0^3(r_0 - 3M)}{M(r_0 - 6M)} \frac{d\tilde{f}_1^r}{d\tilde{t}} \\ + \frac{2r_0^2(r_0 - 3M)^2(r_0^2 - 6Mr_0 + 6M^2)}{M^2(r_0 - 6M)^2} \tilde{f}_1^r \tilde{f}_1^t \\ + \frac{4(r_0 - 3M)(r_0^2 - 10Mr_0 + 18M^2)}{M(r_0 - 6M)^2} r_1 \tilde{f}_1^t. \end{aligned} \quad (106)$$

We see here that the slow evolution of  $r_1$  and  $\Omega_1$  is driven by  $\tilde{f}_{2,\text{diss}}^\mu$ , the rate of change  $\partial_{\tilde{t}} \tilde{f}_{1,\text{cons}}^\mu$ , and products of  $\tilde{f}_{1,\text{diss}}^\mu$  with first-order time-symmetric quantities.

We could next write a formula relating  $\Omega_2$  to  $r_2$  and  $\tilde{f}_2^\mu$ ; this relation is contained in the order- $\varepsilon^2$  term in the  $r$  component of Eq. (94). However, without an equation for  $dr_2/d\tilde{t}$ , this relation would not contribute to a calculation of the gravitational waveform. Our primary requirement for that calculation is an accurate evolution of the waveform phase. As we sketched in Fig. 3, the waveform phase is obtained directly from the orbital phase  $\phi_p$ , which in turn is obtained from the orbital frequency via Eq. (93). Note that by changing the integration variable in that equation, we can also write  $\phi_p$  as an expansion of the form (3),

$$\begin{aligned} \phi_p(t, \varepsilon) = \frac{1}{\varepsilon} \int_0^{\varepsilon t} [\Omega_0(\tilde{s}) + \varepsilon \Omega_1(\tilde{s}) + \mathcal{O}(\varepsilon^2)] d\tilde{s} \\ + \phi_p(0, \varepsilon) \end{aligned} \quad (107)$$

$$:= \frac{1}{\varepsilon} [\phi_0(\varepsilon t) + \varepsilon \phi_1(\varepsilon t) + \mathcal{O}(\varepsilon^2)] + \phi_p(0, \varepsilon). \quad (108)$$

This equation, in combination with the above analysis, provides a simple demonstration of what is required for adiabatic and post-adiabatic accuracy:

1. The adiabatic approximation only requires  $\Omega_0(\tilde{t})$ . According to Eqs. (102) and (103), the only necessary input is  $\tilde{f}_{1,\text{diss}}^\mu$ .
2. The post-adiabatic approximation requires  $\Omega_1(\tilde{t})$  (in addition to  $\Omega_0$ ). According to Eqs. (104) and



(106), the necessary input is  $\tilde{f}_{2,\text{diss}}^\mu$  and both  $\tilde{f}_{1,\text{diss}}^\mu$  and  $\tilde{f}_{1,\text{cons}}^\mu$ .

In the next section, we will see how our results here link to the two-timescale expansion of the field equations. However, before proceeding we make one comment. Previous treatments of eternal circular orbits have typically expanded the orbit in powers of  $\varepsilon$  at fixed frequency [80]. This corresponds to setting  $\Omega_0 = \Omega$  and  $\Omega_n = 0$  for  $n > 0$ . Equation (104) then becomes a formula for  $r_1$  in terms of  $\tilde{f}_1^r$ :

$$r_1 = -\frac{r_0^3 \tilde{f}_1^r}{3MU_0^2 f_0}. \quad (109)$$

However, in the two-timescale expansion presented in this section, the relationship (109) can only be freely enforced at a single value of  $\tilde{t}$ , say  $\tilde{t}_0$ ;  $\Omega_n(\tilde{t}_0) = 0$  is simply a choice of initial condition for the evolution. After that initial time,  $\Omega_1$  changes with time in a uniquely determined way, as dictated by Eqs. (104) and (106). An alternative approach, as noted several times above, would be to perform all our expansions in powers of  $\varepsilon$  while holding  $\Omega$ , rather than  $\tilde{t}$ , fixed. Equation (109) would then automatically hold true throughout the evolution, with all functions on the right of the equation becoming functions of  $\Omega$ . We detail this approach in Appendix A.

## V. TWO-TIMESCALE EXPANSION OF THE FIELD EQUATIONS

We now turn to the expansion of the metric perturbation and its field equations. We first define our choice of slow and fast time variables off the worldline; our definitions are tailored to hyperboloidal slicings, but they are sufficiently general to encompass  $t$  slicing and null slicings. We then derive the two-timescale expansion of the field equations with this generic slicing. We close with discussions of transformations between modes on different slicings and boundary conditions for the expanded fields.

### A. Choice of slow and fast times off the worldline

To perform a two-timescale expansion for the metric perturbation, we need a suitable choice of fast and slow variables. One option is to remain with the choice in Eq. (92). But as shown in Ref. [35], an asymptotically null slicing improves the behavior of the second-order source and eliminate infrared divergences in the retarded integral of that source. The advantage of such a judicious choice was also highlighted by Mino and Price [52]. Since similar problems arise at the horizon, this suggests

taking our basic time variable to be [66]<sup>7</sup>

$$s := t - k(r^*). \quad (110)$$

Here, for suitable choices of height function  $k(r^*)$ , surfaces of constant  $s$  foliate the spacetime with horizon-penetrating hyperboloidal slices. Figure 1 shows an example of a constant- $s$  slice for a particular choice of height function.

In the present section, for the purposes of expanding the field and the field equations, we remain agnostic about our choice of time variable. We instead denote by  $s$  the generic time variable as defined in Eq. (110), leaving  $k$  unspecified. This means our versions of the field equations can be specialized to  $t$  slicing (with  $k = 0$ ), null slicing (with  $k = \pm r^*$ ), or any choice of hyperboloidal slicing.

Given our choice of basic time variable, we define our slow time to be  $\tilde{s} = \varepsilon s$ . As for the fast variable, we use the azimuthal phase  $\phi_p$ , as given in Eq. (93), but extended off the worldline such that it is constant on each slice of constant  $s$ :

$$\phi_p(s, \varepsilon) = \int_0^s \Omega(z, \varepsilon) dz + \phi_p(0, \varepsilon). \quad (111)$$

The next subsection further motivates these choices of slow and fast time based on the form of the source terms in the field equations.

### B. Tensor-harmonic decomposition

We now turn to the field equations (40)–(41). Before performing the two-timescale expansion, we first slightly modify the field equations and decompose them into a basis of tensor spherical harmonics.

Our first step is to add a gauge-damping term to the field equations, following Barack and Lousto [81, 82]. We do this by replacing  $E_{\alpha\beta}$  with

$$\check{E}_{\alpha\beta}[\bar{h}] := E_{\alpha\beta}[\bar{h}] + \frac{4M}{r^2} t_{(\alpha} \check{Z}_{\beta)}[\bar{h}], \quad (112)$$

where  $t_\alpha = -\delta_\alpha^t$ ,  $Z_\alpha[\bar{h}] := \nabla^\beta \bar{h}_{\alpha\beta}$ , and  $\check{Z}_\alpha$ 's components in Schwarzschild coordinates are  $\check{Z}_\alpha = (Z_r, 2Z_r, Z_\theta, Z_\phi)$ . The field equations then become

$$\check{E}_{\alpha\beta}[\bar{h}^{1\mathcal{R}}] = -\check{E}_{\alpha\beta}[\bar{h}^{1\mathcal{P}}] =: \check{S}_{\alpha\beta}^{1\text{eff}}, \quad (113)$$

$$\check{E}_{\alpha\beta}[\bar{h}^{2\mathcal{R}}] = 2\delta^2 G_{\alpha\beta}[h^1] - \check{E}_{\alpha\beta}[\bar{h}^{2\mathcal{P}}] =: \check{S}_{\alpha\beta}^{2\text{eff}}. \quad (114)$$

The solutions to these equations slightly differ from the solutions to the old ones because the individual fields do

<sup>7</sup> For the height function  $k(r)$ , to avoid confusion with the metric perturbation we use a notation different from that of Ref. [66], in which  $h(r)$  is used.

not satisfy  $Z_\alpha[\bar{h}^n] = 0$ . However, when the fields are summed,  $\varepsilon\bar{h}_{\alpha\beta}^1 + \varepsilon^2\bar{h}_{\alpha\beta}^2$  does remain a solution to the original relaxed Einstein equation (32) [or (69)] because the fields *do* satisfy  $\varepsilon Z_\alpha[\bar{h}^1] + \varepsilon^2 Z_\alpha[\bar{h}^2] = \mathcal{O}(\varepsilon^3)$ . This step of altering the field equations is entirely nonessential for our analysis; we perform it solely because, as pointed out by Barack and Lousto, it partially decouples the field equations after the tensor-harmonic decomposition, as we review later in this section.

We next decompose the fields into tensor spherical harmonic modes, using the Barack-Lousto-Sago basis of harmonics [81, 82]:

$$\bar{h}_{\alpha\beta}^{n\mathcal{R}} = \sum_{i\ell m} \frac{a_{i\ell}}{r} \bar{h}_{i\ell m}^{n\mathcal{R}}(s, r, \varepsilon) Y_{\alpha\beta}^{i\ell m}(r, \theta^A), \quad (115)$$

and analogously for  $\bar{h}_{\alpha\beta}^{n\mathcal{P}}$ , where  $i = 1, \dots, 10$ ,  $\ell \geq 0$ , and  $m = -\ell, \dots, \ell$ . The harmonics  $Y_{\alpha\beta}^{i\ell m}$ , given explicitly in Appendix B, provide an orthogonal basis for symmetric rank-2 tensors, satisfying the orthogonality property (B5).  $a_{i\ell}$  is a convenient numerical factor defined in Eq. (B8). Following Barack-Lousto-Sago, we have also pulled out a factor of  $1/r$  to simplify the field equations. The orthogonality property implies

$$\bar{h}_{i\ell m}^n = \frac{r}{a_{i\ell}\kappa_i} \oint d\Omega \eta^{\alpha\mu} \eta^{\beta\nu} \bar{h}_{\alpha\beta}^n Y_{\mu\nu}^{i\ell m*}, \quad (116)$$

with  $\oint d\Omega := \int_0^{2\pi} d\phi \int_0^\pi d\theta \sin\theta$ ,  $\kappa_i$  given by Eq. (B7), and  $\eta^{\alpha\beta}$  given by Eq. (B6). The coefficients  $h_{i\ell m}^n$ , if needed, are related to  $\bar{h}_{i\ell m}^n$  by the interchange  $i = 3 \leftrightarrow i = 6$ , with all other coefficients unchanged.

We similarly expand the source terms  $\check{S}_{\mu\nu}^{n\text{eff}}$  as

$$\check{S}_{\mu\nu}^{n\text{eff}} = \sum_{i\ell m} \frac{-4a_{i\ell}}{rf} S_{i\ell m}^{n\text{eff}}(s, r) Y_{\mu\nu}^{i\ell m}(r, \theta^A), \quad (117)$$

where we have again introduced a factor,  $\frac{-4a_{i\ell}}{rf}$ , to simplify later expressions. From the orthogonality condition (B5), the coefficients are

$$S_{i\ell m}^{n\text{eff}}(s, r) = \frac{-rf}{4a_{i\ell}\kappa_i} \oint d\Omega \eta^{\alpha\mu} \eta^{\beta\nu} \check{S}_{\alpha\beta}^{n\text{eff}} Y_{\mu\nu}^{i\ell m*}, \quad (118)$$

Here and below we define mode coefficients without breves to lessen the notational load.

With these harmonic expansions, Eqs. (113) and (114) each separate into a set of ten coupled partial differential equations for the coefficients  $\bar{h}_{i\ell m}^n$  [82], which read

$$E_{ij\ell m} \bar{h}_{j\ell m}^{n\mathcal{R}} = S_{i\ell m}^{n\text{eff}}. \quad (119)$$

Recall that the repeated basis label  $j$  is summed over. The decomposed wave operator is given by

$$E_{ij\ell m} \bar{h}_{j\ell m} := \square_{sc}^{2d} \bar{h}_{i\ell m} + \mathcal{M}^{ij} \bar{h}_{j\ell m}, \quad (120)$$

where  $\square_{sc}^{2d}$  is the scalar-field wave operator  $\partial_{uv} + V_\ell$ , or

$$\square_{sc}^{2d} = \frac{1}{4} (\partial_t^2 - \partial_{r^*}^2) + V_\ell, \quad (121)$$

with potential

$$V_\ell(r) = \frac{f}{4} \left( \frac{2M}{r^3} + \frac{\ell(\ell+1)}{r^2} \right), \quad (122)$$

and  $\mathcal{M}^{ij}$  with  $i, j = 1, \dots, 10$  are a set of matrices comprised of first-order differential operators that couple between the various  $\bar{h}_{j\ell m}$ 's. Note that the coupling is only between different  $j$ 's; there is no coupling between modes of different  $\ell$  and  $m$ . Also note that the only effect of our added gauge-damping term is to alter these coupling matrices. The explicit form of the coupling matrices can be found in Appendix A of [82].

In addition to these wave equations, we have the four gauge constraints,  $\nabla^\alpha \bar{h}_{\alpha\beta} = 0$ , which separate into four conditions for  $\bar{h}_{i\ell m}$ . These also appear in Appendix A of [82]. We delay discussion of them until after presenting the two-timescale expansion of the wave equations.

To more easily justify our two-timescale ansatz, rather than using the first-order field equation with an effective source, we use the equivalent form

$$E_{ij\ell m} \bar{h}_{j\ell m}^1 = -16\pi T_{i\ell m}^1, \quad (123)$$

where the modes of the stress-energy tensor are defined in analogy with Eq. (118). We can evaluate the integral in the point-particle stress-energy (44) by changing to  $t$  as the integration variable, yielding

$$T_{\alpha\beta}^1 = \frac{\mu u_\alpha(\varepsilon t, \varepsilon) u_\beta(\varepsilon t, \varepsilon)}{U(\varepsilon t, \varepsilon) r_p^2(\varepsilon t, \varepsilon)} \times \delta[r - r_p(\varepsilon t, \varepsilon)] \delta(\theta - \pi/2) \delta[\phi - \phi_p(t, \varepsilon)]. \quad (124)$$

The harmonic modes are then

$$\begin{aligned} T_{i\ell m}^1 &= \frac{-r\mu f}{4a_{i\ell}\kappa_i} \oint d\Omega \eta^{\alpha\mu} \eta^{\beta\nu} \frac{u_\alpha u_\beta}{U r_p^2} \delta(r - r_p) \delta(\theta - \pi/2) \\ &\quad \times \delta(\phi - \phi_p) Y_{\mu\nu}^{i\ell m*}(r_p, \theta, \phi) \\ &= t_{i\ell m}^1 e^{-im\phi_p} \delta(r - r_p), \end{aligned} \quad (125)$$

where

$$t_{i\ell m}^1 := \frac{-r_p \mu f_p}{4a_{i\ell}\kappa_i} \eta^{\alpha\mu} \eta^{\beta\nu} \frac{u_\alpha u_\beta}{U r_p^2} Y_{\mu\nu}^{i\ell m*}(r_p, \pi/2, 0). \quad (126)$$

Here  $\eta^{\alpha\beta}$  is evaluated at  $r = r_p$  and  $\theta = \pi/2$ , and all of  $r_p$ ,  $f_p := 1 - 2M/r_p$ ,  $U$ , and  $u_\alpha$  are evaluated at  $(\varepsilon t, \varepsilon)$ . Note, however, that since  $s$  reduces to  $t$  in a neighborhood of the worldline, we have  $s(t, r_p) = t$ . Since the radial delta function forces evaluation at  $r = r_p$ , we can hence freely treat all functions of  $t$  in  $T_{i\ell m}$ , including  $\phi_p$ , as functions of  $s$ . Each mode  $T_{i\ell m}$  is therefore proportional to the fast-time phase factor  $e^{-im\phi_p(s, \varepsilon)}$ . This motivates adopting  $\phi_p(s, \varepsilon)$  as our fast time in  $\bar{h}_{i\ell m}^1$ , as outlined above.

This feature extends to the second-order source as well. The source term  $\delta^2 G_{i\ell m}$  can be written in terms of the first-order field modes as

$$\delta^2 G_{i\ell m} = \sum_{i_1 \ell_1 m_1 \atop i_2 \ell_2 m_2} \mathcal{G}_{i_1 \ell_1 m_1 i_2 \ell_2 m_2}^{i\ell m} \bar{h}_{i_1 \ell_1 m_1}^1 \bar{h}_{i_2 \ell_2 m_2}^1, \quad (127)$$

where  $\mathcal{G}_{i_1\ell_1m_1i_2\ell_2m_2}^{ilm}$  is a bilinear differential operator that acts linearly on  $\bar{h}_{i_1\ell_1m_1}^1$  and, separately, linearly on  $\bar{h}_{i_2\ell_2m_2}^1$ . The explicit form of Eq. (127) will be presented in Ref. [71]. Its only essential characteristic here is that it enforces  $m = m_1 + m_2$ ; this condition arises in the usual way from the coupling of two spherical harmonics, via  $\delta^2 G_{ilm} \propto \int d\phi e^{-im\phi} e^{im_1\phi} e^{im_2\phi} d\phi \propto \delta_{m,m_1+m_2}$ . Since  $\bar{h}_{i_1\ell_1m_1}^1 \propto e^{-im_1\phi_p(s,\varepsilon)}$  and  $\bar{h}_{i_2\ell_2m_2}^1 \propto e^{-im_2\phi_p(s,\varepsilon)}$ , it follows that  $\delta^2 G_{ilm} \propto e^{-im\phi_p(s,\varepsilon)}$ . Although we do not explicitly calculate the modes of the puncture  $\bar{h}_{ilm}^{2\mathcal{P}}$  in this paper, they too inherit the fast-time dependence of the orbit. Hence, we are again motivated to adopt  $\phi_p(s, \varepsilon)$  as our fast time in  $\bar{h}_{\alpha\beta}^2$ .

In summary, our mode-decomposed field equations are

$$E_{ij\ell m} \bar{h}_{j\ell m}^1 = -16\pi t_{ilm}^1 e^{-im\phi_p(s,\varepsilon)} \delta(r - r_p), \quad (128a)$$

$$E_{ij\ell m} \bar{h}_{j\ell m}^{2\mathcal{R}} = 2\delta^2 G_{ilm} - E_{ij\ell m} \bar{h}_{j\ell m}^{2\mathcal{P}}, \quad (128b)$$

where  $r_p = r_p(\varepsilon s, \varepsilon)$ ,  $t_{ilm}^1 = t_{ilm}^1(\varepsilon s, \varepsilon)$  is given by Eq. (126), and  $\delta^2 G_{ilm}$  is given by Eq. (127). The first-order source is a function of  $\tilde{s}$ ,  $\phi_p$ , and  $\varepsilon$ , periodic in  $\phi_p$  (with period  $2\pi/m$ ). If we assume  $\bar{h}_{ilm}^1$  inherits those properties, then the second-order source  $\delta^2 G_{ilm}$  does as well.

### C. Two-timescale expansion

Prompted by the form of the source, and following the outline in Sec. III, we assume that  $\bar{h}_{ilm}^n(s, r, \varepsilon)$  is approximated (uniformly in  $s$  at fixed  $r$ ) by a function  $\tilde{\bar{h}}_{ilm}^n(\varepsilon s, \phi_p(s, \varepsilon), r, \varepsilon)$  that is periodic in  $\phi_p$ . Expanding that function in powers of  $\varepsilon$  gives us<sup>8</sup>

$$\begin{aligned} \tilde{\bar{h}}_{ilm}^n(\tilde{s}, \phi_p, r, \varepsilon) &= \tilde{\bar{h}}_{ilm}^n(\tilde{s}, \phi_p, r, 0) \\ &+ \varepsilon \partial_\varepsilon \tilde{\bar{h}}_{ilm}^n(\tilde{s}, r, \phi_p, 0) + \mathcal{O}(\varepsilon^2). \end{aligned} \quad (129)$$

We then move the first subleading term from  $\tilde{\bar{h}}_{ilm}^1$  into  $\tilde{\bar{h}}_{ilm}^2$ , defining new first- and second-order fields

$$\tilde{\bar{h}}_{ilm}^1(\tilde{s}, \phi_p, r) := \tilde{\bar{h}}_{ilm}^1(\tilde{s}, \phi_p, r, 0), \quad (130)$$

$$\begin{aligned} \tilde{\bar{h}}_{ilm}^2(\tilde{s}, \phi_p, r) &:= \tilde{\bar{h}}_{ilm}^2(\tilde{s}, \phi_p, r, 0) \\ &+ \partial_\varepsilon \tilde{\bar{h}}_{ilm}^1(\tilde{s}, \phi_p, r, 0). \end{aligned} \quad (131)$$

<sup>8</sup> These expressions assume smooth functions of  $\varepsilon$ . However, in reality,  $\ln \varepsilon$  terms appear in  $\tilde{\bar{h}}_{ilm}^2(\tilde{s}, \phi_p, r, \varepsilon)$ . Such logarithms arise from two effects: the mass  $\mu$  deforms the light cones on which solutions propagate, potentially introducing  $\ln(r/\mu)$  terms into the solution [30]; and the waves in  $h_{\alpha\beta}^1$  introduce curvature that falls off slowly at large distance, which in turn leads to hereditary effects that introduce  $\ln(\varepsilon r/M)$  terms into the solution [35]. Here we blithely absorb these logarithms into the ‘ $\varepsilon$ -independent’ coefficients. Note that this is simply a choice to streamline the notation; we do not discard logarithms from the solutions.

Given the  $2\pi/m$  periodicity of the source, we also adopt the ansatz

$$\tilde{\bar{h}}_{ilm}^n(\tilde{s}, \phi_p, r) = R_{ilm}^n(\tilde{s}, r) e^{-im\phi_p}. \quad (132)$$

Altogether, this means the two-timescale expansion of the total (trace reversed) metric perturbation is

$$\bar{h}_{\mu\nu} = \sum_{n\ell m} \frac{\varepsilon^n a_{i\ell}}{r} R_{ilm}^n(\varepsilon s, r) e^{-im\phi_p(s,\varepsilon)} Y_{\mu\nu}^{ilm}(r, \theta^A). \quad (133)$$

We assume analogous expansions of the puncture and residual fields.

Next we must expand  $E_{ij\ell m}$  and the sources  $T_{ilm}$  and  $\delta^2 G_{ilm}$ . We begin with  $E_{ij\ell m}$ , applying the chain rule for derivatives as described in Sec. III. In Eq. (121), we expressed  $\square_{sc}^{2d}$  in terms of  $t$  derivatives at fixed  $r$  and  $r^*$  derivatives at fixed  $t$ . When acting on a function of  $(\varepsilon s, \phi_p(s, \varepsilon), r)$ , a  $t$  derivative at fixed  $r$  becomes

$$\begin{aligned} \left( \frac{\partial}{\partial t} \right)_r &= \frac{\partial \phi_p}{\partial t} \frac{\partial}{\partial \phi_p} + \frac{\partial \tilde{s}}{\partial t} \frac{\partial}{\partial \tilde{s}} \\ &= \Omega \frac{\partial}{\partial \phi_p} + \varepsilon \frac{\partial}{\partial \tilde{s}} \\ &= \Omega_0 \partial_{\phi_p} + \varepsilon (\partial_{\tilde{s}} + \Omega_1 \partial_{\phi_p}) + \mathcal{O}(\varepsilon^2). \end{aligned} \quad (134)$$

Similarly,  $r^*$  derivatives act according to

$$\begin{aligned} \left( \frac{\partial}{\partial r^*} \right)_t &= \left( \frac{\partial}{\partial r^*} \right)_{\tilde{s}, \phi_p} + \frac{\partial \phi_p}{\partial r^*} \frac{\partial}{\partial \phi_p} + \frac{\partial \tilde{s}}{\partial r^*} \frac{\partial}{\partial \tilde{s}} \\ &= \left( \frac{\partial}{\partial r^*} \right)_{\tilde{s}, \phi_p} - H \Omega_0 \partial_{\phi_p} \\ &\quad - \varepsilon H (\Omega_1 \partial_{\phi_p} + \partial_{\tilde{s}}) + \mathcal{O}(\varepsilon^2), \end{aligned} \quad (135)$$

where we have defined

$$H(r^*) := \frac{dk}{dr^*} \quad (136)$$

and used  $\partial \phi_p / \partial r^* = (\partial \phi_p / \partial s) (\partial s / \partial r^*) = -H \Omega$  and  $\partial \tilde{s} / \partial r^* = -\varepsilon H$ .

From Eq. (134), a  $t$  derivative acting on a Fourier series of the form (132) becomes

$$\left( \frac{\partial}{\partial t} \right)_r \rightarrow -i\omega_m + \varepsilon \tilde{\partial}_{\tilde{s}} + \mathcal{O}(\varepsilon^2). \quad (137)$$

where  $\omega_m := m\Omega_0$  and

$$\tilde{\partial}_{\tilde{s}} := \partial_{\tilde{s}} - im\Omega_1. \quad (138)$$

Similarly, from Eq. (135), an  $r^*$  derivative becomes

$$\left( \frac{\partial}{\partial r^*} \right)_t \rightarrow \partial_{r^*} + i\omega_m H - \varepsilon H \tilde{\partial}_{\tilde{s}} + \mathcal{O}(\varepsilon^2), \quad (139)$$

where the radial derivative is taken at fixed  $\tilde{s}$  and  $\phi_p$ .

We are now ready to expand  $E_{ij\ell m}$ . Substituting Eqs. (137) and (139) into (121), we split the term  $\square_{sc}^{2d}$  into  $\square_{sc}^{2d} = \square_0 + \varepsilon \square_1 + \mathcal{O}(\varepsilon^2)$ , where

$$\square_0 = -\frac{1}{4} \left[ \partial_{r^*}^2 + i\omega_m (2H\partial_{r^*} + H') + (1 - H^2) \omega_m^2 - 4V_\ell(r) \right], \quad (140)$$

$$\square_1 = \frac{1}{4} \left[ (2H\partial_{r^*} + H') \tilde{\partial}_{\tilde{s}} - (1 - H^2) (2i\omega_m \tilde{\partial}_{\tilde{s}} + i\dot{\omega}_m) \right]. \quad (141)$$

Here

$$H' := \frac{dH}{dr^*} \quad \text{and} \quad \dot{\omega}_m := m\dot{\Omega}_0(\tilde{s}). \quad (142)$$

In the same way the coupling matrices are expanded as  $\mathcal{M}^{ij} = \mathcal{M}_0^{ij} + \varepsilon \mathcal{M}_1^{ij} + \mathcal{O}(\varepsilon^2)$ . The leading-order matrices  $\mathcal{M}_0^{ij}$  are obtained from  $\mathcal{M}^{ij}$  in Appendix A of [82] with the replacements  $\tilde{h}_i \rightarrow R_i$ ,  $\partial_r \rightarrow \partial_r + if^{-1}\omega_m H$  and  $\partial_v \rightarrow \frac{1}{2} [f\partial_r - i(1 - H)\omega_m]$ . We give the explicit forms of  $\mathcal{M}_0^{ij}$  and  $\mathcal{M}_1^{ij}$  in Eqs. (C1) and (C2).

Combining these results we obtain

$$E_{ij\ell m} \rightarrow E_{ij\ell m}^0 + \varepsilon E_{ij\ell m}^1 + \mathcal{O}(\varepsilon^2), \quad (143)$$

where

$$E_{ij\ell m}^n = \delta_{ij} \square_n + \mathcal{M}_n^{ij}. \quad (144)$$

For  $H = 0$  (i.e.,  $t$  slicing),  $E_{ij\ell m}^0$  is precisely the operator that appears in the standard frequency-domain Lorenz-gauge linearized field equations for a metric perturbation  $\tilde{h}_{\mu\nu} = \sum_{i\ell m} \frac{a_{i\ell}}{r} R_{i\ell m}(r) Y_{\mu\nu}^{i\ell m} e^{-i\omega_m t}$ , as in Refs. [34, 68]. For  $H \neq 0$ ,  $E_{ij\ell m}^0$  is precisely the operator that would appear in the frequency-domain Lorenz-gauge linearized field equations for a metric perturbation  $\tilde{h}_{\mu\nu} = \sum_{i\ell m} \frac{a_{i\ell}}{r} R_{i\ell m}(r) Y_{\mu\nu}^{i\ell m} e^{-i\omega_m s}$ .

We next expand the source terms in the field equations. The point-particle stress-energy modes (125) are expanded as

$$T_{i\ell m}^1 = \left[ \tilde{T}_{i\ell m}^1(\tilde{s}, r) + \varepsilon \tilde{T}_{i\ell m}^2(\tilde{s}, r) + \mathcal{O}(\varepsilon^2) \right] e^{-im\phi_p}, \quad (145)$$

where

$$\tilde{T}_{i\ell m}^1 = \tilde{t}_{i\ell m}^1(\tilde{s}) \delta[r - r_0(\tilde{s})], \quad (146)$$

$$\begin{aligned} \tilde{T}_{i\ell m}^2 &= \tilde{t}_{i\ell m}^2(\tilde{s}) \delta[r - r_0(\tilde{s})] \\ &\quad + \tilde{t}_{i\ell m}^1(\tilde{s}) r_1(\tilde{s}) \delta'[r - r_0(\tilde{s})]. \end{aligned} \quad (147)$$

$\tilde{t}_{i\ell m}^1(\tilde{s}) := t_{i\ell m}^1(\tilde{s}, 0)$  and  $\tilde{t}_{i\ell m}^2(\tilde{s}) := \frac{\partial t_{i\ell m}^1}{\partial \varepsilon}(\tilde{s}, 0)$  are easily written in terms of  $r_0$ ,  $r_1$ ,  $dr_0/d\tilde{s}$ , and  $\Omega_1$  using the explicit expression (126) and the expansions (90), (91), and (96).

The leading term (146) is identical to the stress-energy mode of a particle on a circular geodesic of radius  $r_0$ . Its

exact form, also given in [68, 82], is

$$\tilde{t}_{i\ell m}^1 = -\frac{1}{4} \mathcal{E}_0 \alpha_{i\ell m} \begin{cases} Y_{\ell m}^*(\pi/2, 0) & i = 1, \dots, 7, \\ \partial_\theta Y_{\ell m}^*(\pi/2, 0) & i = 8, 9, 10. \end{cases} \quad (148)$$

Here  $\mathcal{E}_0 = \mu f_0 U_0$ , with  $f_0 = 1 - 2M/r_0$  is the leading-order orbital energy, and the  $\alpha_{i\ell m}$ 's are given by (suppressing  $\ell m$  labels)

$$\alpha_1 = f_0^2/r_0, \quad \alpha_{2,5,9} = 0, \quad (149a)$$

$$\alpha_3 = f_0/r_0, \quad \alpha_4 = 2if_0 m \Omega_0, \quad (149b)$$

$$\alpha_6 = r_0 \Omega_0^2, \quad \alpha_7 = r_0 \Omega_0^2 [\ell(\ell + 1) - 2m^2], \quad (149c)$$

$$\alpha_8 = 2f_0 \Omega_0, \quad \alpha_{10} = 2imr_0 \Omega_0^2, \quad (149d)$$

with  $r_0 = r_0(\tilde{s})$ . For this source the  $i = 2, 5, 9$  equations are sourceless.

The subleading term (147) in the stress-energy is also straightforwardly calculated. However, rather than use it directly, we will incorporate it into the field equations through a redefinition of the puncture modes  $\tilde{h}_{i\ell m}^{2\mathcal{P}}$ . The two-timescale expansion of the puncture will be presented in the followup paper [72].

Finally, we perform the multiscale expansion of  $\delta^2 G_{i\ell m}$ ,

$$\delta^2 G_{i\ell m} = \delta^2 G_{i\ell m}^0(\varepsilon s, r) e^{-im\phi_p(s, \varepsilon)} + \mathcal{O}(\varepsilon). \quad (150)$$

The leading-order term is

$$\delta^2 G_{i\ell m}^0 = \sum_{\substack{i_1 \ell_1 m_1 \\ i_2 \ell_2 m_2}} \mathcal{G}_{i_1 \ell_1 m_1 i_2 \ell_2 m_2}^{(0)i\ell m} R_{i_1 \ell_1 m_1}^1 R_{i_2 \ell_2 m_2}^1, \quad (151)$$

where  $\mathcal{G}_{i_1 \ell_1 m_1 i_2 \ell_2 m_2}^{(0)i\ell m}$  is given by  $\mathcal{G}_{i_1 \ell_1 m_1 i_2 \ell_2 m_2}^{i\ell m}$  with derivatives that act on  $\tilde{h}_{i_1 \ell_1 m_1}^1$  replaced by  $\partial_t \rightarrow -i\omega_{m_1}(\tilde{s})$  and  $\partial_r \rightarrow \partial_r + i\omega_{m_1}(\tilde{s}) f^{-1} H$ , and analogously for derivatives of  $\tilde{h}_{i_2 \ell_2 m_2}^1$ .

With all the necessary expansions in hand, we now obtain the two-timescale form of the field equations. Substituting Eqs. (129), (132), (143), (145), and (150) into the field equations (128), moving subleading terms from the first-order equation into the second-order equation, and treating  $\tilde{s}$  and  $\phi_p$  as independent variables, we find the following equations for the mode amplitudes  $R_{i\ell m}^n$ :

$$E_{ij\ell m}^0 R_{j\ell m}^1 = -16\pi \tilde{t}_{i\ell m}^1 \delta(r - r_0), \quad (152)$$

$$\begin{aligned} E_{ij\ell m}^0 R_{j\ell m}^{2\mathcal{P}} &= 2\delta^2 G_{i\ell m}^0 - E_{ij\ell m}^0 R_{j\ell m}^{2\mathcal{P}} \\ &\quad - E_{ij\ell m}^1 R_{j\ell m}^1, \end{aligned} \quad (153)$$

where  $E_{ij\ell m}^n$  is given by Eq. (144),  $\tilde{t}_{i\ell m}^1$  by Eq. (148),  $\delta^2 G_{i\ell m}^0$  by Eq. (151), and as mentioned above, we have absorbed the effect of  $\tilde{T}_{i\ell m}^2$  into a redefinition of  $R_{i\ell m}^{2\mathcal{P}}$ .

#### D. Expanded gauge condition

To satisfy the Einstein equations, the solutions to the wave equations also have to satisfy the gauge condition  $Z_\mu := \nabla^\nu \tilde{h}_{\mu\nu} = 0$ .

We expand this condition in the same way we did the wave equation. After substituting the decomposition (133), we obtain an expansion (suppressing  $\ell, m$  labels)

$$\varepsilon Z_{kj}^0 R_j^1 + \varepsilon^2 (Z_{kj}^0 R_j^2 + Z_{kj}^1 R_j^1) = \mathcal{O}(\varepsilon^3), \quad (154)$$

$k = 1, 2, 3, 4$ , where

$$Z_{1j}^0 R_j = i\omega_m (R_1 + fR_3 + HR_2) + \frac{f}{r} (rR_{2,r} + R_2 - R_4), \quad (155a)$$

$$Z_{2j}^0 R_j = i\omega_m (R_2 + HR_1 - HfR_3) + f \left( R_{1,r} - fR_{3,r} + \frac{1}{r} [R_1 - R_5 - fR_3 - 2fR_6] \right), \quad (155b)$$

$$Z_{3j}^0 R_j = i\omega_m (R_4 + HR_5) + \frac{f}{r} [rR_{5,r} + 2R_5 + \ell(\ell+1)R_6 - R_7], \quad (155c)$$

$$Z_{4j}^0 R_j = i\omega_m (R_8 + HR_9) + \frac{f}{r} (rR_{9,r} + 2R_9 - R_{10}), \quad (155d)$$

and

$$Z_{1j}^1 R_j = -\tilde{\partial}_s (R_1 + fR_3 + HR_2), \quad (156a)$$

$$Z_{2j}^1 R_j = -\tilde{\partial}_s (R_2 + HR_1 - fHR_3), \quad (156b)$$

$$Z_{3j}^1 R_j = -\tilde{\partial}_s (R_4 + HR_5), \quad (156c)$$

$$Z_{4j}^1 R_j = -\tilde{\partial}_s (R_8 + HR_9). \quad (156d)$$

Hence, at first order, the conditions are

$$Z_{kj}^0 R_j^1 = 0. \quad (157)$$

At second order, the conditions are

$$Z_{kj}^0 R_j^2 = -Z_{kj}^1 R_j^1. \quad (158)$$

It is easily checked that Eqs. (157) are identical to the gauge constraints in [82] with  $\partial_t \rightarrow -i\omega_m$  and  $\partial_r \rightarrow \partial_r + i\omega_m f^{-1}H$ .

In Sec. VI we discuss how these gauge conditions combine with the field equations (152)–(153) and the equations of motion (102), (103), (104), and (106) to determine (i) the mode amplitudes at fixed values of slow time, and (ii) the slow evolution of the system.

### E. Hierarchical structure

The wave equations (152) and (153) each comprise 10 coupled ODEs for each value of  $\ell$  and  $m$ . However, several properties reduce the level of coupling.

First, as is true in any gauge in Schwarzschild space-time, the seven equations for the even-parity modes ( $i = 1, \dots, 7$ ) decouple from the three equations for the odd-parity modes ( $i = 8, 9, 10$ ): we have  $\mathcal{M}^{ij} = 0$  for any  $i = 1, \dots, 7$  with  $j = 8, 9, 10$ , and for any  $i = 8, 9, 10$  with

$j = 1, \dots, 7$ . (Note, however, that even- and odd-parity first-order modes do couple in the second-order source.)

Second, thanks to the gauge-damping terms added to the wave equations in Eqs. (113) and (114), the wave equations (152) and (153) partially decouple into a hierarchical structure. The equations for the  $i = 1, 3, 5, 6, 7$  modes decouple from the  $i = 2, 4$  modes, and the equations for the  $i = 9, 10$  modes decouple from the  $i = 8$  mode. Because of this partial decoupling, one can first solve for the  $i = 9, 10$  modes and then solve for the  $i = 8$  mode (in the odd sector), and first solve for the  $i = 1, 3, 5, 6, 7$  modes and then solve for the  $i = 2, 4$  modes (in the even sector).

Third, as again is always true, the field equations are overdetermined: we have 14 equations (10 wave equations and 4 gauge conditions) for our 10 variables. (For  $\ell = 0$  this reduces to 6 equations for 4 variables, and for  $\ell = 1$  it reduces to 10 equations for 8 variables.) This excess of equations allows us to use the gauge condition to solve for some modes, rather than using the wave equations. For example, if  $m \neq 0$ , one can use the gauge condition (157) or (158) to compute these modes algebraically from the other seven modes, specifically utilizing the form of (155b) to obtain the  $i = 2$  mode, (155d) to obtain the  $i = 8$  mode, and either (155a) or (155c) to obtain the  $i = 4$  mode.

If  $m = 0$ , the frequency  $\omega_m$  vanishes, making that procedure impossible. However, one can still utilize the gauge condition to further decouple the equations. For example, in the even sector, one can use the gauge conditions to solve for the  $i = 6, 7$  modes in terms of the  $i = 1, 3, 5$  modes, and the  $i = 4$  in terms of the  $i = 2$ . Substituting those results into the wave equations yields a coupled set of wave equations for the 1, 3, 5 modes. The solutions for the  $i = 4, 6, 7$  modes can then be substituted into the wave equation for the  $i = 2$  mode.

For  $\ell = 0$  or 1, some modes do not appear, but the decoupling procedure is otherwise the same.

There are also two additional properties that reduce the number of modes that need to be calculated. Because  $\bar{h}_{\alpha\beta}$  is real, the amplitudes satisfy  $R_{i\ell, -m}^n = (-1)^m R_{i\ell m}^{n*}$ , meaning we need only calculate modes with  $m \geq 0$ . This property follows from the identity  $Y_{\alpha\beta}^{i\ell m*} = (-1)^m Y_{\alpha\beta}^{i\ell, -m}$ . Because the system is symmetric under reflection through the equatorial plane ( $\mathcal{R} : \theta \mapsto \pi - \theta$ ), we also have that the even-parity modes vanish for odd values of  $\ell + m$ , and that odd-parity modes vanish for even values of  $\ell + m$ . This follows from the property  $\mathcal{R} Y_{\alpha\beta}^{i\ell m} = -(-1)^{\ell+m} Y_{\alpha\beta}^{i\ell m}$  for even-parity harmonics and  $\mathcal{R} Y_{\alpha\beta}^{i\ell m} = (-1)^{\ell+m} Y_{\alpha\beta}^{i\ell m}$  for odd-parity harmonics.

Putting all of these properties together, in Table I we summarize how each mode can be calculated.

Once one has obtained all the modes, as a consistency check one can verify that they satisfy whichever wave equations or gauge conditions were not used in obtaining them. Since the wave equations and gauge conditions are only consistent with each other if the source in the field

| Sector   | Coupled ODEs        | Gauge condition | Hierarchical ODEs |
|--|---------------------|-----------------|-------------------|
| $\ell \geq 2, \ell + m \text{ odd}, m \neq 0$  | $i = 9, 10$         | $i = 8$         |                   |
| $\ell \geq 2, \ell + m \text{ even}, m \neq 0$ | $i = 1, 3, 5, 6, 7$ | $i = 2, 4$      |                   |
| $\ell = 1, m = \pm 1$                          | $i = 1, 3, 5, 6$    | $i = 2, 4$      |                   |
| $\ell \geq 2 \text{ odd}, m = 0$               | $i = 9, 10$         |                 | $i = 8$           |
| $\ell \geq 2 \text{ even}, m = 0$              | $i = 1, 3, 5$       | $i = 4, 6, 7$   | $i = 2$           |
| $\ell = 1, m = 0$                              | $i = 9$             |                 | $i = 8$           |
| $\ell = 0$                                     | $i = 1, 3$          | $i = 6$         | $i = 2$           |

TABLE I. Summary of how the solutions for nonvanishing modes are obtained. The first column displays the different sectors into which the modes are divided. The second column lists the modes  $R_{\ell m}^n$  that are calculated by directly solving a coupled subset of the wave equations (152) or (153). The third column lists the modes that can be extracted algebraically from the coupled modes using the gauge condition (157) or (158). The fourth column lists the modes that must be extracted from the coupled modes using the remaining, hierarchically decoupled subset of the wave equations (152) or (153). All unlisted modes identically vanish.

equations is conserved, this consistency check serves as a check on the source.

## F. Transformations between slicings

If  $|r^*| \gtrsim M/\varepsilon$ , then  $t$  and  $s$  can differ by a large amount, of order  $\gtrsim 1/\varepsilon$ . Hence, at large distances, the two-timescale expansion of the metric in one slicing can differ significantly from the two-timescale expansion of it in another slicing. This is related to the fact that the two-timescale expansion breaks down at large distances, as described in Ref. [35]. This breakdown is minor in hyperboloidal slicing but dramatic in  $t$  slicing.

However, within the region  $|r^*| \lesssim M/\varepsilon$ , the choice of slicing should be free. Given the results in one slicing, we should be able to recover the fields in another. In fact, we can straightforwardly determine the relationship between them. To eliminate ambiguities in the resulting equations, we introduce labels in square brackets, such as  $[t]$  and  $[s]$ , to denote quantities defined with respect to a given a slicing. For example,  $R_{\ell m}^{[s]n}(\tilde{s}, r)$  will denote a coefficient in the two-timescale expansion based on  $s$  slicing.  $R_{\ell m}^{[s]n}(\tilde{t}, r)$  denotes that same function evaluated at  $\tilde{t}$ ; this differs from  $R_{\ell m}^{[t]n}(\tilde{t}, r)$  because if  $k \neq 0$  then  $R_{\ell m}^{[s]n}$  and  $R_{\ell m}^{[t]n}$  are different functions of  $r$  (and for  $n > 1$ , also different functions of their first argument, as we see below).

The transformations between slicings follow from  $\tilde{s} = \tilde{t} - \varepsilon k(r^*)$  and  $\phi_p(s, \varepsilon) = \phi_p(t, \varepsilon) - \Omega(\tilde{t}, \varepsilon)k(r^*) + \frac{1}{2}\varepsilon\dot{\Omega}(\tilde{t})k^2(r^*) + \mathcal{O}(\varepsilon^2)$ . Using these expansions to re-expand  $R_{\ell m}^n(\tilde{s}, r)e^{-im\phi_p(s, \varepsilon)}$  at fixed  $\tilde{t}$  and  $\phi_p(t, \varepsilon)$ , we

find

$$R_{\ell m}^{[s]n}e^{-im\phi_p(s, \varepsilon)} = e^{im\Omega_0(\tilde{t})k} \left\{ 1 - \varepsilon k \left[ \tilde{\partial}_{\tilde{t}} + \frac{1}{2}im\dot{\Omega}_0(\tilde{t})k \right] + \mathcal{O}(\varepsilon^2) \right\} R_{\ell m}^{[s]n}(\tilde{t}, r)e^{-im\phi_p(t, \varepsilon)}. \quad (159)$$

For a sufficiently well-behaved metric, the reexpansion of  $\sum_n \varepsilon^n R_{\ell m}^{[s]n}(\tilde{s}, \varepsilon)e^{-im\phi_p(s, \varepsilon)}$  must agree with  $\sum_n \varepsilon^n R_{\ell m}^{[t]n}(\tilde{t}, \varepsilon)e^{-im\phi_p(t, \varepsilon)}$  at each order in  $\varepsilon$  at fixed  $\tilde{t}$  and  $\phi_p(t, \varepsilon)$ . This equality gives us

$$R_{\ell m}^{[t]1} = e^{im\Omega_0(\tilde{t})k(r^*)} R_{\ell m}^{[s]1}(\tilde{t}, r), \quad (160)$$

$$R_{\ell m}^{[t]2} = e^{im\Omega_0(\tilde{t})k(r^*)} \left\{ R_{\ell m}^{[s]2}(\tilde{t}, r) - \varepsilon k(r^*) \left[ \tilde{\partial}_{\tilde{t}} + \frac{1}{2}im\dot{\Omega}_0(\tilde{t})k(r^*) \right] R_{\ell m}^{[s]1}(\tilde{t}, r) \right\}. \quad (161)$$

If we first calculate the metric perturbations in hyperboloidal slicing, these relationships allow us to transform those results into  $t$  slicing. The inverse relationships are also easily derived.

These transformations also affect the second-order source terms  $\delta^2 G_{\ell m}^0$  and  $E_{ij\ell m}^1 R_{j\ell m}^1$ . In particular, they alter the asymptotic behavior of  $E_{ij\ell m}^1 R_{j\ell m}^1$ . We will return to this issue in Sec. V H.

## G. Boundary conditions

For any given boundary conditions we can solve the frequency-domain field equations (152)–(153) using the method of variation of parameters, as done at first order in Refs. [34, 60]. We will provide additional details about the application of this method at second order, with hyperboloidal slicing, in Ref. [70].

The boundary conditions themselves are also seemingly obvious: our solutions should satisfy retarded boundary conditions, with no radiation coming into the system from infinity and none coming out of the black hole's past horizon. However, determining the form of these boundary conditions in the two-timescale expansion is nontrivial because, as alluded to above, the two-timescale expansion breaks down at large  $r$  and near the horizon. Here we sketch the form of the boundary conditions for (i) homogeneous solutions, which are required to obtain inhomogeneous solutions in the method of variation of parameters, (ii) first-order inhomogeneous solutions, and (iii) second-order inhomogeneous solutions.

For homogeneous solutions, we construct a complete basis. For nonstationary ( $m \neq 0$ ) modes, half of the basis solutions represent regular, purely outgoing waves at infinity, behaving like  $\sim e^{-i\omega_m u}$  as  $r \rightarrow \infty$ ; the other half represent regular, purely ingoing waves at the horizon, behaving like  $\sim e^{-i\omega_m v}$  as  $r \rightarrow 2M$ . For stationary ( $m = 0$ ) modes, half of the basis solutions are regular at infinity, and half are regular at the future horizon.

Here by regularity at the horizon we mean regularity of each component in ingoing Eddington-Finkelstein coordinates  $(v, r, \theta^A)$ . By expressing the Eddington-Finkelstein components of a generic (symmetric) tensor  $A_{\mu\nu}$  in terms of its tensor-harmonic coefficients  $A_{i\ell m}$ , we find the following:  $A_{\mu\nu}$  is smooth at the future horizon if and only if

$$\text{all } A_{i\ell m}(v, r) \text{ are smooth at } r = 2M, \quad (162a)$$

$$A_{2\ell m}(v, r) = A_{1\ell m}(v, r) + \mathcal{O}(f^2), \quad (162b)$$

$$A_{i\ell m}(v, r) = A_{i+1, \ell m}(v, r) + \mathcal{O}(f) \quad \text{for } i = 4, 8. \quad (162c)$$

Here the mode amplitudes are defined without overall factors, by  $A_{i\ell m} := \frac{1}{\kappa_i} \oint d\Omega \eta^{\alpha\mu} \eta^{\beta\nu} A_{\alpha\beta} Y_{\mu\nu}^{i\ell m*}$ ; this implies that the conditions (162) are slightly modified for the source modes defined by Eq. (118), for example. Since the conditions describe the behavior of the mode coefficients as functions of  $v$  and  $r$ , not as functions of  $t$ , and  $r$ , they also do not quite apply for  $R_{i\ell m}^{[t]n}$ , but they do apply to the combined quantities  $R_{i\ell m}^{[t]n} e^{-i\omega_m t} = \left( R_{i\ell m}^{[t]n} e^{i\omega_m r^*} \right) e^{-i\omega_m v}$ . Given smooth mode coefficients, the Lorenz-gauge field equations tend to enforce the remaining two conditions in (162).

For each  $\ell m$  mode there are  $2d$  independent homogeneous solutions, where  $d$  is the number of coupled equations to be solved for the given  $\ell m$  mode, as per the second column in Table I. We write these basis solutions as  $R_{i\ell m}^{k-}$  and  $R_{i\ell m}^{k+}$ , with  $k = 1, \dots, d$ . For the nonstationary modes in  $[t]$  slicing, these satisfy the boundary conditions

$$R_{i\ell m}^{[t]k+}(r) = e^{i\omega_m r^*} \sum_{n=0}^{\infty} a_{i\ell m}^{k,n} r^{-n}, \quad (163)$$

$$R_{i\ell m}^{[t]k-}(r) = e^{-i\omega_m r^*} \sum_{n=0}^{\infty} b_{i\ell m}^{k,n} (r - 2M)^n, \quad (164)$$

for any value of  $r$  near infinity or near  $r = 2M$ , respectively. Here both  $\omega_m$  and the coefficients  $a_{i\ell m}^{k,n}$  and  $b_{i\ell m}^{k,n}$  implicitly depend on slow time. For the stationary modes ( $m = 0$ ), the solutions satisfy

$$R_{i\ell 0}^{[t]k+}(r) = \sum_{n=\ell}^{\infty} \left( a_{i\ell m}^{k,n} + \bar{a}_{i\ell m}^{k,n} \ln r \right) r^{-n}, \quad (165)$$

$$R_{i\ell 0}^{[t]k-}(r) = \sum_{n=0}^{\infty} b_{i\ell m}^{k,n} (r - 2M)^n, \quad (166)$$

near the respective boundaries; in this case because the field equations do not depend on  $\Omega_0$ , the coefficients  $a_{i\ell m}^{k,n}$ ,  $\bar{a}_{i\ell m}^{k,n}$ , and  $b_{i\ell m}^{k,n}$  do not depend on slow time. In both cases the coefficients are different for each  $i\ell m$  and are determined from recurrence relations derived by substituting the ansatzes into the field equations. These can be found in Appendix A of Ref. [68]. That reference also describes

how to choose the leading coefficients,  $a_{i\ell m}^{k,0}$ ,  $\bar{a}_{i\ell m}^{k,0}$ , and  $b_{i\ell m}^{k,0}$ , to generate the  $2d$  independent solutions. For the stationary modes, the homogeneous solutions are known analytically [61]. For the nonstationary modes, the homogeneous solutions can be obtained numerically, enforcing the boundary condition (163) or (164) exactly at some radius  $r_{\text{in}}$  near the horizon or  $r_{\text{out}}$  near infinity.

The boundary conditions on  $s$  slices are easily obtained from those on  $t$  slices using the transformation (160). That transformation implies  $R_{i\ell m}^{[s]k+} = R_{i\ell m}^{[t]k+} e^{-i\omega_m (i)k(r^*)}$  for  $m \neq 0$ , and  $R_{i\ell m}^{[s]k+} = R_{i\ell m}^{[t]k+}$  for  $m = 0$ . If the boundaries are placed sufficiently near  $r = 2M$  or  $r \rightarrow \infty$ , where  $s$  is sufficiently close to  $v$  or  $u$ , then for the  $m \neq 0$  modes this amounts to using the same boundary conditions as in (163) and (164), but without the plane-wave factor  $e^{\pm i\omega_m r^*}$  in front of the sum. In that case, all modes, nonstationary and stationary, behave as simple power series near the boundaries (modulo the logarithms at large  $r$  in the stationary modes).

Now moving onto the first-order inhomogeneous solutions, we note that these solutions are homogeneous at all points away from  $r_0$ . Hence, at the future horizon they should be a linear combination of the basis solutions  $R_{i\ell m}^{k-}$ , and at large  $r$  they should be a linear combination of the basis solutions  $R_{i\ell m}^{k+}$ . The analysis in Ref. [35] showed that this is true despite the breakdown of the two-timescale expansion at large  $r$ , and a similar analysis, to be presented elsewhere, shows the same at the horizon. In other words, at first order the correct boundary conditions in the two-timescale expansion are the standard ones imposed in the literature on the first-order solution in the frequency domain.

At second order, the situation is more complicated. The sources are now non-compact, extending all the way to the boundaries. This means that the solutions are not homogeneous anywhere. One might guess that integrating the source against a frequency-domain retarded Green's function would still yield the correct result, but Ref. [35] showed that this is not the case. For some modes, the retarded integral fails to converge. Even when it does converge, one can show that it does not necessarily give the physically correct result; the companion paper [70] derives the necessary and sufficient conditions under which it does give the correct solution.

Ref. [35] showed, in a scalar toy model, how to obtain physical boundary conditions by using a retarded time-domain post-Minkowski solution at large  $r$ . This solution is obtained analytically, using a time-domain retarded Green's function in the asymptotic region, up to some desired post-Minkowskian order. It is then re-expanded in the two-timescale form to determine the physical boundary conditions for the two-timescale solution. A similar iterative, analytical, time-domain method can be used to generate physical boundary conditions near the future horizon. These methods were used in the calculation in Ref. [39]. Their full details will be presented in followup papers [64] [65].



Near future null infinity and the future horizon, the analytical time-domain solutions take the form

$$\bar{h}_{ilm}^{1\pm}(s, r, \varepsilon) = \bar{h}_{ilm}^{1H\pm}(s, r, \varepsilon), \quad (167)$$

$$\bar{h}_{ilm}^{2\pm}(s, r, \varepsilon) = \bar{h}_{ilm}^{2P\pm}(s, r, \varepsilon) + \bar{h}_{ilm}^{2H\pm}(s, r, \varepsilon), \quad (168)$$

where ‘+’ is the solution near future null infinity and ‘−’ is the solution near the future horizon. The terms  $\bar{h}_{ilm}^{nH\pm}$  are homogeneous solutions satisfying retarded boundary conditions:  $\bar{h}_{ilm}^{nH+}$  is an outgoing, regular wave at infinity, and  $\bar{h}_{ilm}^{nH-}$  is an ingoing, regular wave at the future horizon. These homogeneous solutions are not determined internally within the analytical time-domain calculations; instead they are determined by matching the time-domain solution to the two-timescale solution.

$\bar{h}_{ilm}^{2P\pm}$  is an analytically known particular solution to the second-order field equations (35), fully determined by the first-order modes  $\bar{h}_{ilm}^{1H\pm}$ .  $\bar{h}_{ilm}^{2P+}$  may or may not be regular at future null infinity, and  $\bar{h}_{ilm}^{2P-}$  may or may not be regular at the future horizon, but they are (by construction) physical, causal solutions. Any irregularity arises from the choice of gauge. For example, outgoing waves at nonlinear orders in the Lorenz gauge contain logarithms at large  $r$ , as is well known from post-Minkowski theory in the harmonic gauge [83, 84]. As another example, at second order every single mode is impacted (through the source  $\delta^2 G_{ilm}$ ) by the known pathologies of the first-order  $\ell = 0$  mode in the Lorenz gauge, reviewed in Appendix D.

To match the time-domain solutions at the boundaries to the two-timescale solution in the bulk of the space-time, Eqs. (167)–(168) are re-expanded for  $\varepsilon \ll 1$  at fixed  $(\tilde{s}, \phi_p(s, \varepsilon), r)$ , as in Eq. (129). This re-expansion can be highly nontrivial due to the complexity of the time-domain solution (as described in Ref. [35]), but once it is complete, the results must agree term by term with the two-timescale expansion (133). This provides the matching conditions

$$R_{ilm}^1 e^{-im\phi_p} = \tilde{h}_{ilm}^{1H\pm}(\tilde{s}, \phi_p, r, 0), \quad (169)$$

$$R_{ilm}^2 e^{-im\phi_p} = \tilde{h}_{ilm}^{2P\pm}(\tilde{s}, \phi_p, r, 0) + \partial_\varepsilon \tilde{h}_{ilm}^{1H\pm}(\tilde{s}, \phi_p, r, 0) + \tilde{h}_{ilm}^{2H\pm}(\tilde{s}, \phi_p, r, 0), \quad (170)$$

where the ‘+’ refers to the matching condition at large  $r$ , and the ‘−’ to the matching condition near the horizon. Equation (169) is used to match ‘outward’: the output of the leading-order two-timescale solution,  $R_{ilm}^1$ , fixes the homogeneous solutions  $\bar{h}_{ilm}^{1H\pm}$ . That in turn fixes the particular solutions  $\bar{h}_{ilm}^{2P\pm}$ . Equation (170) is then used to match ‘inward’, providing boundary conditions for the second-order two-timescale solution  $R_{ilm}^2$ . To enforce those conditions, one can treat the first two terms on the right as a puncture  $R_{ilm}^{2P\pm}$ , which is utilized in the same manner as the puncture on  $z^\mu$ . The residual field  $R_{ilm}^{2\mathcal{R}\pm} = R_{ilm}^2 - R_{ilm}^{2P\pm}$  must then be the homogeneous solution  $\tilde{h}_{ilm}^{2H\pm}(\tilde{s}, \phi_p, r, 0)$ . This implies that, just like the

first-order field  $R_{ilm}^1$ ,  $R_{ilm}^{2\mathcal{R}\pm}$  must reduce to a linear combination of the basis solutions  $R_{ilm}^{k+}$  at large  $r$  and of the basis solutions  $R_{ilm}^{k-}$  near  $r = 2M$ . Numerically solving for the residual field then fixes the homogeneous solutions  $\bar{h}_{ilm}^{2H\pm}$ , fully determining the full, physical fields (168) at the future horizon and future null infinity.

## H. Asymptotics of the second-order source

Finally, we consider the behavior of the second-order source at large  $r$  and near the horizon. Our analysis will verify the expectation that the source is significantly better behaved with hyperboloidal slicing.

We focus our attention on the term  $E_{ij\ell m}^1 R_{ilm}^1$  in the source; we will briefly discuss  $\delta^2 G_{ilm}^0$  at the end of the section.

First we look at the large- $r$  behavior. In  $t$  slicing, the first-order field behaves as  $R_{ilm}^{[t]1} \sim e^{i\omega_m(t)r^*}$ , from Eq. (163). When a slow time derivative acts on this field, it yields  $\partial_t R_{ilm}^{[t]1} \sim i\dot{\omega}_m r^* e^{i\omega_m r^*}$ . According to Eqs. (141) and (C2), the terms in  $E_{ij\ell m}^1 R_{ilm}^1$  then behave as

$$\square_1 R_{ilm}^{[t]1} = \left[ \frac{\omega_m \dot{\omega}_m}{2} r^* + \mathcal{O}(r^0) \right] R_{ilm}^{[t]1}, \quad (171)$$

$$\mathcal{M}_1^{ij} R_{j\ell m}^{[t]1} = \left[ \frac{i\dot{\omega}_m r^*}{r^2} + \mathcal{O}(r^{-2}) \right] C^{ij} R_{j\ell m}^{[t]1}, \quad (172)$$

where  $C^{ij}$  is an  $r$ -independent coupling matrix. Therefore

$$E_{ij\ell m}^1 R_{j\ell m}^{[t]1} = \left[ \frac{\omega_m \dot{\omega}_m}{2} r^* + \mathcal{O}(r^0) \right] R_{ilm}^{[t]1}, \quad (173a)$$

$$\sim r^* e^{i\omega_m r^*}. \quad (173b)$$

We provide a further discussion of this behavior toward the end of the section.

Next consider the large- $r$  behavior in  $s$  slicing. Assume  $r$  is sufficiently large that  $s = u$ . From the discussion in the previous section,  $R_{ilm}^{[u]1}$  is a simple power series in  $1/r$  at large  $r$ , with coefficients that depend on  $\tilde{u}$ . According to Eqs. (141) and (C2), the terms in  $E_{ij\ell m}^1 R_{ilm}^1$  then behave as

$$\square_1 R_{ilm}^{[u]1} \sim \frac{1}{r^2} R_{ilm}^{[u]1}, \quad (174)$$

$$\mathcal{M}_1^{ij} R_{j\ell m}^{[u]1} \sim \frac{1}{r^2} C^{ij} R_{j\ell m}^{[u]1}, \quad (175)$$

and so

$$E_{ij\ell m}^1 R_{j\ell m}^{[u]1} \sim \frac{1}{r^2}. \quad (176a)$$

This is an improvement over Eq. (173b) by *three orders* in  $1/r$ .

Now we look at the behavior near the horizon. In  $t$  slicing, the first-order field behaves as  $R_{ilm}^{[t]1} \sim e^{-i\omega_m(t)r^*}$ ,

from Eq. (164). When a slow time derivative acts on this field, it yields  $\partial_t R_{i\ell m}^{[t]1} \sim i\dot{\omega}_m r^* e^{-i\omega_m r^*}$ . According to Eqs. (141) and (C2), the terms in  $E_{ij\ell m}^1 R_{i\ell m}^1$  then behave as

$$\square_1 R_{i\ell m}^{[t]1} = \left[ -\frac{\omega_m \dot{\omega}_m}{2} r^* + \mathcal{O}(f^0) \right] R_{i\ell m}^{[t]1}, \quad (177)$$

$$\mathcal{M}_1^{ij} R_{j\ell m}^{[t]1} = \left[ -\frac{i\dot{\omega}_m r^*}{r^2} + \mathcal{O}(f^0) \right] C^{ij} R_{j\ell m}^{[t]1}, \quad (178)$$

and so

$$E_{ij\ell m}^1 R_{j\ell m}^{[t]1} = \left[ -\frac{\omega_m \dot{\omega}_m}{2} r^* + \mathcal{O}(f^0) \right] D^{ij} R_{j\ell m}^{[t]1}, \quad (179a)$$

$$\sim r^* e^{-i\omega_m r^*}. \quad (179b)$$

where we have used  $f = (r - 2M)/r$  to count orders, and  $D^{ij}$  is an  $\mathcal{O}(f^0)$  coupling matrix.

Next consider the near-horizon behavior in  $s$  slicing. Assume  $r$  is sufficiently near  $2M$  that  $s = v$ . From the discussion in the previous section,  $R_{i\ell m}^{[v]1}$  is a simple power series in  $(r - 2M)$  near the horizon, with coefficients that depend on  $\tilde{v}$ . According to Eqs. (141) and (C2), the terms in  $E_{ij\ell m}^1 R_{i\ell m}^1$  then behave as

$$\square_1 R_{i\ell m}^{[v]1} \sim f, \quad (180)$$

$$\mathcal{M}_1^{ij} R_{j\ell m}^{[v]1} \sim f. \quad (181)$$

In the first equation, we have used  $\partial_{r^*} = f \partial_r$ . In the second, we have used the regularity conditions (162). Hence,

$$E_{ij\ell m}^1 R_{j\ell m}^{[v]1} \sim f. \quad (182)$$

This is an improvement over Eq. (179b) from a logarithmic divergence to something that vanishes at the horizon.

What is the physical cause of these wildly differing behaviors between the different slicings? In  $t$  slicing, our choice of slow and fast time has left rapid oscillations in  $r$  in the wave zones. Because the frequencies of those oscillations vary with slow time, varying the slow time a small amount can actually change the field by a large amount, leading to slow time derivatives having large effects. On the other hand, in  $s$  slicing, our choice of slow and fast time in this case has correctly factored out rapid oscillations not only in  $t$  at fixed  $r$ , but also in  $r$  at fixed  $t$ . The amplitudes  $R_{i\ell m}^{[s]1}$  are hence genuinely slowly varying functions in spacetime.

On the other hand, different slicings have only a trivial effect on  $\delta^2 G_{i\ell m}^0$ . The reason is that this source term contains no slow time derivatives. As a consequence,  $\delta^2 G_{i\ell m}^0$  transforms in the same way as  $R_{i\ell m}^1$ :

$$\delta^2 G_{i\ell m}^{[t]0} = \delta^2 G_{i\ell m}^{[s]0}(\tilde{t}, r) e^{im\Omega_0(\tilde{t})k(r^*)}. \quad (183)$$

To assess the behavior at large  $r$ , we can use the facts that (i)  $h_{\alpha\beta}^1 \sim \frac{e^{-im\phi_P(u, \varepsilon)}}{r}$ , (ii)  $\delta^2 G_{\alpha\beta} \sim \partial h^1 \partial h^1 + h^1 \partial^2 h^1$ , and (iii) for first-order modes with mode numbers  $m_1$  and

$m_2$ , the mode number of  $\delta^2 G_{i\ell m}$  is  $m = m_1 + m_2$ . From these facts, we see that at large  $r$ ,

$$\delta^2 G_{i\ell m}^{[t]0} \sim (rf) \frac{e^{i\omega_m r^*}}{r^2} \quad (184)$$

and

$$\delta^2 G_{i\ell m}^{[s]0} \sim (rf) \frac{1}{r^2}. \quad (185)$$

The factor  $(rf)$  arises from the overall factor introduced in the definition (118); the behaviors  $\frac{e^{i\omega_m r^*}}{r^2}$  and  $1/r^2$  are the ‘natural’ behaviors in the two slicings. These scalings hold even for  $m = 0$  modes, due to the beating of first-order waves against each other (i.e., the combination of modes with  $m_2 = -m_1$ ).

At the horizon, we can appeal to the fact that for a horizon-regular first-order metric perturbation,  $\delta^2 G_{\alpha\beta}$  is necessarily horizon-regular as well. The regularity conditions (162) then apply, up to the overall factor of  $(rf)$  in the definition (118). So

$$\delta^2 G_{i\ell m}^{[t]0} \sim (rf) e^{-i\omega_m r^*} \quad (186)$$

and

$$\delta^2 G_{i\ell m}^{[s]0} \sim (rf). \quad (187)$$

We summarize with two points. First, hyperboloidal slicing dramatically increases the falloff of the source term  $E_{ij\ell m}^1 R_{j\ell m}^1$  near the boundaries. This lessens the numerical burden of solving the second-order field equation and simplifies the task of finding physical boundary conditions. We will show this in more detail in Ref. [70]. Our second point is that hyperboloidal slicing has no significant impact on the nonlinear source  $\delta^2 G_{i\ell m}^0$ . For this source term, we must derive boundary conditions as described in the previous section, and as derived explicitly in the scalar model of Ref. [35]. This was the procedure used to compute the second-order binding energy in Ref. [39].

## VI. STRUCTURE OF THE SOLUTION: COMBINED EVOLUTION OF THE FIELD AND THE TRAJECTORY

Collecting the expanded equation of motion and the expanded field equations, we end up with the coupled set (102), (103), (104), (106), (152), (153), (157), and (158). Here we outline in some detail how they combine to provide the adiabatic and post-adiabatic waveform-generation schemes displayed in Fig. 3.

### A. Adiabatic order

At first order, the solution is fully determined by the values of  $r_0$ ,  $M_1$ , and  $S_1$  (together with the boundary

conditions and the relation  $\Omega_0 = \sqrt{M/r_0^3}$ . However,  $M_1$  and  $S_1$  are not relevant for the evolution equations at this order, and unless one is interested in post-adiabatic evolution, one can freely set them to zero. In this section we outline how this structure emerges.

For  $\ell \geq 2$ , and for  $\ell = 1, m = \pm 1$ , we use the same solutions as in Refs. [34, 68]. These solutions are uniquely determined by the source (146) and the retarded boundary conditions described above. Hence, once boundary conditions are selected, each of these mode amplitudes is fully determined by the value of  $r_0$ . Following the notation in the introduction, we write the  $\ell \geq 2$  and  $\ell = 1, m = \pm 1$  solutions as

$$R_{i\ell m}^1(\tilde{s}, r) = R_{i\ell m}^{\text{PP}}(r_0(\tilde{s}), r). \quad (188)$$

However, for  $\ell = 0$  and for  $\ell = 1, m = 0$ , our solutions differ from the traditional ones. For these modes, the solution is *not* uniquely determined by regularity conditions: to any linear perturbation of Schwarzschild space-time, one can always freely add a linear perturbation toward a Kerr spacetime with mass  $M + \varepsilon M_1$  and spin  $\varepsilon S_1$ . In standard linear perturbation theory, the corrections  $M_1$  and  $S_1$  must be constants; if they were not, they would violate the linearized Einstein equation. But in the two-timescale expansion, at first order they can have *arbitrary* dependence on slow time. This is because regardless of their slow time dependence, they will still satisfy the leading-order field equations, in which no slow-time derivatives appear. Their dependence on  $\tilde{s}$  is only determined by the *second*-order field equations. Hence our total first-order solution for these modes is

$$R_{i00}^1(\tilde{s}, r) = R_{i00}^{\text{PP}}(r_0(\tilde{s}), r) + \bar{x}_{i00}(M_1(\tilde{s}), r), \quad (189)$$

$$R_{i10}^1(\tilde{s}, r) = R_{i10}^{\text{PP}}(r_0(\tilde{s}), r) + \bar{x}_{i10}(S_1(\tilde{s}), r), \quad (190)$$

for some yet-to-be-determined functions  $M_1(\tilde{s})$  and  $S_1(\tilde{s})$ . We present  $R_{i00}^{\text{PP}}$ ,  $R_{i10}^{\text{PP}}$ ,  $\bar{x}_{i00}$ , and  $\bar{x}_{i10}$  explicitly in Appendix D.

The mode amplitudes at a given value of slow time feed into the evolution equations that drive the system to future slow times. To evolve the system, we require evolution equations for  $r_0$ ,  $M_1$ , and  $S_1$ . The first of these is given by Eq. (103), for which we require the two-timescale expansion of the self-force.

To assist in writing this expansion, we first write the two-timescale expansions of  $z^\mu$ ,  $\dot{z}^\mu = dz^\mu/dt$ ,  $u^\mu$ , and  $h_{\alpha\beta}^{\mathcal{R}}$ :

$$z^\alpha = z_0^\alpha(t, \tilde{t}, \phi_p) + \varepsilon z_1^\alpha(\tilde{t}, \phi_p) + \mathcal{O}(\varepsilon^2), \quad (191)$$

$$\dot{z}^\alpha = \dot{z}_0^\alpha(\tilde{t}, \phi_p) + \varepsilon \dot{z}_1^\alpha(\tilde{t}, \phi_p) + \mathcal{O}(\varepsilon^2), \quad (192)$$

$$u^\alpha = u_0^\alpha(\tilde{t}, \phi_p) + \varepsilon u_1^\alpha(\tilde{t}, \phi_p) + \mathcal{O}(\varepsilon^2), \quad (193)$$

$$h_{\alpha\beta}^{\mathcal{R}} = \sum_n \varepsilon^n \tilde{h}_{\alpha\beta}^{n\mathcal{R}}(\tilde{s}, \phi_p, r, \theta^A), \quad (194)$$

where

$$z_0^\alpha = (t, r_0, \pi/2, \phi_p), \quad (195)$$

$$z_1^\alpha = (0, r_1, 0, 0), \quad (196)$$

$$\dot{z}_0^\alpha = (1, 0, 0, \Omega_0), \quad (197)$$

$$\dot{z}_1^\alpha = (0, dr_0/d\tilde{t}, 0, \Omega_1), \quad (198)$$

$$u_0^\alpha = U_0 \dot{z}_0^\alpha, \quad (199)$$

$$u_1^\alpha = U_1 \dot{z}_0^\alpha + U_0 \dot{z}_1^\alpha. \quad (200)$$

The quantity  $\tilde{h}_{\alpha\beta}^{n\mathcal{R}}$  here is given by the trace reversal of the coefficient of  $\varepsilon^n$  in Eq. (133) (or more precisely, the analogue of that equation for the residual field), summed over  $i\ell m$ . It can be calculated from the solutions to Eqs. (152)–(153) as

$$\tilde{h}_{\alpha\beta}^{n\mathcal{R}} = \sum_{i\ell m} \frac{a_{i\ell}}{r} R_{i\ell m}^{n\mathcal{R}}(\tilde{s}, r) e^{im[\phi - \phi_p(s, \varepsilon)]} Y_{\alpha\beta}^{i\ell m}(r, \theta, 0), \quad (201)$$

where the trace reversal is achieved with

$$\underline{i} = \begin{cases} 6 & \text{if } i = 3, \\ 3 & \text{if } i = 6, \\ i & \text{otherwise.} \end{cases} \quad (202)$$

The amplitudes  $R_{i\ell m}^{1\mathcal{R}}$  can be computed from the solution to Eq. (152) as  $R_{i\ell m}^{1\mathcal{R}} - R_{i\ell m}^{1\mathcal{P}}$ , or directly, using a puncture scheme. In Eq. (201), we have pointedly pulled out the factor  $e^{im\phi}$  from  $Y_{\alpha\beta}^{i\ell m}$ ; this makes manifest that  $\tilde{h}_{\alpha\beta}^{n\mathcal{R}}$  only depends on  $\phi_p$  and  $\phi$  in the combination  $\phi - \phi_p$ . When evaluated on  $z_0^\alpha$ , where  $\phi = \phi_p$ , the residual field hence becomes independent of fast time. The same statement naturally applies to all derivatives of the residual field, which confirms that the self-force is independent of  $\phi_p$ , as assumed in the expansion (99).

Substituting the above expansions into the self-force (42), we find that the first- and second-order terms in the expansion (99) are given by

$$\tilde{f}_1^\alpha(\tilde{t}) = \frac{1}{2} g^{\alpha\beta} \tilde{h}_{u_0 u_0, \beta}^{1\mathcal{R}}, \quad (203a)$$

$$\begin{aligned} \tilde{f}_2^\alpha(\tilde{t}) = & \frac{1}{2} g^{\alpha\beta} \tilde{h}_{u_0 u_0, \beta}^{2\mathcal{R}} + \frac{1}{2} \left[ r_1 (\partial_r g^{\alpha\beta} \tilde{h}_{u_0 u_0, \beta}^{1\mathcal{R}} + g^{\alpha\beta} \tilde{h}_{u_0 u_0, r\beta}^{1\mathcal{R}} \right. \\ & + 2g^{\alpha\beta} \tilde{h}_{u_0 u_1, \beta}^{1\mathcal{R}} + P_0^{\alpha\beta} (2\Gamma^\gamma \tilde{h}_{\beta\gamma}^{1\mathcal{R}} - 2U_0 u_0^\gamma \partial_{\tilde{t}} \tilde{h}_{\beta\gamma}^{1\mathcal{R}} \\ & \left. - \tilde{h}_{\beta}^{1\mathcal{R}} \gamma \tilde{h}_{u_0 u_0, \gamma}^{1\mathcal{R}}) \right]. \end{aligned} \quad (203b)$$

Here a comma denotes a derivative at fixed slow time, all fields are evaluated on  $z_0^\alpha$ ,  $\tilde{h}_{u_0 u_0, \beta \dots \gamma}^{n\mathcal{R}} := \tilde{h}_{\mu\nu, \beta \dots \gamma}^{n\mathcal{R}}|_{z_0^\alpha} u_0^\mu u_0^\nu$ ,  $P_0^{\alpha\beta} := g^{\alpha\beta} + u_0^\alpha u_0^\beta$ , and  $\Gamma^\alpha := U_0^2 (2\Gamma_{\dot{z}_0}^\alpha - 3\Omega_0^2 f_0 r_1 \delta_r^\alpha)$  is the leading nonzero term in the two-timescale expansion of  $\Gamma_{\beta\gamma}^\alpha(z^\mu) u^\beta u^\gamma$ .

In the evolution equation for  $r_0$ , (103), we require  $\tilde{f}_1^t$ .

From Eq. (203a), this evaluates to

$$\tilde{f}_1^t = -\frac{1}{2}f_0^{-1}\Omega_0\tilde{h}_{u_0u_0,\phi_p}^{1\mathcal{R}}, \quad (204a)$$

$$= \frac{1}{2r_0f_0} \sum_{\ell m} a_{\ell\ell}i\omega_m(\tilde{s})R_{\ell\ell m}^{1\mathcal{R}}(\tilde{s}, r_0) \times Y_{\alpha\beta}^{\ell m}(r_0, \pi/2, 0)u_0^\alpha u_0^\beta. \quad (204b)$$

Since this formula only involves  $m \neq 0$  modes,  $\bar{x}_{\ell m}$  does not contribute. Therefore, the entirety of the right-hand side is fully determined by the value of  $r_0$ , and we can write  $\tilde{f}_1^t = \tilde{f}_1^t(r_0)$ . Again because it has no fast time dependence, we also need not include  $\bar{x}_{\ell m}$  in the leading-order waveform.

Once the evolution of  $r_0$  is determined, it determines  $\Omega_0$  and hence the adiabatic phase  $\int \Omega_0 dt$ . Therefore this collection of results provides us with the adiabatic evolution scheme outlined in the introduction and summarized in the upper box of Fig. 3. In this scheme, we can entirely neglect  $\dot{M}_1$  and  $\dot{S}_1$ .

## B. Post-adiabatic order

Before moving onto the second-order field equations, we can extract more from the first-order solutions. Specifically, if we know the values of  $M_1$  and  $S_1$ , we can compute the other independent component of  $\tilde{f}_1^\alpha$ :

$$\tilde{f}_1^r = \frac{1}{2}f_0\tilde{h}_{u_0u_0,r}^{1\mathcal{R}} \quad (205a)$$

$$= \frac{1}{2}f_0 \sum_{\ell m} a_{\ell\ell}\partial_{r_0}[R_{\ell\ell m}^{1\mathcal{R}}(\tilde{s}, r_0)Y_{\alpha\beta}^{\ell m}(r_0, \pi/2, 0)]u_0^\alpha u_0^\beta. \quad (205b)$$

This gets contributions from both  $m \neq 0$  and  $m = 0$  modes, meaning it depends on  $\bar{x}_{\ell m}(M_1, S_1, r)$ . We can therefore write it as  $\tilde{f}_1^r = \tilde{f}_1^r(r_0, M_1, S_1)$ . (If we were interested in conservative effects on an orbit around a nonspinning black hole on a given slice of slow time, we would evaluate this with  $S_1$  set to zero.)

Next we turn to the second-order field equations (153) and (158). In the source term  $\delta^2 G_{\ell m}^0$  in Eq. (153), the only required input is  $\Omega_0$  and the first-order amplitudes. In the source term  $E_{ij\ell m}^0 R_{j\ell m}^{2\mathcal{P}}$ , we require  $h_{\alpha\beta}^{1\mathcal{R}}$  evaluated on  $z_0^\mu$ , as we see from Eq. (39) (derivatives of  $h_{\alpha\beta}^{1\mathcal{R}}$  are also required for higher-order terms in that equation). As mentioned below Eq. (149), we also transfer the effect of  $\tilde{T}_{\ell m}^1$  into  $R_{\ell m}^{2\mathcal{P}}$ . These terms are proportional to  $r_1$ ,  $\Omega_1$ , and  $\dot{r}_0$ .  $r_1$  we can consider freely specified until we begin to evolve it;  $\Omega_1$  is then determined from  $r_1$  and  $\tilde{f}_1^r$  via Eq. (104); and  $\dot{r}_0$  is determined from  $\tilde{f}_1^t$  via Eq. (103).

The final source term,  $E_{ij\ell m}^1 R_{j\ell m}^1$ , given by Eq. (144) with (141) and (C2), requires as input the slow time derivative of the mode amplitudes  $R_{\ell m}^1$ . From the form

of  $R_{\ell m}^1$  in Eqs. (188)–(190), this derivative is given by

$$\partial_{\tilde{s}} R_{\ell m}^1 = \dot{r}_0 \frac{\partial}{\partial r_0} R_{\ell m}^{\text{pp}} + \dot{M}_1 \frac{\partial}{\partial M_1} \bar{x}_{\ell m} + \dot{S}_1 \frac{\partial}{\partial S_1} \bar{x}_{\ell m}. \quad (206)$$

$\frac{\partial}{\partial M_1} \bar{x}_{\ell m}$  and  $\frac{\partial}{\partial S_1} \bar{x}_{\ell m}$  are easily calculated analytically from the formulas for  $x_{\ell m}$  in Appendix D. In Ref. [70] we present a method of calculating  $\frac{\partial}{\partial r_0} R_{\ell m}^{\text{pp}}$  at a given value of  $r_0$  (without taking a numerical derivative, which would require  $R_{\ell m}^{\text{pp}}$  in a range of neighboring  $r_0$  values).  $\dot{r}_0$  is computed from Eq. (103). We will return to  $\dot{M}_1$  and  $\dot{S}_1$  momentarily.

For  $\ell \geq 2$ , and for  $\ell = 1, m = \pm 1$ , the  $\dot{M}_1$  and  $\dot{S}_1$  terms in Eq. (206) are not involved, and additional corrections to the black hole's mass and spin cannot appear. Given the form of the source described above, the solution to Eqs. (153) and (158) for these modes hence takes the form

$$R_{\ell m}^2(\tilde{s}, r) = \hat{R}_{\ell m}^2(r_0, r_1, M_1, S_1, r), \quad (207)$$

where  $r_0, r_1, M_1$ , and  $S_1$  are functions of  $\tilde{s}$ . Note that  $M_1$  and  $S_1$  appear because the second-order source for any given  $\ell m$  gets contributions from every first-order mode. Concretely,  $\bar{x}_{\ell m}(M_1, S_1, r)$  contributes to every mode of  $\delta^2 G_{\ell m}^0$  by coupling to other modes, and it contributes to  $R_{\ell m}^{2\mathcal{P}}$  through its contribution to  $\tilde{h}_{\alpha\beta}^{1\mathcal{R}}$  and  $\tilde{f}_1^r$ .

For the  $\ell = 0$  and  $\ell = 1, m = 0$  modes, two things change: mass and spin perturbations,  $\bar{x}_{\ell m}(M_2, S_2, r)$ , appear exactly as at first order, and  $\dot{M}_1$  and  $\dot{S}_1$  appear as sources. Just as we can neglect  $\bar{x}_{\ell m}(M_1, S_1, r)$  in an adiabatic evolution, we can neglect  $\bar{x}_{\ell m}(M_2, S_2, r)$  in a first post-adiabatic evolution. In fact, we can neglect the second-order  $\ell = 0$  and  $\ell = 1, m = 0$  modes entirely. However, we do require  $\dot{M}_1$  and  $\dot{S}_1$  regardless, as they determine the evolution of  $\bar{x}_{\ell m}(M_1, S_1, r)$ , which must be included in  $\tilde{f}_1^r$ ,  $R_{\ell m}^{2\mathcal{P}}$ , and  $\delta^2 G_{\ell m}^0$ .

$\dot{M}_1$  and  $\dot{S}_1$  are determined directly from the  $\ell = 0$  and  $\ell = 1, m = 0$  field equations. The wave equation (153) with  $\ell = 0$  can be solved with an arbitrary value of  $\dot{M}_1$ , and the wave equation with  $\ell = 1, m = 0$  can be solved with an arbitrary value of  $\dot{S}_1$ . However, the gauge condition (158) then uniquely fixes the values of  $\dot{M}_1$  and  $\dot{S}_1$ . The result is unsurprising: these quantities grow at precisely the rate that gravitational-wave energy and angular momentum enter the black hole,

$$\dot{M}_1 = \dot{E}_H, \quad \text{and} \quad \dot{S}_1 = \dot{L}_H, \quad (208)$$

where the energy and angular momentum fluxes down the horizon,  $\dot{E}_H$  and  $\dot{L}_H$ , are obtained from the  $\ell \geq 2$  amplitudes  $R_{\ell m}^{\text{pp}}$ ; see, for example Eq. (76) of Ref. [68]. This is in agreement with traditional analyses of the slow evolution of the central black hole [77, 78]. But we emphasize here that the results are obtained directly from the vacuum Einstein equations, rather than from the dynamics

of the horizon generators as in those traditional analyses. In fact, the results do not even require any knowledge of the metric anywhere near the horizon. (Physically, this is a consequence of the fact that the total flux through a surface of constant  $r$  is the same for all  $r < r_0$ .) We present the complete derivation in the next section.

Once the mode amplitudes  $R_{ilm}^{2\mathcal{R}}$  are known, one can calculate  $\tilde{f}_2^t$ , the principal input for  $\dot{r}_1$  in Eq. (106). From Eq. (203b), it is given by

$$\begin{aligned} \tilde{f}_2^t(\tilde{t}) = & -\frac{1}{2}f_0^{-1}\Omega_0\tilde{h}_{u_0u_0,\phi_p}^{2\mathcal{R}} - f_0^{-1}\Omega_0\tilde{h}_{u_0u_1,\phi_p}^{1\mathcal{R}} \\ & + \frac{1}{2}f_0^{-2}\Omega_0r_1\left(f_0'\tilde{h}_{u_0u_0,\phi_p}^{1\mathcal{R}} - f_0\tilde{h}_{u_0u_0,r\phi_p}^{1\mathcal{R}}\right) \\ & + \frac{1}{2}\left(-f_0^{-1}\delta_t^\beta + U_0u_0^\beta\right)\left(2\Gamma^\gamma h_{\beta\gamma}^{1\mathcal{R}} - 2U_0u_0^\gamma\partial_t\tilde{h}_{\beta\gamma}^{1\mathcal{R}}\right. \\ & + f_0^{-1}\Omega_0\tilde{h}_{\beta t}^{1\mathcal{R}}\tilde{h}_{u_0u_0,\phi_p}^{1\mathcal{R}} - f_0\tilde{h}_{\beta r}^{1\mathcal{R}}\tilde{h}_{u_0u_0,r}^{1\mathcal{R}} \\ & \left.+ r_0^{-2}\tilde{h}_{\beta\phi}^{1\mathcal{R}}\tilde{h}_{u_0u_0,\phi_p}^{1\mathcal{R}}\right). \end{aligned} \quad (209)$$

The notation here is as described below Eq. (203b).  $f_0' = 2M/r_0^2$ , and derivatives are evaluated following the examples of Eqs. (204) and (205). We have used the facts that  $\tilde{h}_{u_0u_0,\phi}^{1\mathcal{R}} = -\tilde{h}_{u_0u_0,\phi_p}^{1\mathcal{R}}$  and  $\tilde{h}_{u_0u_0,\theta}^{1\mathcal{R}} = 0$ ; the latter follows from the system's up-down symmetry. Note that because of the fast time derivative in the first term, to evaluate  $\tilde{f}_2^t$  we only require the  $m \neq 0$  modes of  $\tilde{h}_{\alpha\beta}^{2\mathcal{R}}$ . Both in  $\tilde{f}_2^t(\tilde{t})$  and in  $d\tilde{f}_1^r/d\tilde{t}$ , which appears in Eq. (106), we require the slow-time derivative of  $\tilde{h}_{\alpha\beta}^{1\mathcal{R}}$ ; this is calculated as in Eq. (206).

With  $\dot{r}_1$  determined, one can evolve  $r_1$  and hence evolve  $\Omega_1$  [through Eq. (104)] and the post-adiabatic phase  $\int(\Omega_0 + \varepsilon\Omega_1)dt$ . Therefore this collection of results provides us with the post-adiabatic evolution scheme outlined in the introduction and summarized in Fig. 3. In this scheme, we can entirely neglect  $M_2$  and  $S_2$ . The leading-order waveform amplitudes  $R_{ilm}^{\text{pp}}$  are also independent of  $M_1$  and  $S_1$ , but  $M_1$  and  $S_1$  do contribute to the post-adiabatic phase in the numerous ways mentioned above.

## VII. APPLICATION: LEADING-ORDER BALANCE LAWS

As a first demonstration of our scheme, in this section we derive the standard balance laws relating the system's evolution to the flux of energy and angular momentum carried out to infinity and into the black hole. The end result has been well established for decades [17, 19, 85], but our derivation, a slight variant of the one sketched in Ref. [39], is novel in that it proceeds directly from the two-timescale field equations.

### A. Evolution of mass and orbital energy

We begin with the evolution of the mass  $M_1$  and leading-order orbital energy  $\mathcal{E}_0 = \mu f_0 U_0$ . These evolution equations follow specifically from the field equations for the  $i = 2, \ell = 0$  field  $R_{200}^2$ ; this corresponds to the angle-averaged  $t$ - $r$  component of the field equations.

Before writing down the field equations, let us motivate their connection to energy fluxes. Note that if we consider  $-\frac{1}{8\pi}\delta^2 G_{\alpha\beta}^0$  as an effective stress-energy tensor that sources  $\tilde{h}_{\alpha\beta}^2$ , then the energy crossing outward (toward larger radii) through a surface of constant  $r$  is

$$\Delta E_{S_r} = \frac{1}{8\pi} \int_{S_r} \delta^2 G_{\alpha\beta}^0 t^\beta dS^\alpha \quad (210a)$$

$$= \frac{1}{8\pi} \int \delta^2 G_{\alpha\beta}^0 t^\beta r^\alpha r^2 dt d\Omega. \quad (210b)$$

Here  $t^\alpha = \delta_t^\alpha$  and  $r_\alpha = \partial_\alpha r$ . The average rate of energy transfer across the surface is therefore

$$\dot{E}_{S_r} = \frac{1}{8\pi} f r^2 \int \delta^2 G_{tr}^0 d\Omega \quad (211a)$$

$$= -\frac{r}{\sqrt{4\pi}f} \delta^2 G_{200}^0, \quad (211b)$$

where we have used Eq. (118). This demonstrates the connection between the energy flux and the  $i = 2, \ell = 0$  mode of the source.

We can make the connection to the specific fluxes  $\dot{E}_H$  and  $\dot{E}_\infty$  by appealing to the concrete form of  $\delta^2 G_{200}$ . At any vacuum point (i.e., at all points off the worldline), the contracted Bianchi identity reads  $\nabla^\beta \delta^2 G_{\alpha\beta} = 0$ .<sup>9</sup> The mode decomposition of this is identical to the mode decomposition of  $\nabla^\beta \tilde{h}_{\alpha\beta} = 0$ , and the analog of Eq. (155a), specialized to  $\ell = 0$ , is

$$r\partial_r(f^{-1}\delta^2 G_{200}^0) + f^{-1}\delta^2 G_{200}^0 = 0. \quad (212)$$

This applies for all  $r \neq r_0$ . The extra factors of  $f^{-1}$ , relative to Eq. (155a), arise from the differing factors in the decompositions (115) and (117) of  $\tilde{h}_{\alpha\beta}$  and  $\delta^2 G_{\alpha\beta}^0$ . The solution to Eq. (212) is

$$\delta^2 G_{200}^0 = \frac{f s_{200}^-}{r} \theta(r_0 - r) + \frac{f s_{200}^+}{r} \theta(r - r_0) \quad (213)$$

<sup>9</sup> This equality holds at fixed slow time for the coefficient of  $e^{-im\phi_p}$  in  $\nabla^\beta \delta^2 G_{\alpha\beta}$ ; we do not need to include slow-time derivatives from  $\nabla^\beta \delta G_{\alpha\beta} = 0$ , as one might expect from performing a two-timescale expansion of the Bianchi identity. The reason is that the coefficient of each  $e^{-im\phi_p}$  in the first-order field is identical to what the coefficient of  $e^{-i\omega_m s}$  would be in ordinary perturbation theory. It follows that if we omit all slow time derivatives, the coefficient of each  $e^{-im\phi_p}$  in  $\nabla^\beta \delta^2 G_{\alpha\beta}$  is also identical to what the coefficient of  $e^{-i\omega_m s}$  would be in ordinary perturbation theory, and therefore it satisfies the same identities.

for some constants  $s_{200}^\pm$ . Together with Eq. (211), this result implies that  $\dot{E}_{S_r} = -s_{200}^+/\sqrt{4\pi}$  for all  $r > r_0$ , and  $\dot{E}_{S_r} = -s_{200}^-/\sqrt{4\pi}$  for all  $r < r_0$ . Therefore, for all radii smaller than  $r_0$ , we have  $\dot{E}_{S_r} = -\dot{E}_H$ , and for all radii larger than  $r_0$ , we have  $\dot{E}_{S_r} = \dot{E}_\infty$ . (The minus sign appears in front of  $\dot{E}_H$  because we define  $\dot{E}_H$  to be an inward flux, toward smaller radii.) Combining these results, we obtain the following simple expression for  $\delta^2 G_{200}^0$ :

$$\delta^2 G_{200}^0 = \frac{\sqrt{4\pi}f}{r} \left[ \dot{E}_H \theta(r_0 - r) - \dot{E}_\infty \theta(r - r_0) \right]. \quad (214)$$

A more detailed analysis is required, starting from the form of the puncture field, to establish that no term proportional to  $\delta(r - r_0)$  appears. For the purpose of our simple demonstration, we omit those details here.

Because  $\delta^2 G_{200}^0$  is manifestly integrable, we can simplify the field equations by using the second-order Detweiler stress-energy (47) instead of a puncture. That stress-energy is proportional to  $\mu u_\alpha u_\beta$ , meaning the  $t$ - $r$  component is proportional to  $\mu u_r \sim \varepsilon^2$ . Explicitly, the  $t$ - $r$  component has the simple form

$$T_{tr} = \mu \frac{U_0 \dot{r}_0}{r_0^2} \delta(r - r_0) \delta(\theta - \pi/2) \delta(\phi - \phi_p) + \mathcal{O}(\varepsilon^3); \quad (215)$$

corrections due to  $h_{\mu\nu}^{\mathcal{R}}$  and  $r_1$  appear only at third order. From this we can straightforwardly read off  $T_{200} = \varepsilon^2 \tilde{t}_{200}^2 \delta(r - r_0)$ , following the notation in Eq. (145), where

$$\tilde{t}_{200}^2 = -\frac{\mathcal{E}_0 f_0 \dot{r}_0}{4\sqrt{\pi} r_0}. \quad (216)$$

We have used  $\mathcal{E}_0 = \mu f_0 U_0$  to express the source in terms of the orbital energy.

Now, there are two field equations involving  $R_{200}^2$ : the wave equation (153) with  $i = 2$  and  $\ell = 0$ , which reads

$$\begin{aligned} -\frac{1}{4}(\partial_{r^*}^2 - 2Mf/r^3)R_2^2 + \mathcal{M}_0^{2j}R_j^2 \\ = -16\pi\tilde{t}_{200}^2\delta(r - r_0) + 2\delta^2 G_{200}^0 \\ + \frac{M}{r^2} \left[ fH\dot{R}_3^1 + (1 - H)\dot{R}_1^1 \right], \end{aligned} \quad (217)$$

and the gauge condition (158) with  $k = 1$  and  $\ell = 0$ , which reads

$$\mathcal{L}_2 R_2^2 = \dot{R}_1^1 + f\dot{R}_3^1, \quad (218)$$

where  $\mathcal{L}_2 = f(\partial_r + 1/r)$ . Here and below we omit the  $\ell m$  label on  $R_{ilm}^n$ , and we adopt a first-order monopole solution in which  $R_{200}^1 = 0$  (implying  $\dot{R}_{200}^1 = 0$ ). The term  $\mathcal{M}_0^{2j}R_j^2$  in the field equation couples  $R_2^2$  to  $R_1^1$ ,  $R_3^2$ , and  $R_6^2$ , but we can eliminate that coupling by substituting the gauge condition (158) with  $k = 2$ , which reduces  $\mathcal{M}_0^{2j}R_j^2$  to

$$\mathcal{M}_0^{2j}R_j^2 = \frac{ff'}{2}\partial_r R_2^2 + \frac{f^2}{2r^2}R_2^2 - \frac{f'}{2}H(\dot{R}_1^1 - f\dot{R}_3^1). \quad (219)$$

Equation (217) then becomes

$$\mathcal{D}_2 R_2^2 = S_2^2, \quad (220)$$

where  $\mathcal{D}_2 = \left( f^2 \partial_r^2 - ff' \partial_r - \frac{ff'}{r} - \frac{2f^2}{r^2} \right)$  and

$$\begin{aligned} S_2^2 = \frac{16\sqrt{\pi}\mathcal{E}_0 f_0 \dot{r}_0}{r_0} \delta(r - r_0) - \frac{4M}{r^2} \dot{R}_1^1 \\ - \frac{16\sqrt{\pi}f}{r} \left[ \dot{E}_H \theta(r_0 - r) - \dot{E}_\infty \theta(r - r_0) \right]. \end{aligned} \quad (221)$$

Note that the dependence on the height function has now vanished, implying that the evolution equations will be independent of the choice of slow-time slicing.

In the field equation (220),  $\dot{\mathcal{E}}_0$  (or equivalently,  $\dot{r}_0$ ) and  $\dot{M}_1$  act as sources for  $R_2^2$ . We could solve this field equation for arbitrary values of those sources. The gauge condition (218) would then determine the relationships between  $\dot{\mathcal{E}}_0$ ,  $\dot{M}_1$ ,  $\dot{E}_H$ , and  $\dot{E}_\infty$ . That was the general method sketched in Ref. [39], and it is in line with the general descriptions in Sec. II.

However, the explicit analytical solution to Eq. (220) is distractingly lengthy. For compactness, we take another approach. It is well known that a solution to  $E_{\alpha\beta}[\bar{h}] = S_{\alpha\beta}$  satisfies  $\nabla^\beta \bar{h}_{\alpha\beta} = 0$  if and only if  $\nabla^\beta S_{\alpha\beta} = 0$ ; this follows from  $\nabla^\beta S_{\alpha\beta} = \nabla^\beta E_{\alpha\beta}[\bar{h}] = \square \nabla^\beta \bar{h}_{\alpha\beta}$ . So any information about the system's evolution that we can glean from the gauge condition, we can equally well extract directly from the source. Inspired by that, we will derive a differential equation for the source, which will yield the desired evolution equations.

We first note the commutation relation

$$\mathcal{L}_2 \mathcal{D}_2 - \mathcal{D}_2 \mathcal{L}_2 = \frac{2f}{r^2} \mathcal{L}_2. \quad (222)$$

Applying  $\mathcal{L}_2$  to the field equation (220), using this commutation relation, and substituting Eq. (218) for  $\mathcal{L}_2 R_2^2$ , we obtain

$$\left( \mathcal{D}_2 + \frac{2f}{r^2} \right) (\dot{R}_1^1 + f\dot{R}_3^1) = \mathcal{L}_2 S_2^2. \quad (223)$$

We write each side of this equation in the form  $a(r_0)\delta'(r - r_0) + b(r_0)\delta(r - r_0) + c^+(r)\theta(r - r_0) + c^-(r)\theta(r_0 - r)$ , allowing us to equate the coefficients of  $\delta'$ ,  $\delta$ , and  $\theta[\pm(r - r_0)]$ . The coefficients of  $\delta'$  yield

$$[\dot{R}_1^1 + f_0 \dot{R}_3^1] = \frac{16\sqrt{\pi}\mathcal{E}_0 \dot{r}_0}{r_0}, \quad (224)$$

where  $[x] := \lim_{r \rightarrow r_0^+} x - \lim_{r \rightarrow r_0^-} x$ . The coefficients of  $\delta$  yield

$$\begin{aligned} [(\partial_r - f'f^{-1})(\dot{R}_1^1 + f\dot{R}_3^1)] = -\frac{4M}{r_0^2 f_0} [\dot{R}_1^1] + \frac{16\sqrt{\pi}\mathcal{E}_0 \dot{r}_0}{r_0^2 f_0} \\ + \frac{16\sqrt{\pi}}{r_0} (\dot{E}_H + \dot{E}_\infty), \end{aligned} \quad (225)$$

and the coefficients of the Heaviside functions yield

$$\left(\mathcal{D}_2 + \frac{2f}{r^2}\right)(\dot{R}_1^{\pm} + f\dot{R}_3^{\pm}) = \mathcal{L}_2 S_2^{\pm}, \quad (226)$$

where the  $+$  applies for  $r > r_0$  and the minus for  $r < r_0$ .

The two equations (226) suffice to determine  $\dot{M}_1$  and  $\dot{\mathcal{E}}_0$ . Substituting the first-order monopole field  $R_{i00}^1 = R_{i00}^{pp} + \bar{x}_{i00}(M_1)$  given in Appendix D 2 b, and using  $\dot{\mathcal{E}}_0 = (d\mathcal{E}_0/dr_0)\dot{r}_0$ , we quickly obtain<sup>10</sup>

$$\dot{M}_1 = \dot{E}_H \quad (227)$$

from the  $r < r_0$  equation, and

$$\dot{\mathcal{E}}_0 + \dot{M}_1 = -\dot{E}_\infty \quad (228)$$

from the  $r > r_0$  equation. The first of these, as discussed in previous sections, tells us that the black hole's mass grows at the rate that energy flows down the horizon. The second tells us that the system's total energy changes at the rate that energy is carried out of it. We can also combine the two equations to write

$$\dot{\mathcal{E}}_0 = -(\dot{E}_\infty + \dot{E}_H), \quad (229)$$

which says that at leading order, all the energy leaving the region  $2M < r < \infty$  comes from the particle's orbital energy. We stress that obtaining these equations for  $\dot{\mathcal{E}}_0$  and  $\dot{M}_1$  does not require evaluating Eqs. (226) at any particular values of  $r$ ; Eqs. (229) and (227) ensure that Eqs. (226) are satisfied at *all*  $r$ .

Although we can also solve Eqs. (224) and (225), they do not contain any additional information. Eq. (224) is satisfied for any of the solutions  $R_{i00}^1 = R_{i00}^{pp} + \bar{x}_{i00}(M_1)$  in Appendix D 2, regardless of the value of  $\dot{r}_0$ , while Eq. (225) again yields Eq. (229).  $\bar{x}_{i00}(M_1)$  does not enter into either of the equations because it is smooth at  $r = r_0$ .

We have already emphasised that our derivation proceeds directly from the field equations, but there is one other aspect worth drawing attention to: it shows that  $-\frac{1}{8\pi}\delta^2 G_{\alpha\beta}^0$  provides a meaningful notion of gravitational stress-energy throughout the spacetime, in the sense that it defines appropriate fluxes of gravitational energy (and in the next section, of gravitational angular momentum). This may appear obvious, since the Isaacson gravitational-wave stress-energy is derived from  $\delta^2 G_{\alpha\beta}$  [86]. However, earlier analyses of the evolution of a black hole (e.g., Refs. [77, 78]) have stressed that Isaacson's derivation does not apply at the horizon, and

they have instead derived the black hole's evolution from the shear of the horizon generators. The “fluxes” that emerge from such a derivation are then directly the rates of change of the black hole parameters rather than surface integrals of a stress-energy. To our knowledge, our derivation provides the first demonstration that the relevant physical fluxes are precisely those defined from  $-\frac{1}{8\pi}\delta^2 G_{\alpha\beta}^0$ . Reference [87] will extend this result to generic bound orbits around a Kerr black hole (and to post-adiabatic order).

## B. Evolution of spin and orbital angular momentum

We next derive the evolution of the spin  $S_1$  and leading-order orbital angular momentum  $\mathcal{L}_0 = \mu r_0^2 \Omega_0 U_0$ . These evolution equations follow specifically from the field equations for the  $i = 9, \ell = 1, m = 0$  field  $R_{910}^2$ ; this corresponds to the angle-averaged  $r$ - $\phi$  component of the field equations.

The analysis is much the same as in the preceding section, and our presentation will be terse. The angular momentum carried outward across a surface of constant  $r$  is

$$\Delta L_{S_r} = -\frac{1}{8\pi} \int \delta^2 G_{\alpha\beta}^0 \phi^\beta dS^\alpha, \quad (230)$$

where  $\phi^\beta = \delta_\phi^\beta$ , implying that the average rate of angular momentum transfer is

$$\dot{L}_{S_r} = -\frac{1}{8\pi} f r^2 \int \delta^2 G_{r\phi}^0 d\Omega \quad (231a)$$

$$= \frac{r^2}{2\sqrt{3}\pi f} \delta^2 G_{910}^0. \quad (231b)$$

Here we have again used Eq. (118).

The contracted Bianchi identity implies

$$\partial_r(f^{-1}\delta^2 G_{910}^0) + \frac{2}{rf}\delta^2 G_{910}^0 = 0 \quad (232)$$

for all  $r \neq r_0$ ; c.f. Eq. (155d), specialized to  $\ell = 1, m = 0$ . The solution for  $r \neq r_0$  is

$$\delta^2 G_{910}^0 = \frac{s_{910}^- f}{r^2} \theta(r_0 - r) + \frac{s_{910}^+ f}{r^2} \theta(r - r_0) \quad (233)$$

for constants  $s_{910}^\pm$ . Substituting this into Eq. (231), we see that  $\dot{L}_{S_r} = \frac{s_{910}^\pm}{2\sqrt{3}\pi}$ , where the  $+$  applies for all  $r > r_0$  and the minus for all  $r < r_0$ . Therefore, writing  $\dot{L}_{S_{r>r_0}} = \dot{L}_\infty$  and  $\dot{L}_{S_{r<r_0}} = -\dot{L}_H$ , we have

$$\delta^2 G_{910}^0 = -\frac{2\sqrt{3}\pi f}{r^2} \left[ \dot{L}_H \theta(r_0 - r) - \dot{L}_\infty \theta(r - r_0) \right]. \quad (234)$$

<sup>10</sup> The same result is also obtained if we use the monopole solution discussed in Appendix D 2 b. On the other hand, if instead we use the Berndtson solution  $R_{i00}^1 = R_{i00}^{\text{Bern}} + \bar{x}_{i00}(M_1)$ , then we still obtain Eq. (229), but (227) is replaced by Eq. (D30), corresponding to the fact that the black hole's mass is  $M_{\text{BH}} = M + \varepsilon(M_{\text{Bern}} + M_1)$  in this solution.



We once again use the Detweiler stress-energy (47) instead of a puncture. The  $r$ - $\phi$  component is

$$T_{r\phi} = \mu f_0^{-1} U_0 \dot{r}_0 \Omega_0 \delta(r - r_0) \delta(\theta - \pi/2) \delta(\phi - \phi_p) + \mathcal{O}(\varepsilon^3), \quad (235)$$

from which we read off  $T_{910} = \varepsilon^2 \tilde{t}_{910}^2 \delta(r - r_0)$ , where

$$\tilde{t}_{910}^2 = -\frac{1}{4} \mu \sqrt{\frac{3}{\pi}} f_0 U_0 \dot{r}_0 \Omega_0. \quad (236)$$

With these sources in hand, we can write down the two field equations involving  $R_{910}^2$ : First, the wave equation (153) with  $i = 9, \ell = 1, m = 0$ , which reads

$$\mathcal{D}_9 R_{910}^2 = S_9^2, \quad (237)$$

where  $\mathcal{D}_9 = f^2 \partial_r^2 + f f' \partial_r - \frac{2f}{r^3} (3r - 8M)$  and

$$S_9^2 = 64\pi \tilde{t}_{910}^2 \delta(r - r_0) - 8\delta^2 G_{910}^0 + 2f H \partial_r \dot{R}_{910}^1 + H' \dot{R}_9^1. \quad (238)$$

Second, the gauge condition (158) with  $k = 4, \ell = 1, m = 0$ , which reads

$$\mathcal{L}_9 R_{910}^2 = \dot{R}_{810}^1 + H \dot{R}_{910}^1, \quad (239)$$

where  $\mathcal{L}_9 = f(\partial_r + 2/r)$ . Here we have suppressed  $\ell m$  indices as in the previous subsection.

The operators  $\mathcal{L}_9$  and  $\mathcal{D}_9$  satisfy the commutation relation

$$\mathcal{L}_9 \mathcal{D}_9 - \mathcal{D}_9 \mathcal{L}_9 = \frac{4}{r^3} f(r - 4M) \mathcal{L}_9. \quad (240)$$

Applying  $\mathcal{L}_9$  to Eq. (237), using this relation, and substituting Eq. (239), we obtain an equation that relates  $\dot{\mathcal{L}}_0$  and  $\dot{S}_1$  to  $\dot{L}_H$  and  $\dot{L}_\infty$ :

$$\left( \mathcal{D}_9 + \frac{4f}{r^3} (r - 4M) \right) (\dot{R}_{810}^1 + H \dot{R}_{910}^1) = \mathcal{L}_9 S_9^2. \quad (241)$$

This can again be divided into one equation for the coefficients of  $\delta'$ , one for the coefficients of  $\delta$ , and one each for the coefficients of  $\theta[\pm(r - r_0)]$ . As with the energy balance laws, all the information can be obtained from the two equations in the regions  $r < r_0$  and  $r > r_0$ . The former yields

$$\dot{S}_1 = \dot{L}_H, \quad (242)$$

and the latter yields

$$\dot{\mathcal{L}}_0 + \dot{S}_1 = -\dot{L}_\infty. \quad (243)$$

Substituting the first into the second, we obtain

$$\dot{\mathcal{L}}_0 = -(\dot{L}_\infty + \dot{L}_H). \quad (244)$$

We note that once again, the results are independent of the choice of height function in our slow time slicing.

## VIII. CONCLUSION AND OUTLOOK

Since the seminal work of Hinderer and Flanagan [13], the two-timescale approach has provided a common language for discussions of the long-term evolution of EMRIs. However, it has never actually been implemented, nor has any attempt been made (until now) to apply it to the full system of equations in an EMRI, which involve not only the small companion's equation of motion, but also the perturbed Einstein equations.

In this paper, we have taken a first step toward a complete two-timescale treatment of EMRIs. Our analysis is restricted to the narrow case of quasicircular orbits around a Schwarzschild black hole, but it suffices to highlight many of the key features of the full problem:

1. The gravitational field is written as a sum of modes (133), each with a slowly varying amplitude, which evolves on the radiation-reaction time scale, and a rapidly oscillating phase, which varies on the orbital time scale.
2. The field equations split into two sets. One set, given by Eqs. (152) and (153), take the form of standard frequency-domain equations for the mode amplitudes at fixed values of slow time. The other set, given by Eqs. (103), (106), and (208), take the form of evolution equations, determining how the field evolves as a function of slow time.
3. Slow-time derivatives of the first-order field amplitudes also appear as source terms in the second-order field equation (153). The asymptotic behavior of these sources, near infinity as well as near the future horizon, depend strongly on the choice of time foliation.

Our formulation in this restricted case should therefore provide a guide to the full problem.

As part of our two-timescale expansion, we have also introduced a frequency-domain formulation of the field equations with hyperboloidal slicing. Although we have cast our field equations in the language of a two-timescale expansion, the equations (152) apply equally well in an ordinary frequency-domain calculation, with a metric perturbation  $\bar{h}_{\alpha\beta} = \sum_{i\ell m} \frac{a_{i\ell}}{r} R_{i\ell m}(r) e^{-i\omega_m s} Y_{\alpha\beta}^{i\ell m}$ . Such hyperboloidal slicing in the frequency domain has been considered in the past for the Teukolsky equation [66, 67], but this is the first time, to our knowledge, that it has been used for the full linearized Einstein equation. A companion paper will present a practical method of solving these equations with various choices of hyperboloidal slicing [70]. That paper will also apply the same method to calculate the slow-time derivatives of the first-order field.

Our expansions in the body of the paper were formulated in terms of series in  $\varepsilon^n$  at fixed values of a slow time variable. In Appendix A, we also presented an equivalent method based on expansions at fixed values of the

orbital frequency, which works with a nonperturbative  $\Omega$  rather than the series  $\Omega_0 + \varepsilon\Omega_1 + \dots$ . This alternative provides slightly neater equations, and it may look more familiar to researchers working on waveform generation in other regimes of the two-body problem, such as post-Newtonian theory [88].

In sequel papers, we will present several extensions and applications of this work. Still considering quasicircular orbits in Schwarzschild, we will present the complete calculation of the second-order source, including the second-order Einstein tensor and punctures at the particle [71, 72]; the formulation of physical boundary conditions for the second-order field, obtained from a post-Minkowskian expansion near future null infinity and an analogous expansion near the future horizon; the calculation of quasistationary quantities such as the Detweiler redshift [80]; and the post-adiabatic inspiral and emitted waveforms. Going beyond the quasicircular case, we will present a two-timescale expansion for the full problem of generic orbits in Kerr spacetime [62, 63], along with accompanying treatments near the horizon and infinity [64, 65].

This two-timescale expansion will ultimately provide a complete, flexible wave-generation framework, as outlined in the quasicircular case in Sec. IC and Fig. 3. For a generic orbit in Kerr spacetime, the system is characterized by a set of adiabatic-order mechanical parameters  $J_0^a = (\mathcal{E}_0, \mathcal{L}_0, \mathcal{Q}_0, M, S)$  and post-adiabatic ones  $J_1^a = (\mathcal{E}_1, \mathcal{L}_1, \mathcal{Q}_1, M_1, S_1)$ , where  $S = aM$  is the central black hole's leading-order spin,  $\mathcal{L}$  is the orbital angular momentum around the black hole's spin axis, and  $\mathcal{Q}$  is the Carter constant. With the exception of the background quantities  $M$  and  $S$ , each of these evolves as a function of slow time (after eliminating any fast-time dependence using near-identity averaging transformations [48]). One can pre-compute the rates of change  $\frac{dJ_0^a}{dt}(J_0^b)$  and  $\frac{dJ_1^a}{dt}(J_0^b, J_1^b)$ , along with the waveform amplitudes  $h_{\alpha\beta}^{\omega_{mkn}}(J_0^B)$ , across the parameter space. One can then quickly evolve through the parameter space by integrating simple ODEs. The trajectory through parameter space then determines the post-adiabatic evolution of the frequencies  $\omega_{mkn} = \omega_{mkn}^{(0)}(J_0^b) + \varepsilon\omega_{mkn}^{(1)}(J_0^b, J_1^b)$  and the waveform  $\sum_{mkn} h_{\alpha\beta}^{\omega_{mkn}}(J_0^b(\varepsilon t))e^{-i\int\omega_{mkn}(J_0^b(\varepsilon t), J_1^b(\varepsilon t))dt}$ . This framework is described by one of us in Ref. [38] (which was submitted some time after the present paper). It is easily extended to include additional effects, such as the small object's spin, in a modular way; such effects can again be precomputed before simulating an evolution. In principle this framework can stand alone, but to cope with the very large parameter space, its outputs can also be used as input for effective-one-body and surrogate models or as training data for neural networks [89–91].

## ACKNOWLEDGMENTS

We are grateful to Leor Barack, Niels Warburton, and Barry Wardell for helpful discussions and to the anonymous referee for correcting numerous typos. AP additionally thanks Eanna Flanagan and Jordan Moxon for enlightening comments on the central black hole's evolution, and he acknowledges support from a Royal Society University Research Fellowship.

### Appendix A: Two-timescale expansion with nonperturbative mechanical degrees of freedom

In Sec. II's description of the self-consistent framework, the governing philosophy is to express the metric perturbations as functionals of the system's mechanical degrees of freedom, and then to obtain approximate evolution equations for those degrees of freedom. While the evolution equations are approximate, the mechanical degrees of freedom themselves are treated nonperturbatively (i.e., never expanded in an asymptotic series).

In the two-timescale framework developed in the body of the paper, and summarized in Fig. 3, we abandon that nonperturbative treatment and treat the metric and mechanical parameters on an equal footing: all quantities are expanded in powers of  $\varepsilon$  at fixed  $\tilde{s}$  and  $\phi_p$ .

In this appendix, we present a slightly different formulation of the two-timescale expansion that preserves more of the spirit of the self-consistent expansion. We do so by expanding all quantities at fixed  $J^a := (\Omega, \delta\mathcal{M}, \delta S)$  instead of at fixed  $\tilde{s}$ . Here we have introduced the order-1 quantities  $\delta\mathcal{M} := \delta M/\varepsilon$  and  $\delta S := \delta S/\varepsilon$ . The conceptual distinction is that rather than solving fast-time field equations to find the state of the system at a given value of slow time, and then evolving to the next slow time, here we will solve equations to find the state of the system for a given set of mechanical parameters, and then evolve to new values of those parameters. However, we stress that both expansions yield solutions that are uniformly accurate on the slow time scale. They differ only in details of implementation.

In the new expansion considered here, in place of Eq. (89), the worldline becomes  $z^\mu(t, \varepsilon) = (t, r_p(J^a(t, \varepsilon), \varepsilon), \pi/2, \phi_p(t, \varepsilon))$ , where

$$r_p(J^a, \varepsilon) = r_0(\Omega) + \varepsilon r_1(J^a) + \mathcal{O}(\varepsilon^2), \quad (\text{A1})$$

using the fact that  $\delta M$  and  $\delta S$  do not appear in the adiabatic-order term  $r_0$ . Put another way, we expand the worldline as

$$z^\mu = z_0^\mu(t, \Omega, \phi_p) + \varepsilon z_1^\mu(J^a) + \mathcal{O}(\varepsilon^2), \quad (\text{A2})$$

where  $z_0^\mu(t, \Omega, \phi_p) = (t, r_0(\Omega), \pi/2, \phi_p)$  and  $z_1^\mu = (0, r_1(J^a), 0, 0)$ . Rather than a nonperturbative trajectory  $z^\mu$ , as in the self-consistent expansion, here  $J^a$  and  $\phi_p$  are our exact mechanical degrees of freedom.

Similarly, in place of (133) the metric is expanded as

$$\bar{h}_{\mu\nu} = \sum_{n\ell m} \frac{\varepsilon^n a_{i\ell}}{r} R_{i\ell m}^n(J^a(\tilde{s}, \varepsilon), r) e^{-im\phi_p(s, \varepsilon)} Y_{\mu\nu}^{i\ell m}(r, \theta^A). \quad (\text{A3})$$

The metric's time dependence is entirely encoded in  $J^a$  and  $\phi_p$ .

Finally, because *all* functions are expanded at fixed  $J^a$ , we also expand the functions  $\dot{J}^a := dJ^a/d\tilde{s}$ :

$$\dot{J}^a(J^b, \varepsilon) = F_0^a(\Omega) + \varepsilon F_1^a(J^b) + \mathcal{O}(\varepsilon^2), \quad (\text{A4})$$

again using the fact that  $\delta\mathcal{M}$  and  $\delta\mathcal{S}$  do not appear at adiabatic order. The terms on the right are related to (but not identical to)  $\dot{J}_n^a = \frac{dJ_n^a}{d\tilde{s}}$ , the derivative of the  $n$ th-order term in the expansion of  $J^a$  at fixed  $\tilde{s}$ . The two quantities can be related by re-expanding the right-hand side of Eq. (A4) at fixed  $\tilde{s}$  and comparing the result to  $\frac{d}{d\tilde{s}}[J_0^a(\tilde{s}) + \varepsilon J_1^a(\tilde{s}) + \mathcal{O}(\varepsilon^2)]$ . This yields the relationships  $\dot{J}_0^a(\tilde{s}) = F_0^a(\Omega_0(\tilde{s}))$  and  $\dot{J}_1^a(\tilde{s}) = F_1^a(J_0^b(\tilde{s})) + \Omega_1 \partial_{\Omega_0} F_0^a(\Omega_0(\tilde{s}))$ .

Equation (A4) is precisely what governs the system's slow evolution from one state to the next. Its form, an approximate equation for the exact state variables  $J^a$ , is akin to the treatment of the motion in the self-consistent expansion, in which we write an approximate equation  $\frac{D^2 z^\alpha}{d\tau^2} = \varepsilon f_1^\alpha + \varepsilon^2 f_2^\alpha + \mathcal{O}(\varepsilon^3)$  for the exact worldline  $z^\alpha$ .

Given the above expansions, the equation of motion and field equations are solved just as in Secs. IV and V. We substitute the expansions, apply the chain rules  $(\frac{\partial}{\partial t})_r = \varepsilon \dot{J}^a \partial_a + \Omega \partial_{\phi_p}$  and  $(\frac{\partial}{\partial r^*})_t = \partial_{r^*} - \varepsilon H(\dot{J}^a \partial_a + \Omega \partial_{\phi_p})$ , and then solve the resulting equations while treating  $J^a$  and  $\phi_p$  as independent variables.

We will merely summarize the results here. Substituting the expansions (A2), (A4),  $U = U_0(\Omega) + \varepsilon U_1(J^a) + \mathcal{O}(\varepsilon^2)$ , and

$$f^t = \varepsilon \tilde{f}_1^t(\Omega) + \varepsilon^2 \tilde{f}_2^t(J^a) + \mathcal{O}(\varepsilon^3), \quad (\text{A5})$$

$$f^r = \varepsilon \tilde{f}_1^r(J^a) + \mathcal{O}(\varepsilon^2), \quad (\text{A6})$$

into the normalization condition  $g_{\alpha\beta} u^\alpha u^\beta = -1$  and equation of motion  $\frac{D^2 z^\alpha}{d\tau^2} = f^\alpha$ , we obtain a sequence of equations for  $U_n$  (from the normalization condition),  $r_n$  (from the radial component of the equation of motion), and  $F_n^\Omega$  (from the time component). The solutions are

$$U_0 = \frac{1}{\sqrt{1 - 3(M\Omega)^{2/3}}}, \quad U_1 = 0, \quad (\text{A7})$$

$$r_0 = \frac{M}{(M\Omega)^{2/3}}, \quad r_1 = \frac{\tilde{f}_1^r}{3U_0^2 f_0 \Omega^2}, \quad (\text{A8})$$

with  $f_0 = 1 - 2(M\Omega)^{2/3}$ , and

$$F_0^\Omega = -\frac{3f_0 \Omega \tilde{f}_1^t}{(M\Omega)^{2/3} U_0^4 [1 - 6(M\Omega)^{2/3}]}, \quad (\text{A9})$$

$$F_1^\Omega = -\frac{3f_0 \Omega \tilde{f}_2^t}{(M\Omega)^{2/3} U_0^4 [1 - 6(M\Omega)^{2/3}]} - \frac{2\tilde{J}_0^a \partial_a \tilde{f}_1^r}{MU_0^4 [(M\Omega)^{1/3} - 8M\Omega + 12(M\Omega)^{5/3}]} - \frac{4[1 - 6(M\Omega)^{2/3} + 12(M\Omega)^{4/3}] \tilde{f}_1^r \tilde{f}_1^t}{M\Omega U_0^6 f_0 [1 - 6(M\Omega)^{2/3}]^2}. \quad (\text{A10})$$

Note that the expressions for  $U_0$ ,  $U_1$ ,  $r_0$ , and  $r_1$  here are identical to Eqs. (97), (98), (102), and (104) with  $\Omega_0 = \Omega$  and  $\Omega_1 = 0$ .

Next, the expansion of the field equations proceeds just as in Sec. V, and the results in that section apply with only small changes.  $\tilde{\partial}_{\tilde{s}} = \partial_{\tilde{s}} - im\Omega_1$  becomes  $\partial_{\tilde{s}} = F_0^a \partial_a (= \dot{J}_0^a \partial_a)$ .  $\omega_m = m\Omega_0$  becomes  $\omega_m = m\Omega$ .  $\dot{\omega}_m$  becomes  $mF_0^\Omega (= m\dot{\Omega}_0)$ . The two-timescale expansion of  $T_{i\ell m}^1$  (or equivalently, of the puncture) is slightly altered by the condition  $\Omega_1 = 0$ , but we leave a concrete description to a sequel paper.

This naturally provides us with a wave-generation scheme that differs slightly from the one in Fig. 3:

1. Start with initial values of  $\Omega$ ,  $\delta M$ , and  $\delta S$  (along with  $\mu$  and  $M$ ).
2. Calculate  $r_0(\Omega)$  from Eq. (A8).
3. Calculate  $\tilde{T}_{i\ell m}^1(\Omega)$  from  $r_0(\Omega)$ .
4. Solve the field equations for the mode amplitudes  $R_{i\ell m}^1(J^a, r)$  [and, separately, for  $\partial_a R_{i\ell m}^1(J^a, r)$ , as will be described in Ref. [70]].
5. From  $R_{i\ell m}^1(J^a, r)$  and  $\partial_a R_{i\ell m}^1(J^a, r)$ , calculate  $\delta^2 G_{i\ell m}^0(J^a)$ ,  $\tilde{f}_1^\alpha(J^a)$ ,  $F_0^{\delta\mathcal{M}} = \dot{E}_H(\Omega)$ , and  $F_0^{\delta\mathcal{S}} = \dot{L}_H(\Omega)$ . From  $\tilde{f}_1^t(\Omega)$  in (A9), calculate  $F_0^\Omega$ .
6. From  $\delta^2 G_{i\ell m}^0(J^a)$ ,  $E_{i\ell m}^1 \sim F_0^a \partial_a R_{i\ell m}^1$ , and the puncture, construct the second-order source.
7. Solve the second-order field equations for the mode amplitudes  $R_{i\ell m}^2(J^a, r)$ .
8. From  $R_{i\ell m}^2(J^a, r)$ , calculate  $\tilde{f}_2^t(J^a)$ , and from  $\tilde{f}_2^t(J^a)$  in (A10), calculate  $F_1^\Omega$ .
9. Use the post-adiabatic approximations  $d\Omega/d\tilde{s} = F_0^\Omega(\Omega) + \varepsilon F_1^\Omega(J^a)$ ,  $d(\delta\mathcal{M})/d\tilde{s} = F_0^{\delta\mathcal{M}}(\Omega)$ , and  $d(\delta\mathcal{S})/d\tilde{s} = F_0^{\delta\mathcal{S}}(\Omega)$  to evolve forward to new values of  $J^a$ .
10. Repeat the above steps for as long as desired.
11. Once the evolution of the frequency is known, obtain the orbital phase from  $\phi_p = \frac{1}{\varepsilon} \int \Omega(\tilde{s}) d\tilde{s}$  and construct the waveform from Eq. (A3).

## Appendix B: Barack-Lousto-Sago tensor spherical harmonics

The Barack-Lousto-Sago harmonics are given in Schwarzschild coordinates by

$$Y_{\alpha\beta}^{1\ell m} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & f^{-2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} Y^{\ell m}, \quad Y_{\alpha\beta}^{2\ell m} = \frac{f^{-1}}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} Y^{\ell m}, \quad Y_{\alpha\beta}^{3\ell m} = \frac{1}{\sqrt{2}} \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & -f^{-1} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} Y^{\ell m}, \quad (\text{B1a})$$

$$Y_{\alpha\beta}^{4\ell m} = \frac{r}{\sqrt{2\lambda_1}} \begin{pmatrix} 0 & 0 & \partial_\theta & \partial_\phi \\ 0 & 0 & 0 & 0 \\ \partial_\theta & 0 & 0 & 0 \\ \partial_\phi & 0 & 0 & 0 \end{pmatrix} Y^{\ell m}, \quad Y_{\alpha\beta}^{5\ell m} = \frac{rf^{-1}}{\sqrt{2\lambda_1}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \partial_\theta & \partial_\phi \\ 0 & \partial_\theta & 0 & 0 \\ 0 & \partial_\phi & 0 & 0 \end{pmatrix} Y^{\ell m}, \quad (\text{B1b})$$

$$Y_{\alpha\beta}^{6\ell m} = \frac{r^2}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & s^2 \end{pmatrix} Y^{\ell m}, \quad Y_{\alpha\beta}^{7\ell m} = \frac{r^2}{\sqrt{2\lambda_2}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & D_2 & D_1 \\ 0 & 0 & D_1 & -s^2 D_2 \end{pmatrix} Y^{\ell m}, \quad (\text{B1c})$$

$$Y_{\alpha\beta}^{8\ell m} = \frac{r}{\sqrt{2\lambda_1}} \begin{pmatrix} 0 & 0 & s^{-1}\partial_\phi & -s\partial_\theta \\ 0 & 0 & 0 & 0 \\ s^{-1}\partial_\phi & 0 & 0 & 0 \\ -s\partial_\theta & 0 & 0 & 0 \end{pmatrix} Y^{\ell m}, \quad (\text{B2a})$$

$$Y_{\alpha\beta}^{9\ell m} = \frac{rf^{-1}}{\sqrt{2\lambda_1}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & s^{-1}\partial_\phi & -s\partial_\theta \\ 0 & s^{-1}\partial_\phi & 0 & 0 \\ 0 & -s\partial_\theta & 0 & 0 \end{pmatrix} Y^{\ell m}, \quad (\text{B2b})$$

$$Y_{\alpha\beta}^{10\ell m} = \frac{r^2}{\sqrt{2\lambda_2}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & s^{-1}D_1 & -sD_2 \\ 0 & 0 & -sD_2 & -sD_1 \end{pmatrix} Y^{\ell m}, \quad (\text{B2c})$$

where  $s := \sin \theta$ ,  $\lambda_1 := \ell(\ell+1)$ ,  $\lambda_2 := (\ell-1)\ell(\ell+1)(\ell+2)$ ,  $Y^{\ell m} = Y^{\ell m}(\theta, \phi)$  are the standard scalar spherical harmonics, and

$$D_1 := 2(\partial_\theta - \cot \theta)\partial_\phi, \quad (\text{B3})$$

$$D_2 := \partial_{\theta\theta} - \cot \theta \partial_\theta - s^{-2}\partial_{\phi\phi}. \quad (\text{B4})$$

The radial factors involving  $r$  and  $f$  are introduced to make the modes  $\bar{h}_{ilm}$  dimensionless and to ensure that if the components  $\bar{h}_{\alpha\beta}$  are regular at the future horizon in horizon-penetrating coordinates, then each of the modes  $\bar{h}_{ilm}$  is as well.

This basis is orthogonal with respect to a certain inner product, satisfying

$$\oint d\Omega \eta^{\alpha\mu} \eta^{\beta\nu} Y_{\mu\nu}^{ilm} Y_{\alpha\beta}^{*j\ell'm'} = \kappa_i \delta_{ij} \delta_{\ell\ell'} \delta_{mm'}, \quad (\text{B5})$$

where  $d\Omega = \sin \theta d\theta d\phi$  is the surface element on the unit

sphere,

$$\eta^{\alpha\beta} := \text{diag}(1, f^2, r^{-2}, r^{-2} \sin^{-2} \theta), \quad (\text{B6})$$

and

$$\kappa_i := \begin{cases} f^2 & \text{if } i = 3, \\ 1 & \text{otherwise.} \end{cases} \quad (\text{B7})$$

Note that our expression for  $\eta^{\alpha\beta}$  corrects a typo in the original formula in Ref. [81] (also corrected in Ref. [34]).

The coefficients  $a_{i\ell}$  are introduced in Eq. (115) for the purpose of simplifying the field equations. They are defined to be

$$a_{i\ell} = \frac{1}{\sqrt{2}} \times \begin{cases} 1 & \text{for } i = 1, 2, 3, 6, \\ 1/\sqrt{\lambda_1} & \text{for } i = 4, 5, 8, 9, \\ 1/\sqrt{\lambda_2} & \text{for } i = 7, 10, \end{cases} \quad (\text{B8})$$

where  $\lambda_1$  and  $\lambda_2$  are as defined above.

### Appendix C: Field equations

We give here explicit expressions for the matrix elements  $\mathcal{M}_n^{ij}$  appearing in the wave operators (144) and

field equations (152)–(153). For brevity, we omit the labels  $n, \ell, m$  and the arguments  $\tilde{s}, r$  of the fields  $R_{i\ell m}^n(\tilde{s}, r)$ . For the same reason, we use  $\lambda_1 = \ell(\ell + 1)$ , as in Appendix B, and introduce  $\lambda := \lambda_2/\lambda_1 = (\ell + 2)(\ell - 1)$ . The partial derivative  $\partial_r$  is taken with fixed  $\tilde{s}$ .

With those shorthand notations, the quantities  $\mathcal{M}_0^{ij} R_j$  in Eqs. (152) are given by

$$\mathcal{M}_0^{1j} R_j = \frac{f^2 f'}{2} \left( \partial_r R_3 + \frac{i\omega_m H}{f} R_3 \right) + \frac{f \left( 1 - \frac{4M}{r} \right)}{2r^2} (R_1 - R_5 - f R_3) - \frac{f^2}{2r^2} \left( 1 - \frac{6M}{r} \right) R_6, \quad (\text{C1a})$$

$$\begin{aligned} \mathcal{M}_0^{2j} R_j &= \frac{f^2 f'}{2} \left( \partial_r R_3 + \frac{i\omega_m H}{f} R_3 \right) + \frac{f f'}{2} \partial_r (R_2 - R_1) - \frac{i\omega_m}{2} (1 - H) f' (R_2 - R_1) \\ &\quad + \frac{f^2}{2r^2} (R_2 - R_4) - \frac{f f'}{2r} (R_1 - R_5 - f R_3 - 2f R_6), \end{aligned} \quad (\text{C1b})$$

$$\mathcal{M}_0^{3j} R_j = -\frac{f}{2r^2} \left[ R_1 - R_5 - \left( 1 - \frac{4M}{r} \right) (R_3 + R_6) \right], \quad (\text{C1c})$$

$$\mathcal{M}_0^{4j} R_j = \frac{f f'}{4} \partial_r (R_4 - R_5) - \frac{i\omega_m (1 - H) f'}{4} (R_4 - R_5) - \frac{\lambda_1}{2} \frac{f}{r^2} R_2 - \frac{f f'}{4r} (3R_4 + 2R_5 - R_7 + \lambda_1 R_6), \quad (\text{C1d})$$

$$\mathcal{M}_0^{5j} R_j = \frac{f}{r^2} \left[ \left( 1 - \frac{9M}{2r} \right) R_5 - \frac{\lambda_1}{2} (R_1 - f R_3) + \frac{1}{2} \left( 1 - \frac{3M}{r} \right) (\lambda_1 R_6 - R_7) \right], \quad (\text{C1e})$$

$$\mathcal{M}_0^{6j} R_j = -\frac{f}{2r^2} \left[ R_1 - R_5 - \left( 1 - \frac{4M}{r} \right) (R_3 + R_6) \right], \quad (\text{C1f})$$

$$\mathcal{M}_0^{7j} R_j = -\frac{f}{2r^2} (R_7 + \lambda R_5), \quad (\text{C1g})$$

$$\mathcal{M}_0^{8j} R_j = \frac{f f'}{4} \partial_r (R_8 - R_9) - \frac{i\omega_m (1 - H) f'}{4} (R_8 - R_9) - \frac{f f'}{4r} (3R_8 + 2R_9 - R_{10}), \quad (\text{C1h})$$

$$\mathcal{M}_0^{9j} R_j = \frac{f}{r^2} \left( 1 - \frac{9M}{2r} \right) R_9 - \frac{f}{2r^2} \left( 1 - \frac{3M}{r} \right) R_{10}, \quad (\text{C1i})$$

$$\mathcal{M}_0^{10j} R_j = -\frac{f}{2r^2} (R_{10} + \lambda R_9), \quad (\text{C1j})$$

where  $f' = \partial f / \partial r = 2M/r^2$ .

The quantities  $\mathcal{M}_1^{ij} R_j$  in Eqs. (144) and (153) are given by

$$\mathcal{M}_1^{1j} R_j = -\frac{1}{2} f f' H \tilde{\partial}_{\tilde{s}} R_3, \quad (\text{C2a})$$

$$\begin{aligned} \mathcal{M}_1^{2j} R_j &= -\frac{f'}{2} \left[ f H \tilde{\partial}_{\tilde{s}} R_3 \right. \\ &\quad \left. - (1 - H) \tilde{\partial}_{\tilde{s}} (R_2 - R_1) \right], \end{aligned} \quad (\text{C2b})$$

$$\mathcal{M}_1^{4j} R_j = \frac{f'}{4} (1 - H) \tilde{\partial}_{\tilde{s}} (R_4 - R_5), \quad (\text{C2c})$$

$$\mathcal{M}_1^{8j} R_j = \frac{f'}{4} (1 - H) \tilde{\partial}_{\tilde{s}} (R_8 - R_9), \quad (\text{C2d})$$

$$\mathcal{M}_1^{ij} R_j = 0 \quad \text{for } i = 3, 5, 6, 9, 10. \quad (\text{C2e})$$

### Appendix D: Analytic solutions for the $\ell = 0$ and $\ell = 1, m = 0$ first-order modes

In this appendix we present the explicit analytical formulas for the solutions (189) and (190)

#### 1. Dipole solution

We first consider  $\ell = 1, m = 0$ , which only has an odd-parity contribution. For  $R_{i10}^{\text{PP}}$  we use

$$R_{810}^{\text{PP}} = -16 \sqrt{\frac{\pi}{3}} \mathcal{L}_0 \begin{cases} r^2/r_0^3 & \text{for } r \leq r_0, \\ 1/r & \text{for } r \geq r_0, \end{cases} \quad (\text{D1a})$$

$$R_{910}^{\text{PP}} = -\sqrt{\frac{\pi}{3}} \frac{256 M^4 \mathcal{L}_0}{r_0^3 r^2}, \quad (\text{D1b})$$

where  $\mathcal{L}_0 = \mu r_0^2 U_0 \Omega_0$  is the zeroth-order orbital angular momentum. Note that our solution (D1) differs from the one that has been used historically. The historical solution, dating back to Zerilli's classic paper [92], agrees

with our  $R_{810}^{\text{PP}}$  but sets  $R_{910}^{\text{PP}} = 0$ ; though this solution is often assumed to be horizon regular, it actually violates the regularity condition (162c), leading to divergent behavior of the second-order source term  $\delta^2 G_{i\ell m}^0$  at the horizon. The addition of an  $i = 9$  mode (which is pure gauge) satisfying Eq. (162c) cures this ill behavior.<sup>11</sup> In components, the perturbation at large  $r$  reads

$$h_{t\phi}^{pp, \ell=1, m=0} = -\frac{2\mathcal{L}_0 \sin^2 \theta}{r} + \mathcal{O}(1/r^4), \quad (\text{D2})$$

the expected asymptotic form of the metric in a space-time with angular momentum  $\mathcal{L}_0$ . One can straightforwardly verify, using Komar or Abbott-Deser integrals [93] over surfaces of constant  $r$  (at fixed slow time), that the spacetime  $g_{\alpha\beta} + \varepsilon h_{\alpha\beta}^{\text{PP}}$  has zero angular momentum inside every sphere of radius  $r < r_0$ , corresponding to a non-spinning central black hole, and angular momentum  $\varepsilon \mathcal{L}_0$  inside every sphere of radius  $r > r_0$ .

On top of  $R_{i10}^{\text{PP}}$ , we can freely add a perturbation generated by the central black hole's small, evolving spin  $S_1$ , given by

$$\bar{x}_{810} = -\sqrt{\frac{\pi}{3}} \frac{16S_1}{r}, \quad (\text{D3a})$$

$$\bar{x}_{910} = -\sqrt{\frac{\pi}{3}} \frac{32MS_1}{r^2}. \quad (\text{D3b})$$

The  $i = 9$  mode is again added to ensure regularity at the future horizon.  $\bar{x}_{i\ell m}$  corresponds to a metric perturbation with  $t\phi$  component  $x_{t\phi}^{\ell=1, m=0} = -\frac{2S_1 \sin^2 \theta}{r} + \mathcal{O}(1/r^4)$ , and one can verify that the spacetime  $g_{\alpha\beta} + \varepsilon x_{\alpha\beta}$  contains angular momentum  $\varepsilon S_1$  for all  $r > 2M$ .

## 2. Monopole solutions

Next, we consider  $\ell = 0$ . The analysis of this mode is complicated by a well-known pathology of the Lorenz gauge: in the Lorenz gauge, there is no globally regular homogeneous solution with nonzero mass. A homogeneous solution with nonzero mass is always irregular at either the horizon or at infinity (or both) [60, 93]. These singularities are purely a gauge artefact; outside the Lorenz gauge, it is easy to find mass perturbations that are regular at the future horizon and at infinity. For example, in an ‘Eddington-Finkelstein gauge’, a homogeneous perturbation with mass  $M_1$  has the simple, manifestly regular form

$$x_{\alpha\beta}^{\text{EF}} = \frac{2M_1}{r} \delta_\alpha^v \delta_\beta^v \quad (\text{D4})$$

(obtained by replacing  $M$  with  $M + \varepsilon M_1$  in the Eddington-Finkelstein components of the exact Schwarzschild metric). However, here we restrict our attention to Lorenz-gauge solutions. We consider several solutions and their relative merits.

### a. Basis of homogeneous solutions

All possible solutions to Eq. (152) are easily constructed from the complete basis of homogeneous solutions provided by Dolan and Barack [93], the main members of which we denote  $\{h_{\alpha\beta}^{(A)}, h_{\alpha\beta}^{(B)}, h_{\alpha\beta}^{(C)}, h_{\alpha\beta}^{(D)}\}$ , following Dolan and Barack's labelling. With the metric perturbation written as  $h_{\alpha\beta} = \text{diag}(h_{tt}, h_{rr}, Hr^2, Hr^2 \sin^2 \theta)$ , these are given by

$$h_{\alpha\beta}^{(A)} = g_{\alpha\beta}, \quad (\text{D5})$$

$$h_{tt}^{(B)} = -\frac{2MfP(r)}{r^3}, \quad h_{rr}^{(B)} = \frac{2Q(r)}{fr^3}, \quad (\text{D6})$$

$$H^{(B)} = \frac{2fP(r)}{r^2}, \quad (\text{D7})$$

$$h_{tt}^{(C)} = -\frac{2M^4}{r^4}, \quad h_{rr}^{(C)} = -\frac{2M^3(2r - 3M)}{f^2 r^4}, \quad (\text{D8})$$

$$H^{(C)} = \frac{2M^3}{r^3}, \quad (\text{D9})$$

$$h_{tt}^{(D)} = -\frac{2}{3r^4} [rW(r) + MrfP(r) \ln f - 8M^4 \ln(r/M)], \quad (\text{D10})$$

$$h_{rr}^{(D)} = \frac{2}{3r^4 f^2} [L(r) \ln f - rK(r) + 8M^3(2r - 3M) \ln(r/M)], \quad (\text{D11})$$

$$H^{(D)} = \frac{2}{3r^3} \{ (r^3 - 8M^3) \ln f - r[P(r) - Mr] - 8M^3 \ln(r/M) \}, \quad (\text{D12})$$

where

$$P := r^2 + 2rM + 4M^2, \quad (\text{D13})$$

$$Q := r^3 - r^2M - 2rM^2 + 12M^3, \quad (\text{D14})$$

$$W := 3r^3 - 7r^2M - rM^2 - 4M^3, \quad (\text{D15})$$

$$K := r^3 - 5r^2M - 5rM^2 + 12M^3, \quad (\text{D16})$$

$$L := r^4 - 3r^3M + 16rM^3 - 24M^4. \quad (\text{D17})$$

(The  $H$  here should not be confused with  $H = dk/dr^*$ .) There is also a fifth independent solution, consisting solely of a  $t$ - $r$  component,

$$h_{tr}^{(E)} = \frac{M^2}{fr^2}. \quad (\text{D18})$$

This was not identified by Dolan and Barack, but it appears as part of one of the solutions they consider. We will not require explicit use of it.

<sup>11</sup> This longstanding error was discovered in collaboration with Leor Barack, Niels Warburton, and Barry Wardell. We thank Leor Barack for deriving the correct, nonzero  $R_{910}^{\text{PP}}$ .

Given the metric components, the Barack-Lousto-Sago amplitudes are

$$R_{100} = \sqrt{4\pi}r (h_{tt}^{\ell=0} + f^2 h_{rr}^{\ell=0}), \quad (\text{D19a})$$

$$R_{200} = \sqrt{16\pi}r f h_{tr}^{\ell=0}, \quad (\text{D19b})$$

$$R_{300} = \sqrt{16\pi}r H^{\ell=0}, \quad (\text{D19c})$$

$$R_{600} = \sqrt{4\pi}r f^{-1} (h_{tt}^{\ell=0} - f^2 h_{rr}^{\ell=0}). \quad (\text{D19d})$$

Together,  $h_{\alpha\beta}^{(A)}$  through  $h_{\alpha\beta}^{(E)}$  form a complete basis. From Table I, we see that for  $\ell = 0$  there are two coupled field equations for  $i = 1, 3$ , plus the hierarchically decoupled field equation for  $i = 2$ , with the  $i = 6$  mode recovered algebraically from the gauge condition (157) with (155b). The  $i = 1, 3$  equations should have a total of four independent homogeneous solutions; these are provided by the (A)–(D) solutions. When these solutions are substituted into the  $i = 2$  field equation, one finds that the  $i \neq 2$  modes cancel one another in  $\mathcal{M}_0^{ij} R_{j\ell m}$ , leaving a fully decoupled equation for  $i = 2$ . The gauge condition (157) with (155a) also provides a decoupled equation for  $i = 2$ , and it is easy to verify that the sole solution to the gauge condition automatically satisfies the field equation; this is the (E) solution

$h_{\alpha\beta}^{(A)}$  is the only one of the five solutions to contain mass, equal to  $M/2$ ; all others are pure gauge. It is regular at the future horizon but irregular at infinity (where its components go to constants instead of decaying to zero).  $h_{\alpha\beta}^{(B)}$  is regular at the future horizon but irregular at infinity.  $h_{\alpha\beta}^{(C)}$  and  $h_{\alpha\beta}^{(E)}$  are regular at infinity but not at the future horizon.  $h_{\alpha\beta}^{(D)}$  is irregular at both boundaries.

*b. Asymptotically regular, horizon-irregular solution*

We first consider solutions that are asymptotically flat and in which  $\bar{x}_{i00}(M_1, r)$  contains the full perturbation to the black hole's mass.

An inhomogeneous solution to Eq. (152) that is regular at infinity, contains no mass in the region  $r < r_0$  (and therefore no correction to the black hole's mass), and is maximally regular at the horizon is given by

$$h_{\alpha\beta}^{\text{pp}, \ell=0} = \left( 2\mathcal{E}_0 h_{\alpha\beta}^{(A)} + a_+^{(B)} h_{\alpha\beta}^{(B)} + a_+^{(C)} h_{\alpha\beta}^{(C)} + a_+^{(D)} h_{\alpha\beta}^{(D)} \right) \theta^+ + \left( a_-^{(B)} h_{\alpha\beta}^{(B)} + a_-^{(C)} h_{\alpha\beta}^{(C)} + a_-^{(D)} h_{\alpha\beta}^{(D)} \right) \theta^-, \quad (\text{D20})$$

where  $\theta^\pm := \theta[\pm(r - r_0)]$  and

$$a_+^{(B)} = -\frac{4}{3}\mathcal{E}_0, \quad (\text{D21a})$$

$$a_+^{(C)} = -\frac{\mathcal{E}_0}{3Mr_0 f_0} [M^2(8\ln 2 - 44) + r_0^2 + 8M(r_0 - 3M)\ln(r_0/M) + 20Mr_0], \quad (\text{D21b})$$

$$a_+^{(D)} = -\mathcal{E}_0, \quad (\text{D21c})$$

$$a_-^{(B)} = -\frac{(r_0 - 3M)\ln(f_0)\mathcal{E}_0}{3r_0 f_0}, \quad (\text{D21d})$$

$$a_-^{(C)} = -\frac{4M(5 + 8\ln 2)\mathcal{E}_0}{3r_0 f_0}, \quad (\text{D21e})$$

$$a_-^{(D)} = -\frac{M\mathcal{E}_0}{r_0 f_0}. \quad (\text{D21f})$$

The mode amplitudes are then given by Eq. (D19). Since only  $h_{\alpha\beta}^{(A)}$  contains mass, we can read off that in the space-time  $g_{\alpha\beta} + \varepsilon h_{\alpha\beta}^{\text{pp}}$ , a sphere of radius  $r < r_0$  contains mass  $M$ , and a sphere of radius  $r > r_0$  contains mass  $M + \varepsilon\mathcal{E}_0$ . This describes the spacetime of a particle of mass  $\mu$  orbiting a black hole of mass  $M$ .

The perturbation is regular at large  $r$ , where

$$h_{\alpha\beta}^{\text{pp}, \ell=0} = \frac{2\mathcal{E}_0}{r} \text{diag}(1, 1, r^2, r^2 \sin^2 \theta) + \mathcal{O}(1/r^2), \quad (\text{D22})$$

and

$$R_{100}^{\text{pp}} = 8\sqrt{\pi}\mathcal{E}_0 + \mathcal{O}(1/r), \quad (\text{D23a})$$

$$R_{200}^{\text{pp}} = 0, \quad (\text{D23b})$$

$$R_{300}^{\text{pp}} = 8\sqrt{\pi}\mathcal{E}_0 + \mathcal{O}(1/r), \quad (\text{D23c})$$

$$R_{600}^{\text{pp}} = \mathcal{O}(1/r). \quad (\text{D23d})$$

But it diverges logarithmically at the horizon, where

$$R_{100}^{\text{pp}} = \mathcal{O}(f^2), \quad (\text{D24a})$$

$$R_{200}^{\text{pp}} = 0, \quad (\text{D24b})$$

$$R_{300}^{\text{pp}} = \mathcal{O}(f), \quad (\text{D24c})$$

$$R_{600}^{\text{pp}} = \frac{8\sqrt{\pi}}{r_0 f_0} \left[ (r_0 - 3M)\ln f_0 + M \ln \left( \frac{r}{2M} - 1 \right) \right] + \mathcal{O}(f \ln f). \quad (\text{D24d})$$

One can construct other inhomogeneous solutions with the same properties by adding perturbations proportional to  $h_{\alpha\beta}^{(E)}$ . But it is straightforward to check that no such solution can improve upon the above solution's regularity at the horizon without sacrificing regularity at infinity.

Next, an asymptotically flat mass perturbation with maximal regularity at the horizon is given by

$$x_{\alpha\beta}^{\ell=0}(M_1, r) = 2M_1 h_{\alpha\beta}^{(A)} - \frac{4}{3}M_1 h_{\alpha\beta}^{(B)} - \frac{4}{3}(5 + \ln 4)M_1 h_{\alpha\beta}^{(C)} - M_1 h_{\alpha\beta}^{(D)}. \quad (\text{D25})$$



At large  $r$ , it behaves as in Eq. (D23) with the replacement  $\mathcal{E}_0 \rightarrow M_1$ . Near the horizon it behaves as

$$\bar{x}_{100} = \mathcal{O}(f^2), \quad (\text{D26a})$$

$$\bar{x}_{200} = 0, \quad (\text{D26b})$$

$$\bar{x}_{300} = 8\sqrt{\pi}M_1 \left[ 2 + \ln \left( \frac{r}{2M} - 1 \right) \right] + \mathcal{O}(f), \quad (\text{D26c})$$

$$\bar{x}_{600} = 16\sqrt{\pi}M_1 + \mathcal{O}(f). \quad (\text{D26d})$$

Again, one can easily check that this is the maximal regularity that an asymptotically flat Lorenz-gauge mass perturbation can have at the horizon.

The perturbation  $h_{\alpha\beta}^{\text{pp},\ell=0} + x_{\alpha\beta}^{\ell=0}(M_1, r)$  in Eqs. (D20) and (D25) fits neatly into Sec. VI's description of the full two-timescale solution and its evolution. In particular, the black hole's total mass is  $M_{\text{BH}} = M + \varepsilon M_1$ , and it satisfies the flux equation  $dM_{\text{BH}}/d\tilde{s} = \varepsilon dM_1/d\tilde{s} = \varepsilon \dot{E}_H$ .

The disadvantage of this solution is obvious. Because it is singular at the horizon, it makes the second-order source term  $\delta^2 G_{ilm}^0$  also singular there. And because the  $\ell = 0$  mode of  $h_{\alpha\beta}^1$  contributes to every mode  $\delta^2 G_{ilm}^0$ , this irregularity spreads into every second-order mode. Of course once the solution is obtained, the singularity can be eliminated with a gauge transformation. But one must first carefully obtain the correct (singular) boundary conditions for each mode.

### c. Berndtson solution

We next consider asymptotically flat solutions in which  $x_{\alpha\beta}^{\ell=0}(M_1, r)$  does *not* contain the full perturbation to the black hole's mass. This implies that the inhomogeneous solution  $h_{\alpha\beta}^{\text{pp},\ell=0}$  contains part of the correction to the black hole mass, which turns out to allow it to be regular at both boundaries.

The relevant regular inhomogeneous solution was first derived by Berndtson [94]. Denoting it by  $h_{\alpha\beta}^{\text{Bern}}$  to distinguish it from the  $h_{\alpha\beta}^{\text{pp},\ell=0}$  of Eq. (D20), we can write it as

$$\begin{aligned} h_{\alpha\beta}^{\text{Bern}} = & \left[ 2M_{>} h_{\alpha\beta}^{(A)} + b_+^{(B)} h_{\alpha\beta}^{(B)} + b_+^{(C)} h_{\alpha\beta}^{(C)} \right. \\ & + b_+^{(D)} h_{\alpha\beta}^{(D)} \Big] \theta^+ + \left[ 2M_{<} h_{\alpha\beta}^{(A)} + b_-^{(B)} h_{\alpha\beta}^{(B)} \right. \\ & + b_-^{(C)} h_{\alpha\beta}^{(C)} + b_-^{(D)} h_{\alpha\beta}^{(D)} \Big] \theta^-, \end{aligned} \quad (\text{D27})$$

where

$$b_{\pm}^{(B)} = a_{\pm}^{(B)} - \frac{4}{3} M_{<}, \quad (\text{D28a})$$

$$b_{\pm}^{(C)} = a_{\pm}^{(C)} - \frac{4}{3} (5 + \ln 4) M_{<}, \quad (\text{D28b})$$

$$b_{\pm}^{(D)} = a_{\pm}^{(D)} - M_{<}. \quad (\text{D28c})$$

Here

$$M_{<} = -\frac{\mathcal{E}_0}{r_0 f_0} =: M_{\text{Bern}} \quad (\text{D29})$$

is the mass contained in a sphere of radius  $r < r_0$ , and  $M_{>} = (M_{\text{Bern}} + \mathcal{E}_0)$  is the mass contained in a sphere of radius  $r > r_0$ .

$h_{\alpha\beta}^{\text{Bern}}$  is regular at both future null infinity and the future horizon, satisfying all the conditions of (162), as desired. If we exclude any nonzero  $t$ - $r$  components,  $h_{\alpha\beta}^{\text{Bern}}$  is in fact the unique Lorenz-gauge solution that is regular at both boundaries. (If we include  $h_{\alpha\beta}^{(E)}$ , we can construct other regular solutions, but none of them improve upon the Berndtson solution.)  $h_{\alpha\beta}^{\text{Bern}}$  also differs from Eq. (D20) in that it contains mass,  $M_{\text{Bern}}$ , in the region  $r < r_0$ ; this corresponds to a perturbation to the black hole's mass. Contrary to statements in Refs. [60, 93], this does not suggest  $h_{\alpha\beta}^{\text{Bern}}$  is an unphysical solution. It tells us that  $h_{\alpha\beta}^{\text{Bern}}$  describes the physical spacetime of a point mass  $\mu$ , with specific energy  $\mathcal{E}_0$ , on a circular geodesic orbit around a black hole with mass  $M_{\text{BH}} = M + \varepsilon M_{\text{Bern}}$ .

A minor drawback of the Berndtson solution is that it does not fit quite so cleanly into the evolution scheme in Sec. VI. If we take our total first-order  $\ell = 0$  solution to be  $h_{\alpha\beta}^{\text{Bern}} + x_{\alpha\beta}^{\ell=0}(M_1)$ , with  $x_{\alpha\beta}^{\ell=0}(M_1)$  as in Eq. (D25), then the total black hole mass is  $M_{\text{BH}} = M + \varepsilon(M_{\text{Bern}} + M_1)$ . Since  $dM_{\text{BH}}/d\tilde{s} = \varepsilon \dot{E}_H$ , this implies that the evolution equation for  $M_1$  becomes

$$\frac{dM_1}{d\tilde{s}} = \dot{E}_H - \frac{dM_{\text{Bern}}}{d\tilde{s}}, \quad (\text{D30})$$

where  $\frac{dM_{\text{Bern}}}{d\tilde{s}} = \frac{dr_0}{d\tilde{s}} \frac{dM_{\text{Bern}}}{dr_0}$ .

Of course, since  $x_{\alpha\beta}^{\ell=0}$  is irregular in any case, the total solution  $h_{\alpha\beta}^{\text{Bern}} + x_{\alpha\beta}^{\ell=0}$  is just as irregular as the solution  $h_{\alpha\beta}^{\text{pp},\ell=0} + x_{\alpha\beta}^{\ell=0}$  from the previous subsection. The only potential advantage it might have over the previous solution is if the changes in  $M_1$  were numerically negligible; in that case, we could simply neglect  $M_1$ , and  $h_{\alpha\beta}^{\text{Bern}}$  would provide a well-behaved first-order field at the boundaries, leading in turn to a second-order Einstein tensor  $\delta^2 G_{ilm}^0$  that is maximally well behaved at the boundaries. This may in fact be the case, since changes in the black hole parameters are typically extremely small over an inspiral [95].

A major advantage of the Berndtson solution lies in cases where we are not actually concerned with evolving the system but only with calculating some physical quantity at a fixed value of slow time. In that case we can freely omit  $x_{\alpha\beta}$ , and the Berndtson solution provides a first-order field that is regular at both boundaries, leading to a second-order source that is maximally well-behaved at the boundaries. This was the approach taken in Ref. [39].

As a final comment on the Berndtson solution, we note that if one leaves the Lorenz gauge, it provides a simple way to construct a solution that is continuous at  $r = r_0$ , regular at both boundaries, has zero mass for  $r < r_0$ , and has mass  $\mathcal{E}_0$  for  $r > r_0$ . To find this solution, start with the Berndtson solution and add an Eddington-Finkelstein solution (D4) with mass  $M_1 = -M_{\text{Bern}}$ .

*d. Asymptotically irregular, horizon-regular solution*

In most Lorenz-gauge calculations, neither the solution (D20) nor the Berndtson solution has been used. Instead, to our knowledge, all authors have adopted a solution for  $h_{\alpha\beta}^{\text{pp},\ell=0}$  that is regular at the horizon but not at infinity. Since this solution is displayed frequently, we do not repeat it here; it can be found in Eq. (114) of Dolan and Barack, for example. It satisfies the horizon-regularity conditions (162), but at infinity it goes to a constant:

$$\lim_{r \rightarrow \infty} h_{\alpha\beta}^{\text{pp},\ell=0} = 2M_{\text{Bern}} \delta_{\alpha}^t \delta_{\beta}^t. \quad (\text{D31})$$

Here  $M_{\text{Bern}}$  is *not* the mass in the solution; instead, it appears as a measure of the mismatch between intervals of coordinate time  $t$  in the perturbed spacetime and intervals of proper time for asymptotic observers.

A simple mass perturbation that is horizon-regular and minimally asymptotically irregular is given by

$$x_{\alpha\beta}^{\ell=0}(M_1, r) = 2M_1 h_{\alpha\beta}^{(A)} - M_1 h_{\alpha\beta}^{(B)}. \quad (\text{D32})$$

At large  $r$ , it behaves as

$$\lim_{r \rightarrow \infty} x_{\alpha\beta}^{\ell=0} = -2M_1 \delta_{\alpha}^t \delta_{\beta}^t. \quad (\text{D33})$$

Just as the solutions in the previous two sections made the second-order source ill behaved at the horizon, the solution  $h_{\alpha\beta}^{\text{pp},\ell=0} + x_{\alpha\beta}^{\ell=0}$  in this section makes the source ill behaved at infinity, decaying slowly as  $r \rightarrow \infty$ . That slow decay is numerically burdensome to integrate over, and it complicates the procedure of deriving boundary conditions.

This solution also requires special considerations in order to obtain useful results. In practice we are almost always interested in calculating quantities in gauges that admit the same preferred asymptotic reference frame as the background. As a consequence, in a typical first-order calculation, after computing a physical quantity in the asymptotically irregular Lorenz gauge, one performs a gauge transformation to find the value of that quantity

in an asymptotically flat gauge [15]. With the transformation written in the form  $x^{\alpha} \rightarrow x^{\alpha} - \varepsilon \xi^{\alpha}$ , it requires the linearly growing gauge vector

$$\xi^{\alpha} = \delta_{\alpha}^t M_{\text{Bern}} t, \quad (\text{D34})$$

which corresponds to a rescaling of time,

$$t \rightarrow (1 - \varepsilon M_{\text{Bern}})t \quad (\text{D35})$$

and also consequently a rescaling of the frequencies,

$$\omega_m \rightarrow (1 + \varepsilon M_{\text{Bern}})\omega_m. \quad (\text{D36})$$

In the first-order context, this transformation only affects the monopole mode of the metric perturbation, and it leaves that perturbation static. But at second order, such a transformation would affect every mode and lead to spurious, growing solutions in the second-order field, violating the presumed form of the two-timescale expansion.

We can recast the linearly growing transformation in a form consistent with the two-timescale form of the field by writing it as

$$\xi^t = \int M_{\text{Bern}}(\varepsilon t) dt = \frac{1}{\varepsilon} \int M_{\text{Bern}}(\tilde{t}) d\tilde{t}, \quad (\text{D37})$$

if reexpanded at fixed  $t$ , this reduces to the form above. It is a large,  $\mathcal{O}(1)$  transformation of  $t$ , but it is better understood as a small transformation of the slow time,

$$\tilde{t} \rightarrow \tilde{t} - \varepsilon \int M_{\text{Bern}}(\tilde{t}) d\tilde{t}. \quad (\text{D38})$$

Such a transformation filters through the entire two-timescale expansion, altering the frequencies  $\omega_m$  as above, the field equations in which  $\omega_m$  appears, and the fast time  $\phi_p = \int \Omega dt$ . It also requires additional care if we use  $\tilde{t}$  instead of  $t$  as our slow time.

Ref. [62] will discuss the gauge freedom in the two-timescale expansion in greater generality, including this freedom to transform the slow time.

- 
- [1] LIGO Scientific and Virgo Collaborations (B.P. Abbott et al.), Observation of gravitational waves from a binary black hole merger, *Phys. Rev. Lett.* **116**, 061102 (2016), arXiv:1602.03837 [gr-qc].
  - [2] LIGO Scientific and Virgo Collaborations (B.P. Abbott et al.), GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs, *Phys. Rev.* **X9**, 031040 (2019), arXiv:1811.12907 [astro-ph.HE].
  - [3] The catalog of LIGO/Virgo alerts are publicly available at <https://gracedb.ligo.org/superevents/public/O3/>.
  - [4] R. Abbott et al. (LIGO Scientific, Virgo), GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run, arXiv:2010.14527 [gr-qc] (2020).
  - [5] LIGO Scientific and Virgo Collaborations (B.P. Abbott et al.), *Astrophys. J.* **882**, L24 (2019), arXiv:1811.12940 [astro-ph.HE].
  - [6] LIGO Scientific and Virgo Collaborations (B.P. Abbott et al.), GW170817: Measurements of neutron star radii and equation of state, *Phys. Rev. Lett.* **121**, 161101 (2018), arXiv:1805.11581 [gr-qc].
  - [7] LIGO Scientific and Virgo Collaborations (B.P. Abbott et al.), Tests of general relativity with the binary black hole signals from the LIGO-Virgo Catalog GWTC-1, arXiv:1903.04467 [gr-qc] (2019).

- [8] R. Abbott *et al.* (LIGO Scientific, Virgo), Tests of General Relativity with Binary Black Holes from the second LIGO-Virgo Gravitational-Wave Transient Catalog, arXiv:2010.14529 [gr-qc] (2020).
- [9] The eLISA Consortium, The Gravitational Universe, arXiv:1305.5720 [astro-ph.CO] (2013).
- [10] S. Babak, J. Gair, A. Sesana, E. Barausse, C. F. Sopuerta, C. P. L. Berry, E. Berti, P. Amaro-Seoane, A. Petiteau, and A. Klein, Science with the space-based interferometer LISA. V: Extreme mass-ratio inspirals, Phys. Rev. **D95**, 103012 (2017), arXiv:1703.09722 [gr-qc].
- [11] L. Barack *et al.*, Black holes, gravitational waves and fundamental physics: a roadmap, Class. Quant. Grav. **36**, 143001 (2019), arXiv:1806.05195 [gr-qc].
- [12] E. Rosenthal, Construction of the second-order gravitational perturbations produced by a compact object, Phys. Rev. **D73**, 044034 (2006), arXiv:gr-qc/0602066 [gr-qc].
- [13] T. Hinderer and E. E. Flanagan, Two timescale analysis of extreme mass ratio inspirals in Kerr. I. Orbital Motion, Phys. Rev. **D78**, 064028 (2008), arXiv:0805.3337 [gr-qc].
- [14] E. E. Flanagan and T. Hinderer, Transient resonances in the inspirals of point particles into black holes, Phys. Rev. Lett. **109**, 071102 (2012), arXiv:1009.4923.
- [15] L. Barack and A. Pound, Self-force and radiation reaction in general relativity, Rept. Prog. Phys. **82**, 016904 (2019), arXiv:1805.10385.
- [16] S. Drasco, E. E. Flanagan, and S. A. Hughes, Computing inspirals in Kerr in the adiabatic regime. I. The Scalar case, Class. Quant. Grav. **22**, S801 (2005), arXiv:gr-qc/0505075.
- [17] N. Sago, T. Tanaka, W. Hikida, K. Ganz, and H. Nakano, The Adiabatic evolution of orbital parameters in the Kerr spacetime, Prog. Theor. Phys. **115**, 873 (2006), arXiv:gr-qc/0511151.
- [18] R. Fujita, W. Hikida, and H. Tagoshi, An Efficient Numerical Method for Computing Gravitational Waves Induced by a Particle Moving on Eccentric Inclined Orbits around a Kerr Black Hole, Prog. Theor. Phys. **121**, 843 (2009), arXiv:0904.3810.
- [19] S. Isoyama, R. Fujita, H. Nakano, N. Sago, and T. Tanaka, “Flux-balance formulae” for extreme mass-ratio inspirals, PTEP **2019**, 013E01 (2019), arXiv:1809.11118 [gr-qc].
- [20] S. Hughes, ‘Fairly fast’ adiabatic EMRI waveforms, talk given at the 22nd Capra Meeting on Radiation Reaction in General Relativity, CBPF, Rio de Janeiro, June 17–21, 2019. Slides available at [http://www.cbpf.br/~mcasals/Capra22\\_Talks/capra\\_fairlyfast\\_Hughes.pdf](http://www.cbpf.br/~mcasals/Capra22_Talks/capra_fairlyfast_Hughes.pdf).
- [21] A. J. K. Chua, C. J. Moore, and J. R. Gair, Augmented kludge waveforms for detecting extreme-mass-ratio inspirals, Phys. Rev. **D96**, 044005 (2017), arXiv:1705.04259 [gr-qc].
- [22] M. van de Meent, Gravitational self-force on generic bound geodesics in Kerr spacetime, Phys. Rev. D **97**, 104033 (2018), arXiv:1711.09607.
- [23] N. Warburton, T. Osburn, and C. R. Evans, Evolution of small-mass-ratio binaries with a spinning secondary, Phys. Rev. D **96**, 084057 (2017), arXiv:1708.03720.
- [24] V. Witzany, Spin-perturbed orbits near black holes, arXiv:1903.03649 [gr-qc] (2019).
- [25] S. Akcay, S. R. Dolan, C. Kavanagh, J. Moxon, N. Warburton, and B. Wardell, Dissipation in extreme-mass ratio binaries with a spinning secondary, arXiv:1912.09461 [gr-qc] (2019).
- [26] A. Pound, Second-order gravitational self-force, Phys. Rev. Lett. **109**, 051101 (2012), arXiv:1201.5089 [gr-qc].
- [27] E. Rosenthal, Second-order gravitational self-force, Phys. Rev. **D74**, 084018 (2006), arXiv:gr-qc/0609069 [gr-qc].
- [28] S. Detweiler, Gravitational radiation reaction and second order perturbation theory, Phys. Rev. D **85**, 044048 (2012), arXiv:1107.2098.
- [29] S. E. Gralla, Second order gravitational self force, Phys. Rev. **D85**, 124011 (2012), arXiv:1203.3189 [gr-qc].
- [30] A. Pound, Nonlinear gravitational self-force. I. Field outside a small body, Phys. Rev. **D86**, 084019 (2012), arXiv:1206.6538 [gr-qc].
- [31] A. Pound, Nonlinear gravitational self-force: second-order equation of motion, Phys. Rev. **D95**, 104056 (2017), arXiv:1703.02836 [gr-qc].
- [32] A. Pound and J. Miller, Practical, covariant puncture for second-order self-force calculations, Phys. Rev. **D89**, 104020 (2014), arXiv:1403.1843 [gr-qc].
- [33] N. Warburton and B. Wardell, Applying the effective-source approach to frequency-domain self-force calculations, Phys. Rev. D **89**, 044046 (2014), arXiv:1311.3104.
- [34] B. Wardell and N. Warburton, Applying the effective-source approach to frequency-domain self-force calculations: Lorenz-gauge gravitational perturbations, Phys. Rev. **D92**, 084019 (2015), arXiv:1505.07841 [gr-qc].
- [35] A. Pound, Second-order perturbation theory: problems on large scales, Phys. Rev. **D92**, 104047 (2015), arXiv:1510.05172 [gr-qc].
- [36] J. Miller, B. Wardell, and A. Pound, Second-order perturbation theory: the problem of infinite mode coupling, Phys. Rev. **D94**, 104018 (2016), arXiv:1608.06783 [gr-qc].
- [37] J. Miller, *The second-order gravitational self-force*, Ph.D. thesis, University of Southampton (2017).
- [38] A. Pound and B. Wardell, Black hole perturbation theory and gravitational self-force, arXiv:2101.04592 [gr-qc] (2021).
- [39] A. Pound, B. Wardell, N. Warburton, and J. Miller, Second-order self-force calculation of gravitational binding energy in compact binaries, Phys. Rev. Lett. **124**, 021101 (2020), arXiv:1908.07419 [gr-qc].
- [40] S. E. Gralla and R. M. Wald, A rigorous derivation of gravitational self-force, Class. Quant. Grav. **25**, 205009 (2008), [Erratum: Class. Quant. Grav., 28, 159501 (2011)], arXiv:0806.3293 [gr-qc].
- [41] A. Pound, Self-consistent gravitational self-force, Phys. Rev. **D81**, 024023 (2010), arXiv:0907.5197 [gr-qc].
- [42] P. Diener, I. Vega, B. Wardell, and S. Detweiler, Self-consistent orbital evolution of a particle around a Schwarzschild black hole, Phys. Rev. Lett. **108**, 191102 (2012), arXiv:1112.4821.
- [43] Y. Mino, Self-force in the radiation reaction formula, Prog. Theor. Phys. **113**, 733 (2005), arXiv:gr-qc/0506003 [gr-qc].
- [44] A. Pound and E. Poisson, Osculating orbits in Schwarzschild spacetime, with an application to extreme mass-ratio inspirals, Phys. Rev. D **77**, 044013 (2008), arXiv:0708.3033.
- [45] J. R. Gair, E. E. Flanagan, S. Drasco, T. Hinderer, and S. Babak, Forced motion near black holes, Phys. Rev. D **83**, 044037 (2011), arXiv:1012.5111.
- [46] N. Warburton, S. Akcay, L. Barack, J. R. Gair,

- and N. Sago, Evolution of inspiral orbits around a Schwarzschild black hole, *Phys. Rev. D* **85**, 061501 (2012), arXiv:1111.6908 [gr-qc].
- [47] T. Osburn, N. Warburton, and C. R. Evans, Highly eccentric inspirals into a black hole, *Phys. Rev. D* **93**, 064024 (2016), arXiv:1511.01498.
- [48] M. Van De Meent and N. Warburton, Fast Self-forced Inspirals, *Class. Quant. Grav.* **35**, 144003 (2018), arXiv:1802.05281 [gr-qc].
- [49] A. Pound, Motion of small objects in curved spacetimes: An introduction to gravitational self-force, *Proceedings, 524th WE-Heraeus-Seminar: Equations of Motion in Relativistic Gravity (EOM 2013): Bad Honnef, Germany, February 17-23, 2013*, *Fund. Theor. Phys.* **179**, 399 (2015), arXiv:1506.06245 [gr-qc].
- [50] J. K. Keckian and J. Cole, *Multiple scale and singular perturbation methods* (Springer-Verlag New York, 1996).
- [51] A. Pound and E. Poisson, Multi-scale analysis of the electromagnetic self-force in a weak gravitational field, *Phys. Rev. D* **77**, 044012 (2008), arXiv:0708.3037.
- [52] Y. Mino and R. Price, Two-timescale adiabatic expansion of a scalar field model, *Phys. Rev. D* **77**, 064001 (2008), arXiv:0801.0179 [gr-qc].
- [53] A. Pound, Singular perturbation techniques in the gravitational self-force problem, *Phys. Rev. D* **81**, 124009 (2010), arXiv:1003.3954 [gr-qc].
- [54] A. Pound, *Motion of small bodies in general relativity: foundations and implementations of the self-force*, Ph.D. thesis, University of Guelph (2010), arXiv:1006.3903.
- [55] H. Yang, B. Bonga, Z. Peng, and G. Li, Relativistic Mean Motion Resonance, arXiv:1910.07337 [gr-qc] (2019).
- [56] T. Damour, A. Gopakumar, and B. R. Iyer, Phasing of gravitational waves from inspiralling eccentric binaries, *Phys. Rev. D* **70**, 064028 (2004), arXiv:gr-qc/0404128 [gr-qc].
- [57] A. Klein, N. Cornish, and N. Yunes, Gravitational waveforms for precessing, quasicircular binaries via multiple scale analysis and uniform asymptotics: The near spin alignment case, *Phys. Rev. D* **88**, 124015 (2013), arXiv:1305.1932 [gr-qc].
- [58] C. M. Will and M. Maitra, Relativistic orbits around spinning supermassive black holes. Secular evolution to 4.5 post-Newtonian order, *Phys. Rev. D* **95**, 064003 (2017), arXiv:1611.06931 [gr-qc].
- [59] C. M. Will, Compact binary inspiral: Nature is perfectly happy with a circle, *Class. Quant. Grav.* **36**, 195013 (2019), arXiv:1906.08064 [gr-qc].
- [60] S. Akcay, N. Warburton, and L. Barack, Frequency-domain algorithm for the Lorenz-gauge gravitational self-force, *Phys. Rev. D* **88**, 104009 (2013), arXiv:1308.5223 [gr-qc].
- [61] T. Osburn, E. Forseth, C. R. Evans, and S. Hopper, Lorenz gauge gravitational self-force calculations of eccentric binaries using a frequency domain procedure, *Phys. Rev. D* **90**, 104031 (2014), arXiv:1409.4419.
- [62] E. Flanagan, T. Hinderer, J. Moxon, and A. Pound, The two-body problem in general relativity in the extreme-mass-ratio limit via multiscale expansions: Foundations (in preparation).
- [63] E. Flanagan, T. Hinderer, J. Miller, J. Moxon, and A. Pound, The two-body problem in general relativity in the extreme-mass-ratio limit via multiscale expansions: Dynamics of the strong-field region (in preparation).
- [64] E. Flanagan, T. Hinderer, J. Moxon, and A. Pound, The two-body problem in general relativity in the extreme-mass-ratio limit via multiscale expansions: Dynamics of the weak-field region (in preparation).
- [65] S. Isoyama, A. Pound, T. Tanaka, and K. Yamada, The two-body problem in general relativity in the extreme-mass-ratio limit via multiscale expansions: Dynamics of the near-horizon region (in preparation).
- [66] A. Zenginoğlu, A geometric framework for black hole perturbations, *Phys. Rev. D* **83**, 127502 (2011), arXiv:1102.2451 [gr-qc].
- [67] R. Panosso Macedo, Hyperboloidal framework for the Kerr spacetime, arXiv:1910.13452 [gr-qc] (2019).
- [68] S. Akcay, A fast frequency-domain algorithm for gravitational self-force: I. circular orbits in Schwarzschild spacetime, *Phys. Rev. D* **83**, 124026 (2011), arXiv:1012.5860 [gr-qc].
- [69] K. Martel and E. Poisson, *Phys. Rev. D* **71**, 104003 (2005), arXiv:0502028 [gr-qc].
- [70] J. Miller and A. Pound, Frequency-domain methods for first- and second-order self-force calculations in Schwarzschild spacetime (in preparation).
- [71] A. Pound, Tensor-harmonic decomposition of the second-order Einstein equation in Schwarzschild spacetime (in preparation).
- [72] J. Miller, A. Pound, N. Warburton, and B. Wardell, Second-order self-force calculations in Schwarzschild spacetime: calculation of the second-order source (in preparation).
- [73] A. Pound, Gauge and motion in perturbation theory, *Phys. Rev. D* **92**, 044021 (2015), arXiv:1506.02894.
- [74] L. Barack, D. A. Golbourn, and N. Sago, m-Mode Regularization Scheme for the Self Force in Kerr Spacetime, *Phys. Rev. D* **76**, 124036 (2007), arXiv:0709.4588.
- [75] I. Vega and S. L. Detweiler, Regularization of fields for self-force problems in curved spacetime: Foundations and a time-domain application, *Phys. Rev. D* **77**, 084008 (2008), arXiv:0712.4405.
- [76] S. Upton and A. Pound, Second-order self-force in a highly regular gauge, in preparation.
- [77] S. A. Teukolsky and W. H. Press, Perturbations of a rotating black hole. iii. interaction of the hole with gravitational and electromagnetic radiation, *Astrophys. J.* **193**, 443 (1974).
- [78] E. Poisson, Absorption of mass and angular momentum by a black hole: Time-domain formalisms for gravitational perturbations, and the small-hole / slow-motion approximation, *Phys. Rev. D* **70**, 084044 (2004), arXiv:gr-qc/0407050 [gr-qc].
- [79] A. Ori and K. S. Thorne, The transition from inspiral to plunge for a compact body in a circular equatorial orbit around a massive, spinning black hole, *Phys. Rev. D* **62**, 124022 (2000), arXiv:gr-qc/0003032 [gr-qc].
- [80] A. Pound, Conservative effect of the second-order gravitational self-force on quasicircular orbits in Schwarzschild spacetime, *Phys. Rev. D* **90**, 084039 (2014), arXiv:1404.1543 [gr-qc].
- [81] L. Barack and C. Lousto, Perturbations of Schwarzschild black holes in the Lorenz gauge: Formulation and numerical implementation, *Phys. Rev. D* **72**, 104026 (2005), arXiv:gr-qc/0510019 [gr-qc].
- [82] L. Barack and N. Sago, Gravitational self force on a particle in circular orbit around a Schwarzschild black hole, *Phys. Rev. D* **75**, 064021 (2007), arXiv:gr-qc/0701069 [gr-qc].

- [83] L. Blanchet and T. Damour, Radiative gravitational fields in general relativity I. General structure of the field outside the source, *Phil. Trans. R. Soc. Lond. A* **320**, 379 (1986).
- [84] L. Blanchet and T. Damour, Hereditary effects in gravitational radiation, *Phys. Rev. D* **46**, 4304 (1992).
- [85] D. Gal'tsov, Radiation reaction in the Kerr gravitational field, *J. Phys. A: Math. Gen.* **15**, 3737 (1982).
- [86] R. A. Isaacson, Gravitational radiation in the limit of high frequency. ii. nonlinear terms and the effective stress tensor, *Phys. Rev.* **166**, 1272 (1968).
- [87] J. Moxon, E. E. Flanagan, A. Pound, Z. Sam, and J. E. Thompson, Energy and angular momentum balance laws in second-order self-force theory, in preparation.
- [88] L. Blanchet, Gravitational radiation from post-Newtonian sources and inspiralling compact binaries, *Living Reviews in Relativity* **17**, 2 (2014).
- [89] A. Antonelli, M. van de Meent, A. Buonanno, J. Steinhoff, and J. Vines, Quasi-circular inspirals and plunges from non-spinning effective-one-body Hamiltonians with gravitational self-force information, [arXiv:1907.11597 \[gr-qc\]](https://arxiv.org/abs/1907.11597) (2019).
- [90] N. E. M. Rifat, S. E. Field, G. Khanna, and V. Varma, A Surrogate Model for Gravitational Wave Signals from Comparable- to Large- Mass-Ratio Black Hole Binaries, [arXiv:1910.10473 \[gr-qc\]](https://arxiv.org/abs/1910.10473) (2019).
- [91] A. J. K. Chua, C. R. Galley, and M. Vallisneri, Reduced-order modeling with artificial neurons for gravitational-wave inference, *Phys. Rev. Lett.* **122**, 211101 (2019), [arXiv:1811.05491 \[astro-ph.IM\]](https://arxiv.org/abs/1811.05491).
- [92] F. J. Zerilli, Gravitational field of a particle falling in a Schwarzschild geometry analyzed in tensor harmonics, *Phys. Rev.* **D2**, 2141 (1970).
- [93] S. R. Dolan and L. Barack, Self-force via  $m$ -mode regularization and 2+1D evolution: III. Gravitational field on Schwarzschild spacetime, *Phys. Rev. D* **87**, 084066 (2013), [arXiv:1211.4586](https://arxiv.org/abs/1211.4586).
- [94] M. V. Berndtson, *Harmonic gauge perturbations of the Schwarzschild metric*, Ph.D. thesis, University of Colorado (2007), [arXiv:0904.0033](https://arxiv.org/abs/0904.0033).
- [95] S. A. Hughes, Bound orbits of a slowly evolving black hole, *Phys. Rev. D* **100**, 064001 (2019), [arXiv:1806.09022 \[gr-qc\]](https://arxiv.org/abs/1806.09022).