Extreme mass-ratio inspirals in the effective-one-body approach: Quasicircular, equatorial orbits around a spinning black hole

Nicolás Yunes,[1,2,3] Alessandra Buonanno,[4] Scott A. Hughes,[2] Yi Pan,[4] Enrico Barausse,[4]
M. Coleman Miller,[5] and William Throwe[2]

[1]*Department of Physics, Princeton University, Princeton, New Jersey 08544, USA*
[2]*Department of Physics and MIT Kavli Institute, MIT,*
77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA
[3]*Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA*
[4]*Maryland Center for Fundamental Physics & Joint Space-Science Institute, Department of Physics,*
University of Maryland, College Park, Maryland 20742, USA
[5]*Maryland Astronomy Center for Theory and Computation & Joint Space-Science Institute, Department of Astronomy,*
University of Maryland, College Park, Maryland 20742, USA

(Received 29 September 2010; published 24 February 2011)

We construct effective-one-body waveform models suitable for data analysis with the Laser Interferometer Space Antenna for extreme mass-ratio inspirals in quasicircular, equatorial orbits about a spinning supermassive black hole. The accuracy of our model is established through comparisons against frequency-domain, Teukolsky-based waveforms in the radiative approximation. The calibration of eight high-order post-Newtonian parameters in the energy flux suffices to obtain a phase and fractional amplitude agreement of better than 1 rad and 1%, respectively, over a period between 2 and 6 months depending on the system considered. This agreement translates into matches higher than 97% over a period between 4 and 9 months, depending on the system. Better agreements can be obtained if a larger number of calibration parameters are included. Higher-order mass-ratio terms in the effective-one-body Hamiltonian and radiation reaction introduce phase corrections of at most 30 rad in a 1 yr evolution. These corrections are usually 1 order of magnitude larger than those introduced by the spin of the small object in a 1 yr evolution. These results suggest that the effective-one-body approach for extreme mass-ratio inspirals is a good compromise between accuracy and computational price for Laser Interferometer Space Antenna data-analysis purposes.

DOI: 10.1103/PhysRevD.83.044044 PACS numbers: 04.25.Nx, 04.30.Db, 04.30.−w, 04.30.Tv

I. INTRODUCTION

Extreme mass-ratio inspirals (EMRIs) are one of the most promising sources of gravitational waves (GWs) expected to be detected with the proposed Laser Interferometer Space Antenna (LISA) [1–4]. These sources consist of a small compact object, such as a neutron star or stellar-mass black hole (BH), in a close orbit around a spinning, supermassive BH [5]. Gravitational radiation losses cause the small object to spiral closer to the supermassive BH and eventually merge with it. Hence, the GW signal from such events encodes information about strong gravity, allowing tests of general relativity [6] and of the Kerr metric [7–17], as well as measurements of the spins and masses of massive BHs [18].

Unfortunately, EMRIs are very weak sources of GWs at their expected distances from us, and thus, they must be observed over many cycles to be detectable [5]. For example, a typical EMRI at a distance of 3 Gpc would produce GWs with signal-to-noise ratios (SNRs) on the order of 10–200 depending on the observation time. Therefore, matched filtering is essential to extract EMRIs from LISA noise and the foreground of unresolved GWs from white dwarf binaries in our galaxy.

Matched filtering consists of cross-correlating the data stream with a certain noise-weighted waveform template [19]. If the latter is similar to a GW event hidden in the data, then this cross-correlation filters it out of the noise. Of course, for matched filtering to be effective, one must construct accurate template filters. Otherwise, real events can be missed, or if an event is detected, parameter estimation can be strongly biased [20]. The construction of accurate EMRI waveforms is extremely difficult due to the long duration of the signal and the strong-field nature of the orbits. A one-year EMRI signal contains millions of radians in phase information. To avoid significant dephasing, its waveform modeling must be accurate to at least one part in 10^5–10^6 [21].

Such an exquisite accuracy requirement is complicated further by the strong-field nature of the orbit. An EMRI can reach orbital velocities of two-thirds the speed of light and orbital separations as small as a few times the mass of the supermassive companion. This automatically implies that standard, post-Newtonian (PN) Taylor-expanded waveforms fail to model such EMRI orbits [22]. PN theory relies on the assumptions that all orbital velocities are much smaller than the speed of light and that all objects are at separations much larger than the total mass of the

system [23]. A better approximation scheme to model EMRIs is BH perturbation theory, where one only assumes that the mass ratio of the system is much less than unity [24]. This is clearly the case for EMRIs, as the mass ratio is in the range 10^{-4}–10^{-6}. Perturbation theory, however, is computationally and analytically expensive. Only recently have generic orbits been computed around a nonspinning BH to linear order in the mass ratio [25–27], and it is unlikely that these will be directly used for EMRI data analysis [4].

EMRIs involve complicated inspiral analysis, but unlike comparable-mass coalescences, the merger and ringdown phase can be completely neglected. To see this, note that the instantaneous amplitude of the waves from a binary scales as μ, where $\mu = m_1 m_2/M$ is the reduced mass, M is the total mass and $m_{1,2}$ are the component masses. The inspiral lasts for a time $\sim 1/\mu$ and releases an energy flux $\sim \mu^2/\mu \sim \mu$. In contrast, the merger and ringdown last for a time $\sim M$ (independent of μ), and thus, release an energy flux $\sim \mu^2$. For an EMRI, $\mu \ll M$ and the inspiral clearly dominates the signal. Based on this argument, we neglect the merger and ringdown, focusing on the inspiral for our analysis.

A. Summary of previous work

The modeling of EMRIs has been attempted in the past with various degrees of success. One approach is to compute the self-field of the test particle to understand how it modifies the orbital trajectory. This task, however, is quite involved, both theoretically and computationally, as the self-field contains a divergent piece that is difficult to regularize (see, e.g., Ref. [28] for a recent review). Recently, a breakthrough was achieved, with the full calculation of the self-force for generic EMRIs about nonspinning supermassive BHs [25,26]. Such calculations, however, are computationally prohibitive if the goal is to populate a waveform template space.

Another approach is to use more approximate methods to model the EMRI trajectories. One such approach was developed by Hughes [29,30], following the pioneering work of Poisson [31]. In this *radiative-adiabatic* scheme, the inspiral is treated as a sequence of adiabatically shrinking geodesics. The degree of shrinkage is determined by solving the Teukolsky equation on each individual geodesic. Its solution encodes how the *constants* of the motion (the energy, angular momentum, and Carter constant) change due to GW emission. By interpolating across such sequence of geodesics, one then obtains a continuous inspiral and waveform. The calculation of a single waveform, however, is rather computationally expensive, as it requires the mapping of the entire orbital phase space, which for generic orbits is likely to be prohibitive. It is also worth noting that the radiative approximation neglects the impact of conservative effects which, especially for eccentric orbits, are likely to be important [32].

Other, perhaps more rough approximations can also be used to model EMRIs. The templates obtained through these methods are sometimes called *kludge* waveforms to emphasize their approximate nature. The goal of their construction was never to provide sufficiently accurate templates for real data analysis. Instead, kludge waveforms were built to carry out *descoping* or parameter estimation studies to determine roughly the accuracy to which parameters could be extracted, given an EMRI detection with LISA.

The first kludge waveforms were constructed by Barack and Cutler [18]. These waveforms employ the quadrupole formula to build templates as a function of the orbital trajectories. The latter are simply Keplerian ellipses with varying orbital elements. The variation of these is determined by low-order PN expressions, constructed from the GW energy and angular momentum fluxes. An improvement of these fluxes was developed by Gair and Glampedakis [33], who fitted these low-order PN expression to more accurate fluxes constructed from solutions to the Teukolsky equation. A further improvement was developed by Babak *et al.* [34], who modeled the waveforms via a quadrupole-octopole formula and the orbital trajectories via solutions to the geodesic equations, augmented with PN–orbit-averaged evolution equations for the orbital elements.

All of these improvements, however, do not mean that kludge waveforms would be effectual or faithful for realistic data analysis with LISA. One cannot exactly quantify this statement because exact EMRI waveforms are not available and will not be in the near future. One can nonetheless predict that these approaches will be insufficient because critical components of the fluxes are not being taken into account. For example, GWs do not only escape to infinity, but they are also absorbed by the supermassive BH, contributing to the overall fluxes of energy and angular momentum. This contribution is nonnegligible if one considers sufficiently long waveforms (longer than a few weeks). In fact, as we shall show in this paper, even the inclusion of such terms and very high-order PN expressions in the fluxes is still insufficient for accurate waveform models that last more than a couple of months.

B. The effective-one-body approach

The effective-one-body (EOB) formalism was introduced in Refs. [35,36] to model the inspiral, merger, and ringdown of comparable-mass BH binaries. This scheme was then extended to higher PN orders [37], spinning BHs [38–41], small mass-ratio mergers [42–44], and improved by resumming the radiation-reaction force and waveforms [43,45–48]. In the comparable-mass case, phase and amplitude agreement was achieved between EOB and numerical-relativity (NR) waveforms, after calibrating a few parameters [49–51]. By calibrating the EOB model to

the comparable-mass case, one can also improve the agreement of the model with the self-force predictions [25,52]. The combination of EOB and BH perturbation theory tools for LISA data-analysis purposes was first carried out in Refs. [53,54]. In these papers, the EOB scheme was found to be successful for the coherent modeling of EMRIs about a nonspinning background for a 2 yr period. Here we extend these results to nonprecessing EMRIs about a spinning background.

As a first step toward the construction of accurate EMRI waveforms, we concentrate on quasicircular, equatorial EMRIs about a spinning, supermassive Kerr BH. The modeling of such EMRIs is simpler than that of inclined and eccentric ones, as only a single component of the radiation-reaction force is nonvanishing and entirely controlled by the GW energy flux (the Carter constant vanishes by symmetry). Moreover, such EMRIs are expected in at least one astrophysical scenario [55]. In this setup, stellar-mass compact objects are either created in the accretion disk surrounding the supermassive BH or are captured by the disk, and hence move with the disk. The accretion disk is expected to be in the spin equatorial plane within a few hundred gravitational radii of the supermassive BH [56].

We first calibrate the EOB energy flux to the energy flux computed in BH perturbation theory through the solution to the Teukolsky equation. This calibration is done in the *test-particle* limit, i.e., by setting relative mass-ratio terms to zero in the EOB flux function. This calibration is more complicated than for nonspinning systems because it must now be performed globally, i.e., as a function of both spin and velocity. This increases the computational cost of the calibration and the number of calibration parameters, as a bivariate series generically contains more terms than a monovariate one. After calibrating 8 parameters, we find that the fluxes agree to within one part in 10^3 for all spins $[a/M = (-0.99, 0.99)]$ and velocities $[v = (0.01c, v_{\rm ISCO})]$ considered, where $v_{\rm ISCO}$ is the velocity at the innermost circular orbit (ISCO).

Once the energy flux has been calibrated, we evolve the Hamilton equations in the adiabatic approximation and compare the amplitude and phase evolution to that obtained with an approximate BH perturbation theory, numerical result. For the latter, we employ the so-called radiative approximation [29,30], where one models the EMRI as an adiabatic sequence of geodesics with varying orbital elements, as prescribed by the solution to the Teukolsky equation. We find that the EOB and Teukolsky-based waveforms agree in phase and relative amplitude to better than 1 rad and 1%, respectively, after 2 or 6 months of evolution, depending on the system considered. Better agreements can be obtained if a larger number of calibration parameters were included.

Our EOB waveforms differ from previous kludge models on several fronts. First, the radiation-reaction force is here computed differently than in the kludge approach. In the latter this force is calculated from PN, Taylor-expanded fluxes that encode the GWs that escape to infinity only. These fluxes were then improved by fitting a very large number of parameters to more accurate Teukolsky-fluxes with a log-independent, power-series expansion for the fitting functions [33]. In the EOB approach, the radiation-reaction force is computed directly from the factorized resummed waveforms [46,47]. These are enhanced through the addition of BH absorption terms and then the calibration of eight high PN-order parameters to Teukolsky fluxes with a log-dependent, power-series expansion for the fitting functions. Second, the conservative dynamics is also treated here differently than in the kludge approach. In the latter, the Hamiltonian is either a two-body, Newtonian one [18] or the full test-particle limit one, i.e., Schwarzschild or Kerr [34]. In the EOB approach, the conservative dynamics not only encodes the exact test-particle limit Hamiltonian, but also allows for the inclusion of finite mass-ratio terms and of the spin of the small body.

C. Data-analysis implications

The waveforms computed here are thus suitable for coherent data analysis over periods of several months. This can be established by computing the overlap between the EOB and Teukolsky-based waveforms, after maximizing over extrinsic parameters (an overall phase and time shift). We find that, when eight calibration parameters are used, the overlap remains higher than 97% over 4 to 9 months of evolution, depending on the system considered. This is to be compared with numerical kludge waveforms [34] whose overlap drops to 56% and 74% after 4 and 9 months, respectively, even when 45 calibration parameters are used to fit the flux [33]. Of course, one could obtain higher overlaps by maximizing over intrinsic parameters, such as the chirp mass or the spin of the background, but this would naturally bias parameter estimation. Also, when integrating over only two weeks, the overlap increases, remaining higher than 0.999 99 at 1 Gpc regardless of the model used.

The benefit of coherently integrating over longer periods of time is that the recovered SNR naturally increases, thus allowing us to detect signals farther out and improving parameter estimation. One can see this by simply noting that the SNR scales with the square root of the time of observation. For example, coherent integration over 4 or 9 months instead of two weeks increases the SNR at 1 Gpc from 2.8 to 14 and from 6 to 18 for two prototypical EMRIs. Such a large increase in SNR by coherently integrating over long observation times brings EMRIs not only to a confidently detectable range, but would also allow interesting tests of GR.

We conclude the paper by studying the error introduced in these waveforms due to neglecting second-order mass-ratio terms in the radiation-reaction force (dissipative PN self-force) and first-order in the mass-ratio terms in the

Hamiltonian (conservative PN self-force). Such mass-ratio dependent effects can easily be included in the EOB prescription, as they are known in the PN/EOB framework. Of course, since these are known to finite PN order, we cannot include full second-order effects. These effects should be considered estimates, since the complete result may differ from the PN prediction. We find that such PN radiation-reaction effects modify the phase of the waveform by $\mathcal{O}(10)$ radians in a 1 yr evolution, provided the EMRI samples the strong-gravity regime close to the ISCO. In a two-month period, however, the inclusion of finite mass-ratio effects increases the mismatch from 2.9×10^{-5} to 3.6×10^{-5} at 1 Gpc. This implies that such effects will only be seen if one coherently integrates over a sufficiently long time of observation. We find a somewhat smaller final dephasing when we allow the second body to be spinning and neglect any self-force corrections. The relative importance of the conservative or dissipative PN self-force terms and that of the spin of the second body depends somewhat on the EMRI considered. Generically, we find all such corrections to be larger than 1 rad after a full year of integration, while they are negligible over a two-month period.

This paper is organized as follows. Section II describes how we model EMRIs analytically and numerically. Section III discusses how the analytical EOB model is calibrated to the Teukolsky energy flux. Section IV compares EOB evolutions to Teukolsky ones, while Sec. V discusses the data-analysis implications of such a comparison. Section VI estimates the effect of mass-ratio dependent effects and Sec. VII concludes and points to future research. Appendix A presents details on the transformation between spheroidal and spherical tensor harmonics. Appendix B contains expressions for the GW energy flux absorbed by BHs. Finally, in Appendix C we write the EOB Hamiltonian derived in Ref. [41] when BHs carry spins aligned or antialigned with the orbital angular momentum. We use geometric units, with $G = c = 1$, unless otherwise noted.

II. EMRI MODELING

A. Analytical modeling: EOB-based waveforms

Consider a BH binary system with masses m_1 and m_2, total mass $M = m_1 + m_2$, reduced mass $\mu = m_1 m_2/M$, and symmetric mass ratio $\nu = \mu/M$. We assume that the orbital angular momentum is coaligned or counter-aligned with the individual BH spins $S_A = a_A m_A = q_A m_A^2$, where $a_A = S_A/m_A$ denotes the A-th BH's spin parameter and $q_A = a_A/m_A$ denotes the dimensionless spin parameter.

We first discuss the case of a nonspinning BH ($q_2 = 0$) with mass m_2 orbiting a spinning BH with spin parameter q_1 and mass $m_1 \gg m_2$, to leading order in the mass ratio m_2/m_1. Subleading terms in the mass ratio and terms proportional to q_2 introduce conservative corrections that are not included in the Teukolsky waveforms, which we shall use to calibrate our model, and we therefore neglect them during the calibration. Eventually, however, we shall turn these conservative terms on and estimate their effect using the spin EOB Hamiltonian of Ref. [41] (see Appendix C).

In the EOB framework, the orbital trajectories are obtained by solving Hamilton's equations, supplemented by a radiation-reaction force describing the backreaction of GW emission on the orbital dynamics. Neglecting conservative corrections of order $\mathcal{O}(m_2/m_1)$ and $\mathcal{O}(q_2)$, the spin EOB Hamiltonian reduces to the Hamiltonian of a nonspinning test-particle in Kerr, $H_{\rm NS}$:

$$H_{\rm EOB} = H_{\rm NS}[1 + \mathcal{O}(m_2/m_1) + \mathcal{O}(q_2)], \tag{1}$$

$$H_{\rm NS} = \beta^i p_i + \alpha\sqrt{m_2^2 + \gamma^{ij} p_i p_j}, \tag{2}$$

where

$$\alpha = \frac{1}{\sqrt{-g^{tt}}}, \tag{3}$$

$$\beta^i = \frac{g^{ti}}{g^{tt}}, \tag{4}$$

$$\gamma^{ij} = g^{ij} - \frac{g^{ti} g^{tj}}{g^{tt}}, \tag{5}$$

$g_{\mu\nu}$ being the Kerr metric. In Boyer-Lindquist coordinates (t, r, θ, ϕ) and restricting ourselves to the equatorial plane $\theta = 0$, the relevant metric components read

$$g^{tt} = -\frac{\Lambda}{r^2\Delta}, \tag{6a}$$

$$g^{rr} = \frac{\Delta}{r^2}, \tag{6b}$$

$$g^{\phi\phi} = \frac{1}{\Lambda}\left(-\frac{\omega_{\rm fd}^2}{r^2\Delta} + r^2\right), \tag{6c}$$

$$g^{t\phi} = -\frac{\omega_{\rm fd}}{r^2\Delta}, \tag{6d}$$

where $\omega_{\rm fd} = 2q_1 m_1^2 r$, and the metric potentials are

$$\Delta = r^2 + q_1^2 m_1^2 - 2m_1 r, \tag{7}$$

$$\Lambda = (r^2 + q_1^2 m_1^2)^2 - q_1^2 m_1^2 \Delta. \tag{8}$$

Although the EOB formalism includes possible nonadiabaticities in the last stages of the inspiral and plunge, it is not necessary to include nonadiabatic effects here. Generically, for the systems that we consider, we find that the inclusion of nonadiabatic corrections leads to small phase corrections [of $\mathcal{O}(1$ rad) after 1 yr of evolution] [53,54]. The assumption of adiabaticity allows us to simplify the evolution equations that are solved numerically.

This in turn reduces the computational cost of producing EOB waveforms: an adiabatic EOB evolution requires a few CPU seconds, while a nonadiabatic one would require CPU days or weeks (although this is estimated with a nonoptimized MATHEMATICA implementation). The nonadiabatic model is computationally more expensive because one needs to solve all of Hamilton's equations, with radiation-reaction source terms that are expensive to evaluate.

The Hamiltonian of Eq. (2) simplifies drastically when we consider circular, equatorial orbits ($\theta = \pi/2$) with $\mathbf{S}_1$ coaligned or counter-aligned with the orbital angular momentum (see, e.g., [57]). Imposing $p_r = 0$, which is a necessary condition for circular orbits, and inserting Eqs. (6a)–(6d) in Eq. (2), a straightforward calculation returns

$$H_{\rm NS} = p_\phi \frac{\omega_{\rm fd}}{\Lambda} + \frac{m_2 r\sqrt{\Delta}\sqrt{Q}}{\sqrt{\Lambda}}, \tag{9}$$

where

$$Q = 1 + \frac{p_\phi^2 r^2}{m_2^2 \Lambda}, \tag{10}$$

and $p_\phi \equiv L$ is the conjugate momentum to the ϕ Boyer-Lindquist coordinate or simply the orbital angular momentum. Imposing the condition $\dot{p}_r = (\partial H_{\rm NS}/\partial r)_{p_r=0} = 0$, which is also satisfied by circular orbits, we can solve for L as a function of r and q_1 [57]

$$L = \pm m_2 m_1^{1/2} \frac{r^2 \mp 2q_1 m_1^{3/2} r^{1/2} + q_1^2 m_1^2}{r^{3/4}(r^{3/2} - 3m_1 r^{1/2} \pm 2q_1 m_1^{3/2})^{1/2}}, \tag{11}$$

where $\pm$ corresponds to prograde or retrograde orbits, respectively. Inserting the above equation in Eq. (9) yields an expression for the energy $E \equiv H_{\rm NS}^{\rm circ}(r)$ of circular orbits in Kerr [57]

$$E = m_1 + m_2 \frac{1 - 2m_1/r \pm q_1 m_1^{3/2}/r^{3/2}}{\sqrt{1 - 3m_1/r \pm 2q_1 m_1^{3/2}/r^{3/2}}}. \tag{12}$$

The above quantities can also be expressed in terms of the orbital velocity ω, once $r(\omega)$ is derived for circular orbits. Computing $\omega = (\partial H_{\rm NS}/\partial p_\phi)_{p_r=0}$, using Eq. (11), we obtain

$$r = \frac{[1 - q_1(m_1\omega)]^{2/3}}{(m_1\omega)^{2/3}}. \tag{13}$$

We also define the parameter $v \equiv (m_1\omega)^{1/3}$.

In the adiabatic approximation, the orbital evolution is fully determined by the frequency evolution through Eq. (13). Assuming the motion follows an adiabatic sequence of quasicircular orbits, we can use the balance equation $\dot{L} = \dot{E}/\omega = -\mathcal{F}/\omega$ to derive

$$\dot{\omega} = -\frac{1}{\omega}\left(\frac{dL}{d\omega}\right)^{-1} \mathcal{F}(\omega), \tag{14}$$

where $\mathcal{F}$ is the GW energy flux (see, e.g., Ref. [36]). The multipolar-factorized form of this flux, proposed in the nonspinning case in Refs. [43,46] and extended to the spin case in Ref. [47], is given by

$$\mathcal{F}(\omega) \equiv \frac{1}{8\pi} \sum_{\ell=2}^{8} \sum_{m=0}^{\ell} |\dot{h}_{\ell m}|^2, \tag{15}$$

which under the assumption of adiabaticity reduces to

$$\mathcal{F}(\omega) = \frac{1}{8\pi} \sum_{\ell=2}^{8} \sum_{m=0}^{\ell} (m\omega)^2 |h_{\ell m}|^2, \tag{16}$$

with

$$h_{\ell m}(v) = h_{\ell m}^{{\rm Newt},\epsilon_p} S_{\ell m}^{\epsilon_p} T_{\ell m} e^{i\delta_{\ell m}} (\rho_{\ell m})^\ell, \tag{17}$$

where ϵ_p denotes the parity of the multipolar waveform (i.e., $\epsilon_p = 0$ if $\ell + m$ is even, $\epsilon_p = 1$ if $\ell + m$ is odd), and

$$h_{\ell m}^{{\rm Newt},\epsilon_p} \equiv \frac{m_1}{R} n_{\ell m}^{(\epsilon_p)} c_{\ell+\epsilon_p} v^{\ell+\epsilon_p} Y_{\ell-\epsilon_p,-m}(\pi/2, \phi). \tag{18}$$

When spin effects are present, the expressions for all the terms in Eq. (17), namely, $S_{\ell m}^{\epsilon_p}(v)$, $T_{\ell m}(v)$, $\delta_{\ell m}(v)$, and $\rho_{\ell m}(v)$ can be read in Ref. [47] [see Eqs. (24), (25), (26), and (29) therein]. The functions $Y_{\ell,m}(\theta, \phi)$ are the standard spherical harmonics, while $n_{\ell m}^{(\epsilon_p)}$ and $c_{\ell+\epsilon_p}$ are numerical coefficients that depend on the mass ratio [see Eqs. (5)–(7) in Ref. [46]].

The solution to Eq. (14) requires that we prescribe initial data. We here choose post-circular initial conditions, as described in Ref. [36], to set up a *mock* evolution that starts at a separation of $100m_1$ and ends at either the ISCO or whenever the GW frequency reaches 0.01 Hz. This mock evolution is then used to read initial data one year before the end of the mock evolution. This approach leads to an accurate initial data prescription, without any eccentricity contamination. For example, the error in the initial frequency induced by starting the mock evolution at $100m_1$, instead of $200m_1$, is on the order of 10^{-9} Hz, which leads to a difference in accumulated GW cycles of 0.03 rad after a 1 yr evolution.

Finite mass-ratio corrections can be incorporated into the EOB model by including subleading terms of $\mathcal{O}(m_2/m_1)$ and $\mathcal{O}(q_2)$ in the Hamiltonian, angular momentum, and $r(\omega)$ relation. We shall first ignore such terms to compare against Teukolsky-based waveforms. In Sec. VI, we shall study how our results change when we include such terms. To do so, we shall still assume circular, equatorial orbits and an adiabatic evolution, but employ the spin EOB Hamiltonian of Refs. [41,58] (reviewed in Appendix C), instead of the Kerr Hamiltonian of Eq. (9).

Except for this change, the EOB waveform modeling with finite mass-ratio corrections follows closely the derivation presented above. First, we compute the angular momentum associated with the Hamiltonian of Eq. (C22) for circular, equatorial orbits, imposing $\dot{p}_r = (\partial H_{\rm EOB}/\partial r)_{p_r=0} = 0$ and solving for $L \equiv p_\phi$. Then, we derive the orbital frequency $\omega = (\partial H_{\rm EOB}/\partial p_\phi)_{p_r=0}$ to relate r to ω, and to express L in terms of ω. When mass-ratio corrections are present, however, the Hamiltonian becomes much more involved, so solutions for L as a function of r must be searched numerically. We have checked that the discretization and interpolation used to solve these equations numerically do not introduce an error larger than 10^{-10} in the Hamiltonian (C22).

B. Numerical modeling: Teukolsky-based waveforms

Teukolsky-based waveforms is the name we give to radiative models that use the Teukolsky equation to prescribe the radiation-reaction force for an inspiral. We use BH perturbation theory, considering a background spacetime with mass m_1 and spin $|\mathbf{S}_1| = m_1^2 q$ (recall that the spin parameter $q_1 = a_1/m_1$). The inspiraling object is a test body with mass $m_2 \ll m_1$ and no spin ($q_2 = 0$). In principle, the masses and spins used here should be the same as those introduced in the EOB model.

The radiative approximation assumes that EMRIs can be modeled as an adiabatic sequence of geodesics with slowly varying constants of the motion. Consider the discretization of the orbital phase space, each point of which represents a certain geodesic with a given set of constants of motion (energy E, angular momentum L, and Carter constant Q). For quasicircular, equatorial EMRIs, the Carter constant vanishes, while the variation of the energy and angular momentum are related via $\delta L = (\delta E)/\omega$, ω being the orbital frequency. At each point in the orbital phase space, the geodesic equations can be solved to obtain the orbital trajectory of the small compact object, given any spin of the background.

Once the geodesic trajectories are known, one can use these to solve the linearized Einstein equations and obtain the gravitational metric perturbation. This is best accomplished by rewriting the linearized Einstein equations in terms of the Newman-Penrose curvature scalar ψ_4 to yield the Teukolsky equation [59]. One can decompose ψ_4 into spin-weight -2 spheroidal harmonics ${}_{-2}S^{a\omega}_{\ell m}(\theta)$, using Boyer-Lindquist coordinates (t, r, θ, ϕ), in the Fourier domain:

$$\psi_4 = \frac{1}{(r - i m_1 q_1 \cos\theta)^4} \int_{-\infty}^{\infty} d\omega \times \sum_{\ell m} R_{\ell m\omega}(r)\, {}_{-2}S^{m_1 q_1 \omega}_{\ell m}(\theta) e^{-i(m\phi - \omega t)}. \quad (19)$$

The radial functions $R_{\ell m\omega}(r)$ satisfy the radial Teukolsky equation

$$\Delta^2 \frac{d}{dr}\left(\frac{1}{\Delta}\frac{dR_{\ell m\omega}}{dr}\right) - V(r) R_{\ell m\omega} = -\mathcal{T}_{\ell m\omega}, \quad (20)$$

where Δ is given in Eq. (7) and the radial potential is

$$V(r) \equiv -\frac{K^2 + 4i(r - m_1)K}{\Delta} + 8i\omega r + \lambda, \quad (21)$$

with $K \equiv (r^2 + m_1^2 q_1^2)\omega - m m_1 q_1$, $\lambda \equiv \mathcal{E}_{\ell m} - 2 m_1 q_1 m\omega + m_1^2 q_1^2 \omega$, and $\mathcal{E}_{\ell m}$ the spheroidal harmonic eigenvalue. The source function $\mathcal{T}_{\ell m\omega}$ is given explicitly in Eq. (4.26) of Ref. [29] and it depends on the stress-energy tensor for a test particle in a geodesic trajectory.

The Teukolsky equation admits two asymptotic solutions: one outgoing as $r \to \infty$ and one ingoing as one approaches the background's event horizon. These two solutions represent outgoing radiation at future null infinity and ingoing radiation that falls into the BH through the event horizon. Both types of radiation are critical in the modeling of EMRIs; not including BH absorption can lead to errors in the waveform of order 10^4 radians [30,60]. These solutions can then be used to reconstruct both the GW radiated out to infinity, as well as the total energy flux lost in GWs. The energy flux can then be related to the temporal rate of change of the orbital elements, such as the orbital radius.

Solving the Teukolsky equation for a geodesic orbit tells us how that orbit tends to evolve due to the dissipative action of GW emission. By doing so for each point in orbital phase space, we endow this space with a set of vectors that indicate how the binary flows from one orbit to another. We compute these vectors at a large number of points, and use cubic-spline interpolation to estimate the rates of change of orbital constants between these points. This allows us to compute the temporal evolution of all relevant quantities, including the orbital trajectories and gravitational waveforms.

We implemented this algorithm, discretizing the orbital phase space from an initial separation of $r = 10^4 m_1$ to the Kerr ISCO

$$\frac{r_{\rm ISCO}}{m_1} = 3 + Z_2 \mp \sqrt{(3 - Z_1)(3 + Z_1 + 2Z_2)},$$
$$Z_1 = 1 + (1 - q_1^2)^{1/3}[(1 + q_1)^{1/3} + (1 - q_1)^{1/3}],$$
$$Z_2 = (3q_1^2 + Z_1^2)^{1/2}, \quad (22)$$

in a 1000 point grid, equally spaced in

$$v \equiv (m_1\omega)^{1/3} = \left[\frac{q_1 - r^{3/2}/m_1^{3/2}}{q_1^2 - r^3/m_1^3}\right]^{1/3}. \quad (23)$$

We cannot evolve inside the Kerr ISCO with such a frequency-domain code, as stable orbits do not exists in this regime (and so such orbits do not have a well-defined frequency spectrum). The code we use to construct our waves is based on [29,30,61], updated to use the spectral methods introduced by Fujita and Tagoshi

[62,63]. A detailed presentation of this code and its results is in preparation [64].

The dominant error in these Teukolsky-based waveforms is due to truncation of the sums over ℓ and m. In particular, to compare with PN results, we must map this spheroidal decomposition to a spherical one (see Appendix A for more details). Such a mapping requires one to include a buffer region of ℓ modes about the largest mode computed. We have been careful to use a sufficiently wide buffer and total number of ℓ modes such that the energy fluxes are accurate to 10^{-10} for all velocities and spins. In particular, this means that in the strong-field region (close to the ISCO) up to 50 ℓ modes were included. Other sources of error are due to the intrinsic double precision in the numerical solution to the Teukolsky equation, the discretization of the orbital phase space, and its cubic-spline interpolation. All of these amount to errors of order 10^{-14}. All in all and in terms of the GW phase, the Teukolsky-based waveforms are accurate to at least 10^{-2} rad during an entire year of evolution.

III. CALIBRATING THE TEST-PARTICLE ENERGY FLUX

We consider here a calibrated EOB model that is built from the $h_{\ell m}$ functions in Eq. (17), but in which higher-order PN terms are included in the functions $\rho_{\ell m}$ and are calibrated to the numerical results. In particular, we write

$$\begin{aligned}\rho^{22}_{\rm Cal} &= \rho^{22} + [a^{(9,1)}_{22} + b^{(9,1)}_{22}\text{eulerlog}_2 v^2]\bar{q}v^9 \\ &\quad + [a^{(12,0)}_{22} + b^{(12,0)}_{22}\text{eulerlog}_2 v^2]v^{12}, \\ \rho^{33}_{\rm Cal} &= \rho^{33} + [a^{(8,2)}_{33} + b^{(8,2)}_{33}\text{eulerlog}_3 v^2]\bar{q}^2 v^8 \\ &\quad + [a^{(10,0)}_{33} + b^{(10,0)}_{33}\text{eulerlog}_3 v^2]v^{10}, \end{aligned} \tag{24}$$

where $(a^{(N,M)}_{\ell m}, b^{(N,M)}_{\ell m})$ are eulerlog-independent and eulerlog-dependent calibration coefficients that enter the (ℓ, m) mode at $\mathcal{O}(v^N)$ and proportional to $\bar{q}^M$. As in Refs. [46,47] the eulerlog function is defined as

$$\text{eulerlog}_m(x) = \gamma_E + \log(2m\sqrt{x}), \tag{25}$$

where $\gamma_E = 0.577\,215\ldots$ is Euler's constant. Notice that we have introduced 4 calibration parameters in the nonspinning sector of the flux and 4 in the spinning sector. The spin parameter $\bar{q}$ denotes here the spin of the background. When we neglect mass-ratio terms, we choose $\bar{q} = q_1$. However, when we switch on the mass-ratio terms we have an ambiguity on the choice of $\bar{q}$. Following Ref. [47] we choose $\bar{q} = q$, where q is the deformed-Kerr spin parameter defined in Appendix C. Note that since $q = q_1 + \mathcal{O}(m_2/m_1)$, these choices are identical in the test-particle limit, which is when the energy flux is calibrated.

The choice of calibrating function in Eq. (24) is rather special and requires further discussion. We have chosen this function so that leading-order corrections in the two dominant GW modes, (2, 2) and (3, 3), are included. Higher (ℓ, m) modes contribute significantly less to the GW and its associated flux. In each mode, we have included the leading-order unknown terms that are both q independent and q dependent. Since q-independent terms in the energy flux are known to much higher order (5.5 PN) than the q-dependent ones (4 PN) in the test-particle limit [24,65,66], spin-independent calibration coefficients enter at a much higher PN order. The spin dependence of the calibration terms is inferred from known terms at lower PN orders. We have investigated many functional forms for the calibrating functions, with a varying number of degrees of freedom, and found the one above to be optimal in the class studied.

The introduction of 8 additional calibration terms might seem like a lot. In Ref. [53], the knowledge of nonspinning terms up to 5.5 PN order was found to be crucial to obtain a sufficiently good agreement in the flux. Moreover, only 4 additional calibration terms (2 at 6 PN order in ρ_{22} and 2 at 5 PN order in ρ_{33}) were needed to reach a phase agreement of 1 rad after two years of evolution. Similarly here, we expect that if the remaining 4.5, 5, and 5.5 PN-order terms were calculated in the test-particle limit when the central black carries a spin, then the flux would also improve, requiring a smaller number of calibration parameters. It is quite likely that those higher-order PN terms will be computed in the near future, as they involve dramatically less complicated calculations than PN terms in the comparable-mass case.

Having in hand an improved GW energy flux carried away to infinity, this must be enhanced with expressions for the GW energy flux that is absorbed by the background BH. We do so here by simply adding the Taylor-expanded form of the latter (see Appendix B) to the flux of Eq. (15). The BH absorption terms in the GW energy flux depend on polygamma functions, which are computationally expensive to evaluate. We have empirically found that expanding this function in $q \ll 1$ to 30th order is a sufficiently good approximation for our purposes. When performing computationally expensive calculations (like the fits described below) we shall employ such expansions, but when solving for the orbital phase and when computing the waveforms we shall return to the full polygamma expressions.

The resulting EOB energy flux, including BH absorption terms, is then calibrated via a two-dimensional, least-squares minimization relative to numerical data obtained from Teukolsky-based calculations. The fitting routine is two-dimensional because when considering EMRIs about spinning backgrounds, the flux depends on two independent variables: the orbital velocity (or frequency or separation) and the spin of the background. This, in turn, increases the number of points that need to be used by

more than an order of magnitude to properly calibrate Eq. (24). In all fits, we have assumed a data variance of 10^{-11} for all velocities and spins and we have required a relative accuracy of one part in 10^8. Since the data is now two-dimensional, one must search for a *global* minimum in (q, v) space. After doing so, we find the calibration parameters

$$a_{22}^{(9,1)} = -3.1092, \qquad b_{22}^{(9,1)} = -18.786, \tag{26}$$

$$a_{22}^{(12,0)} = 493.08, \qquad b_{22}^{(12,0)} = -247.89, \tag{27}$$

$$a_{33}^{(8,2)} = -17.310, \qquad b_{33}^{(8,2)} = 22.500, \tag{28}$$

$$a_{33}^{(10,0)} = -113.01, \qquad b_{33}^{(10,0)} = 28.125, \tag{29}$$

The computational cost of the calibrations performed in this paper is much larger than those carried out in Ref. [53] for the following reasons. First, we consider twice as many calibration parameters as in Ref. [53], increasing the dimensionality of the fitting space. Second, global minimization routines require nontrivial algorithms that are numerically more expensive than those employed in one-dimensional minimizations. Third, the amount of data fitted increases by at least 1 order of magnitude, due to the intrinsic bi-dimensionality of the problem. Combining all of this, the computational cost of performing the calibration is now more than 100 times larger than in Ref. [53]. Even then, however, these fits require $\mathcal{O}(10)$ CPU minutes to complete. Once they have been carried out, this calculation does not need to be repeated again in the waveform modeling.

Figure 1 plots the fractional difference between the analytical GW energy flux and that computed with Teukolsky-based waveforms as a function of velocity, from an initial value of $v/c = 0.01$ to the velocity at the ISCO, for five different spin values: $q = (-0.9, -0.5, 0.0, 0.5, 0.9)$. All comparisons are here normalized to the Newtonian value of the flux $F_{\rm Newt} = 32/5\nu^2 v^{10}$. The different curve styles differentiate between analytical models: the dotted curves use the total, uncalibrated Taylor expansion; the dashed curves use the uncalibrated ρ-resummed flux with BH absorption terms; the solid curves use the calibrated ρ-resummed flux with BH absorption terms. Notice that the calibrated model does better than the other two by at least 2 orders of magnitude near the ISCO for all spin values.

Several interesting conclusions can be drawn from Fig. 1. First, as obtained in Ref. [47] the uncalibrated ρ-resummed model is better than the Taylor-expanded version of the flux by up to nearly an order of magnitude at the ISCO for all spins. In turn, the calibrated model is better than the uncalibrated one by 1 to 2 orders of magnitude near ISCO for all spins. One could also calibrate the

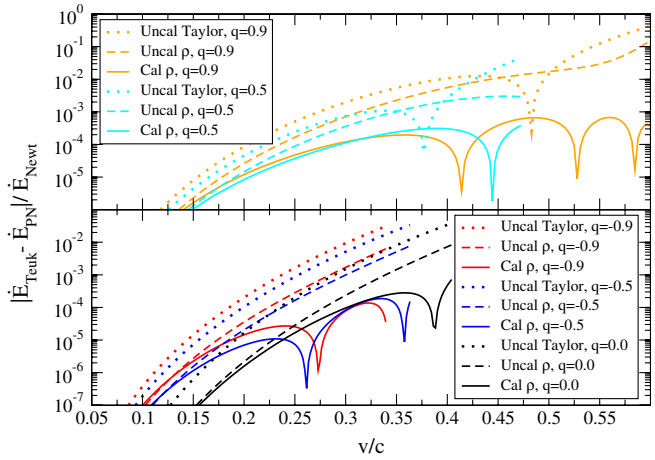

FIG. 1 (color online). Fractional difference between PN and Teukolsky-based fluxes as a function of velocity for spins $q = (0.5, 0.9)$ (top) and $q = (-0.9, -0.5, 0.0)$ (bottom). The dotted curves employ the Taylor-expanded PN flux with BH absorption terms, while the dashed and solid curves use the uncalibrated and calibrated ρ-resummed fluxes with BH absorption terms, respectively.

Taylor-expanded flux (not shown in Fig. 1), but this would not produce such good agreement in the entire (v, q) space. This is clearly because the uncalibrated ρ-resummed model is more accurate than the Taylor one, and thus the calibration terms have to do less work to improve the agreement. For the calibrated Taylor and ρ-resummed models to become comparable in accuracy one would have to include up to at least 16 calibration coefficients in the Taylor model.

The inclusion of BH absorption coefficients is crucial to obtain good agreement with the full Teukolsky-based flux, a result that was not obvious for the case of nonspinning EMRIs. Figure 2 plots the fractional difference between

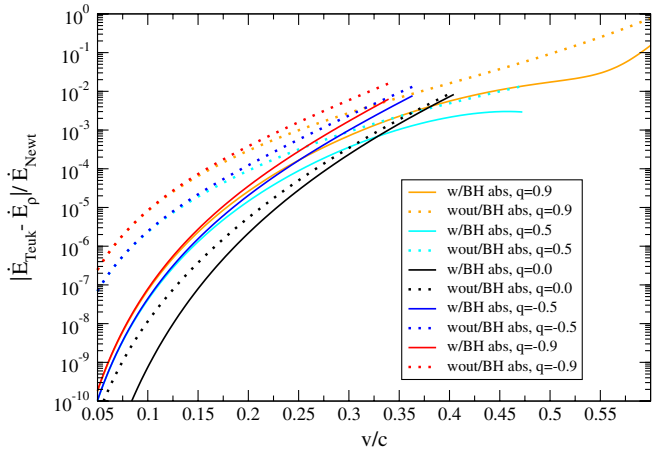

FIG. 2 (color online). Fractional difference between ρ-resummed and Teukolsky-based fluxes as a function of velocity for spins $q = (-0.9, -0.5, 0.0, 0.5, 0.9)$. The dotted curves do not include the Taylor-expanded BH absorption contributions, while the solid lines do.

the uncalibrated EOB GW energy flux and Teukolsky-based one as a function of velocity for five different spin values: $q = (-0.9, -0.5, 0.0, 0.5, 0.9)$, from an initial value of $v/c = 0.01$ at the ISCO, for five different spin values: $q = (-0.9, -0.5, 0.0, 0.5, 0.9)$. For these cases, we have $v_{\rm ISCO} = (0.343, 0.367, 0.408, 0.477, 0.609)$. The solid curves use the uncalibrated EOB model including the Taylor-expanded BH absorption contributions, while the dotted curves do not. For the nonspinning case, observe that there is a very small difference (smaller than 10^{-2}) between adding the BH absorption terms or not.

For the spinning cases, however, this is not the case. For rapidly spinning backgrounds, adding the BH absorption terms improves the agreement by an order of magnitude. Presumably, resumming these terms in a multipolar-factorized manner would improve the agreement even more. The BH absorption terms play a much larger role in the spinning case because spin changes the PN order at which absorption enters in the energy flux. These terms enter at 4 PN order for Schwarzschild black holes, but at 2.5 PN order for nonzero spin. This change of order has a very large and important impact on the system's evolution.

The inclusion of calibration parameters to improve the agreement of PN-inspired fluxes and Teukolsky-based ones for EMRIs is certainly not new. In Ref. [33], a similar, PN-inspired calibration was carried out for circular-inclined orbits (and more generic ones). Before calibration, their fluxes were Taylor expanded to 2 PN order and included only the contribution that escapes to infinity (not the BH absorption terms discussed above). Their fit was then done with Teukolsky data produced by an older version of the code used here, which was accurate only to one part in 10^6. Moreover, the fit was done in the range $r \in (5, 30)M$ [$v \in (0.183, 0.436)$], so the fitted function loses accuracy rapidly outside this regime, particularly close to the ISCO. Inside the fitting regime, the flux was fitted to an accuracy of 3×10^{-2} using 45 calibration coefficients for an inclined, but fixed orbit. The accuracy decreases to 0.1 for orbits which get closer to the ISCO.

To fairly compare the results of Ref. [33] with our results which are restricted to circular, equatorial orbits, we implemented their model and recalibrated it considering only circular, equatorial orbits. Using 45 coefficients, we found an accuracy similar to ours at high velocities close to the ISCO, but worse at low velocities. This is because their fluxes before calibration are not as accurate as the one employed here (by including up to 5.5 PN order terms and BH absorption terms), particularly at low velocities. It is important to emphasize that by calibrating 8 parameters instead of 45 we here obtain better flux accuracies than in Ref. [33] for circular, equatorial EMRIs. We could obtain even better accuracy if we were using a larger number of calibration coefficients, e.g., using 16 coefficients the agreement with the Teukolsky-based flux would be of $\mathcal{O}(10^{-5})$.

IV. COMPARISON OF THE GW PHASE AND AMPLITUDE

The comparison of EOB and Teukolsky evolutions requires that we choose a specific EMRI. We shall here follow Ref. [53] and choose system parameters that define two classes of EMRIs:

(i) *System I* explores a region between orbital separations $r/M \in (16, 26)$, which spans orbital velocities and GW frequencies in the range $v \in (0.2, 0.25)$ and $f_{\rm GW} \in (0.005, 0.01)$ Hz, respectively. Such an EMRI has masses $m_1 = 10^5 M_\odot$ and $m_2 = 10 M_\odot$ for a mass ratio of 10^{-4} and it inspirals for $\sim(6.3\text{–}6.7) \times 10^5$ rad of orbital phase depending on the spin.

(ii) *System II* explores a region between orbital separations $r/M \in (11, r_{\rm ISCO})$, which spans orbital velocities and GW frequencies in the range $v \in (0.3, v_{\rm ISCO})$ and $f_{\rm GW} \in (0.001, f_{\rm GW}^{\rm ISCO})$ Hz respectively. Such an EMRI has masses $m_1 = 10^6 M_\odot$ and $m_2 = 10 M_\odot$ for a mass ratio of 10^{-5} and it inspirals for $\sim(1.9\text{–}4.5) \times 10^5$ rad of orbital phase depending on the spin.

The evolution of Sys. I is stopped around an orbital separation of $16M$, because this coincides with a GW frequency of 0.01 Hz, which is close to the end of the LISA sensitivity band. The evolution of Sys. II is usually stopped at the orbital separation corresponding to the ISCO, or whenever its GWs reach a frequency of 0.01 Hz. For each of these systems, we shall investigate different background spin parameters.

Before proceeding, notice that Sys. I and II should not be compared on a one-to-one basis. One might be tempted to do so, as Sys. I resembles a weak-field EMRI, which inspirals at a larger orbital separation and with smaller orbital velocities than Sys. II, a more strong-field EMRI. Comparisons are not straightforward, however, as these systems accumulate a different total number of GW cycles. In fact, Sys. I usually accumulates almost twice as many GW cycles as Sys. II. Therefore, even though one might expect PN models of Sys. I to agree better with Teukolsky-based evolutions, this need not be the case, as this system has more time (as measured in GW cycles) to accumulate a phase and amplitude difference than Sys. II.

We compare the EOB and the Teukolsky-based waveforms after aligning them in time and phase. Such an alignment is done by minimizing the statistic in Eq. (23) of Ref. [50], just as was done in Ref. [53]. This is equivalent to maximizing the fitting factor over time and phase of coalescence in a matched filtering calculation with white noise [50]. The alignment is done in the low-frequency regime, inside the time interval $(0, 64)\lambda_{\rm GW}$, where $\lambda_{\rm GW}$ is the GW wavelength. This quantity depends on the spin of the background, ranging from $386M$ ($63M$) to $415M$ ($121M$) for Sys. I (Sys. II). This corresponds to aligning

the initial phase and frequency inside a window of length in the range (0.004, 0.01) months depending on the system and spin of the background. We have checked that increasing the size of the alignment window does not affect the final phase and amplitude difference; for example, for a spin of $q = 0.9$ and Sys. I, increasing the alignment window by a factor of 2 changes the final phase difference by 0.002 rad and the relative, fractional amplitude agreement by 0.0004%.

Figure 3 shows the absolute value of the dephasing and relative, fractional amplitude difference of the dominant $(\ell, m) = (2, 2)$ mode for both systems and a variety of background spins. For Sys. I, the calibrated EOB model maintains a 1 rad phase accuracy over at least the first 6 months for all spin values, while for Sys. II the same phase accuracy is maintained for up to only the first 2 months. The amplitude agreement is also excellent for all spin values, with better agreement for Sys. I. As found in Ref. [53] the GW phase agreement is primarily due to the correct modeling of the orbital phase, as the former tracks the latter extremely closely; we find that the difference between the orbital and GW phase over a 1 yr evolution is less than 0.1 rad.

The agreement in the phase as a function of background spin follows closely the flux agreement shown in Fig. 1. This is hard to see in Fig. 1, which is why Fig. 4 zooms into the velocity region sampled by Sys. I and plots all background spin cases for the calibrated ρ-resummed system. Observe that the $q = 0.0$ case has the best flux agreement, which explains why the phase and amplitude agreement is so good for this case in Fig. 3. Observe also that the $q = 0.9$ and $q = -0.9$ cases have the worst flux agreement, which also explains why they disagree the most in phase and amplitude in Fig. 3.

The accuracy of the calibrated EOB model is excellent relative to Taylor-expanded PN models. If one were to use an uncalibrated Taylor-expanded version of the flux, instead of the calibrated ρ-resummed flux, one would find a phase and amplitude disagreement of $\sim 10^1$–10^2 rad [$\sim 10^3$–10^4 rad] and $\sim 0.1\%$ ($\sim 10\%$) for Sys. I (Sys. II) after a 1 yr evolution for different spin values. The above results are consistent with the arguments in Ref. [22], who concluded that 3.5 PN accurate GW phase expressions could lead to phase errors around 10^3–10^4 radians over the last year of inspiral. That analysis reached those conclusions by comparing 3 to 3.5 PN accurate, analytic expressions for the GW phase. Here, we are comparing full-numerical evolutions of the PN equations of motion carried out to much higher order, and, of course, we find that such conclusions depend sensitively on the type of EMRI considered and the spin of the background.

The increase in accuracy of the calibrated ρ-resummed model is due both to the calibration and to the $h_{\ell m}$ factorized resummation. This fact can be appreciated in Fig. 5, where we plot the absolute value of the dephasing for the dominant mode in different PN models and the Teukolsky-based waveforms. Light curves (orange) correspond to background spins of $q = 0.9$, dark curves (red) to $q = -0.9$ and black curves to nonspinning backgrounds. Dotted curves use the uncalibrated Taylor flux model, dashed curves use the uncalibrated ρ-resummed model, and solid curves use the calibrated version. For Sys. I, there is a large gain in accuracy by switching from the uncalibrated Taylor model to the uncalibrated ρ-resummed model, but then the calibration itself does not appear to improve the accuracy substantially. For Sys. II, on the other hand, the calibration can increase the accuracy up to almost 2 orders of magnitude, as in the $q = 0.9$ case.

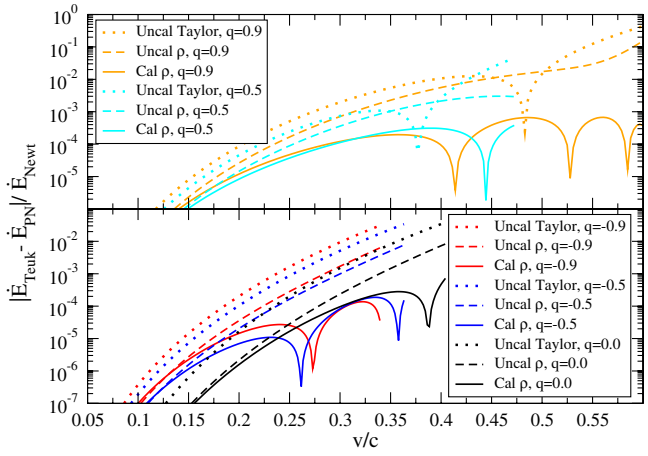

FIG. 3 (color online). Absolute value of the dephasing (left) and relative, fractional amplitude difference (right) computed in the calibrated EOB model and the Teukolsky-based waveforms for the dominant $(\ell, m) = (2, 2)$ mode. Different curves correspond to different background spin values.

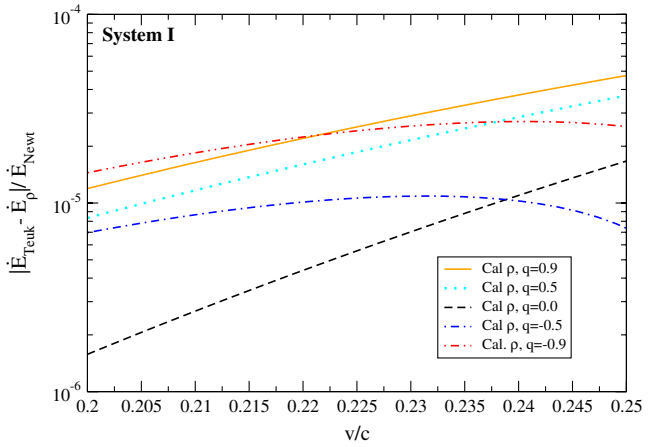

FIG. 4 (color online). Fractional difference between the calibrated ρ-resummed and Teukolsky-based fluxes for spins $q = (-0.9, -0.5, 0.0, 0.5, 0.9)$ as a function of velocity. We plot here only the range of velocities sampled by Sys. I.

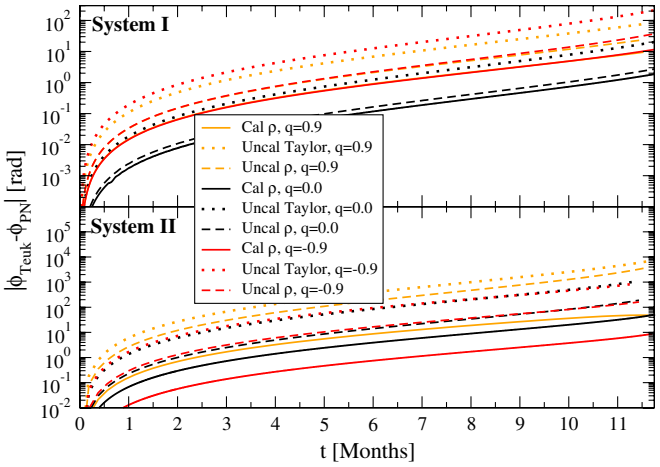

FIG. 5 (color online). Absolute value of the dephasing computed for the dominant mode in different PN models and the Teukolsky-based waveforms. Different curve colors/shades correspond to different background spin values, while different curve types correspond to different PN models.

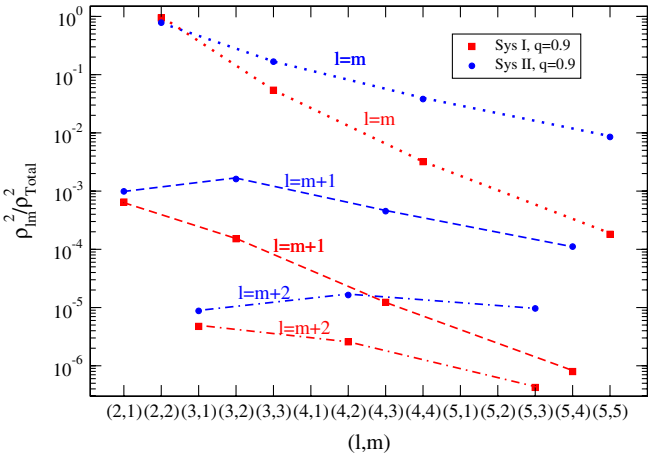

FIG. 7 (color online). Relative fraction between the squared of the SNR computed with only the (ℓ, m) mode and that computed with all (ℓ, m) modes. The lines connecting points are only meant to group points with $\ell = m$ (dotted line), $\ell = m + 1$ (dashed line), and $\ell = m + 2$ (dot-dashed line).

The agreement in the phase and amplitude is not only present in the dominant mode, but also in higher (ℓ, m) ones, as shown in Fig. 6. We here plot the absolute value of the dephasing and the relative, fractional amplitude difference between the calibrated EOB model and Teukolsky-based waveforms for the dominant mode, as well as the (3, 3) and (4, 4) modes. We have here shifted the higher (ℓ, m) modes using the best frequency shift that maximizes the agreement for the dominant mode. This agreement is simply a manifestation of the agreement in the orbital phase. In fact, we find that $\Phi_{\rm GW} \sim m\Phi_{\rm orb}$, with differences that are always less than 1 rad for all systems considered.

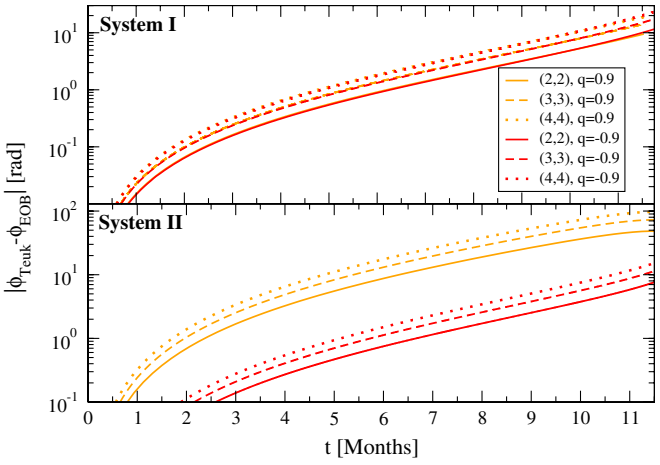

FIG. 6 (color online). Absolute value of the dephasing (left) and relative, fractional amplitude difference (right) computed in the calibrated EOB model and the Teukolsky-based waveforms. The solid curve corresponds to the dominant (2, 2) mode, while the dashed curve is for the (3, 3) mode and the dotted curve for the (4, 4) mode. Different curves stand for different background spins.

Higher-ℓ modes contribute significantly less to the SNR than the dominant (2, 2) mode. Figure 7 plots the relative fraction between the squared of the SNR computed with only the $h_{\ell m}$ component of the waveform and that computed by summing over all modes. This figure uses data corresponding to Systems I and II, both with spin $q = 0.9$ (results for other spin values are almost identical). Clearly, the (2, 2) mode is dominant, followed by the (3, 3) and (4, 4) modes. Because of this feature of quasicircular inspirals, obtaining agreement for the (2, 2) mode implies one can recover over 97% of the SNR.

V. DATA-ANALYSIS IMPLICATIONS

Although the phase agreement presented in the previous section is a good indicator of the validity of the EOB model, one is really interested in computing more realistic data-analysis measures. In this section we compute the mismatch between the Teukolsky and the EOB model, maximized over extrinsic parameters and as a function of observation time.

Let us first introduce some basic terminology. Given any time series $a(t)$ and $b(t)$, we can define the following inner product in signal space:

$$(a|b) = 4\,\mathrm{Re}\int_0^\infty \frac{\tilde{a}(f)\tilde{b}^\star(f)}{S_n(f)}, \tag{30}$$

where the overhead tildes stand for the Fourier transform and the star stands for complex conjugation. The quantity $S_n(f)$ is the spectral noise density curve, where here we follow [67,68]. Notice that we use the *sky-averaged* version of this noise curve here, which is larger than the

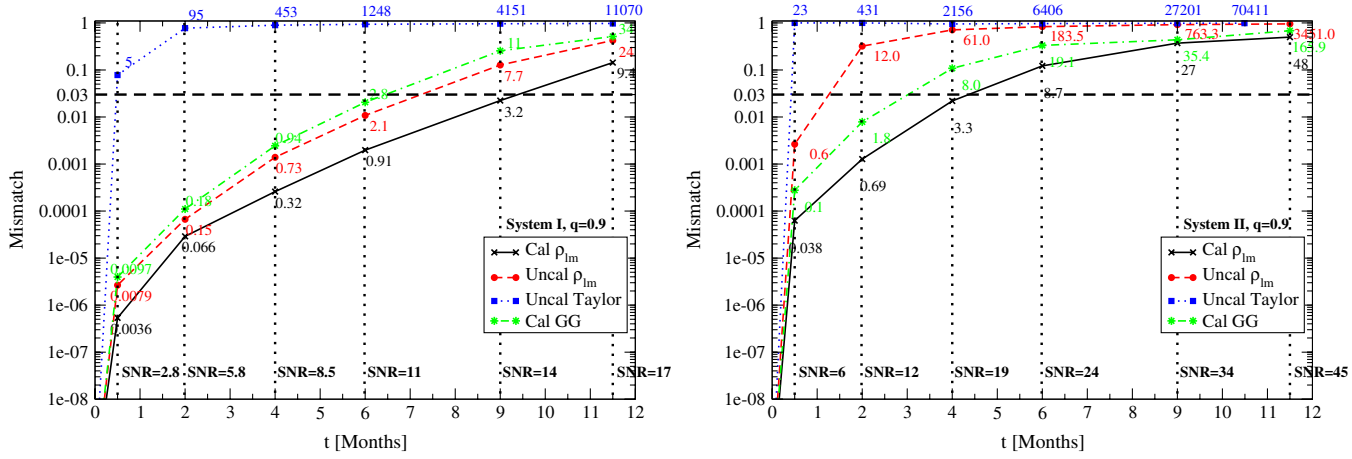

FIG. 8 (color online). Mismatch between Teukolsky and EOB waveforms as a function of waveform duration for Systems I and II and a background spin of $q = 0.9$.

non–sky-averaged version by a factor of $20/3$. In particular, this means that our SNRs are smaller than those one would obtain with a non–sky-averaged noise curve by a factor of $(20/3)^{1/2} \sim 2.6$. Also notice that we compute the SNRs from the sky-averaged Fourier transform of the response function, and not from the Fourier transform of the (2, 2) mode.

Given this inner product, we can now define some useful measures. The SNR of signal a is simply

$$\rho = \sqrt{(a|a)}, \tag{31}$$

while the overlap between signals a and b is simply

$$\mathrm{M} = \max \frac{(a|b)}{\sqrt{(a|a)(b|b)}}, \tag{32}$$

with the mismatch $\mathrm{MM} = 1 - \mathrm{M}$. The max label here is to remind us that this statistic must be maximized over a time shift and a phase shift (see, e.g., Appendix B of [45] for a more detailed discussion).

The data-analysis measures introduced above (ρ and MM) depend on the length of the time-series, i.e., the observation time. Figure 8 plots the mismatch between the Teukolsky-based waveforms and a variety of models for both Sys. I and II and a background spin of $q = 0.9$ as a function of observation time. The vertical lines correspond to observation times of 2 weeks, 2 months, 6 months, 9 months, and 11.5 months, together with their associated SNRs at 1 Gpc. The mismatches are computed with different analytical models: black crosses stand for the calibrated ρ model with 8 calibration coefficients; red circles to the uncalibrated ρ model; blue squares to the uncalibrated Taylor model; green circles to that EOB evolution using the original flux of Ref. [33] which has 45 calibration coefficients (denoted *GG* in the figure). For comparison, we also include the amount of dephasing (numbers next to data points in Fig. 8) at 2 weeks, 2 months, 6 months, 9 months, and 11.5 months. Observe that the calibrated ρ model maintains an overlap higher than 97% over 9 and 4 months for Sys. I and II, respectively. The uncalibrated ρ model performs comparably to the EOB model using the flux of Ref. [33] which has 45 calibration coefficients, both of which have an overlap higher than 97% over 6 and 1 month for Sys. I and II, respectively. In the case of Sys. II the calibrated flux of Ref. [33] performs better than the uncalibrated $\rho_{\ell m}$ model. Also observe that the uncalibrated Taylor model is simply inadequate to model EMRIs for any observation time.

We then see that the use of the calibrated EOB model allows us to integrate over longer observation times, compared to a 2-week period or to other models. In turn, this allows us to recover a higher SNR than we would otherwise. The increase in SNR scales as the square root of the observation time, as expected. For example, since the calibrated EOB model is accurate over 9 months, one would be able to coherently recover an SNR of 14 at 1 Gpc for Sys. I, to be compared with an SNR of 11 obtained after 6 months of coherent integration if using, for example, the fluxes of Ref. [33]. In general, an integration over a period longer than two weeks gains us a large increase in SNR. Such gains in SNR are important because they allow us to see EMRIs farther out. Since the SNR scales as $\rho \sim D_L^{-1}$, where D_L is the luminosity distance, even an SNR increase in a factor of 3 increases our accessible volume by a factor of 27, since the latter scales as D_L^3.

VI. HIGHER-ORDER EFFECTS

Let us now discuss how finite mass-ratio higher-order effects affect the GW phase and amplitude. Those effects are encoded in the ν terms present in the radiation-reaction force and in the Hamiltonian. The former are second-order effects in the dissipative dynamics, while the latter are first-order effects in the conservative dynamics. We have

analytic control over the PN version of such effects within the EOB formalism, but until now we had set both of these to zero when comparing to Teukolsky evolutions, as the latter do not account for such effects. In the EOB model, however, it is straightforward to include such terms, as PN expansion are formally known for all mass ratios at some given order in v. The inclusion of conservative ν terms is achieved by using the spin EOB Hamiltonian of Ref. [41] reviewed in Appendix C. The inclusion of dissipative ν terms is achieved by including relative ν terms in the multipolarly decomposed waveform $h_{\ell m}$ and flux $\mathcal{F}$.

Whether such mass-ratio effects matter depends on the EMRI considered. In Ref. [53], it was found that such effects increase the dephasing between EOB and Teukolsky models by 1 order of magnitude after a 2 yr evolution for nonspinning EMRIs, a result consistent with Ref. [69]. This effect is greatly amplified when considering spinning EMRIs. Table I compares the effect that the inclusion of relative ν terms in the EOB Hamiltonian and the radiation-reaction force has on the final dephasing after a 1 yr evolution. In order to read out the effect of such finite mass-ratio terms in the phasing, one must compare rows two, three, and four to the baseline given in the first row of Table I. For example, the effect of the ν terms in the Hamiltonian are such as to increase the dephasing by $27.20 - 10.04 = 17.16$ radians. Observe that the conservative and dissipative ν terms usually push the dephasing in different directions, partially canceling out when both of them are present. Even then though, the generic effect of high-order ν terms is to increase the rate of dephasing by several tens of radians are a 1 yr evolution. Notice furthermore that the magnitude of the effect is here not very large because we are considering circular equatorial orbits.

Finite mass-ratio effects are clearly suppressed when dealing with circular orbits. This is because, for such orbits outside the ISCO, the effect of the conservative self-force is simply to shift the waveform from one orbital frequency to another. Thus, from an observational standpoint, such an effect is unmeasurable. Even though the conservative self-force shifts the ISCO, this effect is still degenerate with a shift of the system's mass parameters. This discussion, however, neglects radiation reaction, which is crucial to model a true inspiral waveform. One can think of the radiation-reaction force as defining a trajectory through the sequence of orbital energies that an orbit follows. There is gauge-invariant information in this sequence, in the sense that the mapping between energies and orbital frequencies depends on the details of the orbit at each energy level. For a given radiation-reaction force, the sequence of geodesic orbits (and hence the sequence of frequencies) depends on whether the conservative self-force is included or not.

TABLE I. Absolute value of the total dephasing after 11.5 months of evolution. The first row includes no relative ν contributions in the Hamiltonian or the radiation-reaction force. The second row includes relative ν terms in the Hamiltonian, while the third row includes such terms in the radiation-reaction force.

System	I	I	I	II	II	II
q_1	-0.9	0.0	0.9	-0.9	0.0	0.9
No rel. ν	10.04	1.60	9.36	7.63	42.21	48.39
ν in H	30.38	0.083	18.96	4.38	40.35	47.13
ν in $\dot{E}$	19.99	5.31	4.49	7.24	41.50	46.55
ν in H and $\dot{E}$	40.32	6.83	14.08	3.98	39.64	45.29

As is clear from this discussion, the real (gauge-invariant) effect of the conservative self-force on quasicircular inspiral waveforms can be rather subtle. A robust effect, however, does arise if the self-force acts on a more generic orbit, such as an eccentric one. In that case, this force will act separately on the radial and the azimuthal orbital frequencies, which can leave a potentially strong imprint in the waveform. In principle, even for an inclined circular orbit there could be a strong imprint. In practice, however, the azimuthal and polar orbital frequencies are quite similar, which suggests that perhaps, even in this case, the self-force effects will be small.

With all of this in mind, let us discuss the results presented in Table I in more detail. Our study suggests that the overall effect of ν terms in both H and $\mathcal{F}$ leads to only 5.2 and 2.5 additional radians of phase for nonspinning, Sys. I and II, respectively. This is in fact consistent with the results presented in Ref. [53], except that there one considered 2 yr long evolutions. One might wonder whether using the nonspinning Hamiltonian of Ref. [50] (where the deformed-Schwarzschild potential are Padè resummed instead of being given by Eqs. (C18) and (C8)) has an effect on this dephasing. We have investigated this question and found that the additional contribution to the phase is 0.05 (0.03) and 1.06 (1.33) rad for Sys. I and II, respectively, over the entire year of inspiral using the 3 PN (4 PN) Padè form of the deformed potentials. (We notice [70] that the 4PN Padè potentials of Ref. [50] reproduce very closely the ISCO shift of Ref. [25].) This implies that the nonspinning Hamiltonian [50] at 3 PN and 4 PN order is sufficiently close to the Hamiltonian presented in Appendix B for the data analysis of nonspinning EMRIs.

One can also compare the results in Table I to the recent study of Huerta and Gair [71], who investigated the effect of ν^2 corrections in the determination of parameters, given an EMRI signal. Their Table I presents the number of cycles accumulated for a variety of mass ratios. Their last column happens to correspond to our Sys. II with no spin, for which they get a total dephasing of 2.3 rad and 3.8 rad after the last year of inspiral when including only conservative and all second-order corrections. This is to be compared to our results: 1.86 rad in and 2.6 rad after the last year of inspiral when including only conservative and all second-order corrections. These numbers are in excellent agreement, allowing for differences in the modeling.

TABLE II. Absolute value of the total dephasing in rad after 11.5 months of evolution. The second row sets the second BH's spin to zero, while the second row sets it to $q_2 = 0.75$. In both cases, the spin of the background is set to $q_1 = 0.9$ and $\nu \neq 0$ in both the Hamiltonian and the radiation-reaction force.

System	I	I	I	II	II	II
q_1	-0.9	0.0	0.9	-0.9	0.0	0.9
$q_2 = -0.75$	36.34	3.96	16.03	4.41	40.41	47.04
$q_2 = 0.00$	40.32	6.83	14.08	3.98	39.64	45.29
$q_2 = 0.75$	44.31	9.70	12.14	3.56	38.88	43.54

Their analysis suggests that such small differences will not affect parameter estimation for EMRIs similar to Sys II. For Sys I, however, the dephasing is much larger as the mass ratio is less extreme by 1 order of magnitude; thus, parameter estimation might be affected in this case.

Another high-order effect that one can study is the inclusion of the small object's spin in the evolution of the binary. Since the spin angular momentum of the small body scales with its mass, its contribution to the orbital evolution is one order in ν suppressed. We can model this effect by allowing $q_2 \neq 0$ in the EOB Hamiltonian of Ref. [41] and letting $\nu \neq 0$. Doing so for Sys. I (Sys. II) with $q_1 = 0.9$ and $q_2 = 0.75$, we find that the total dephasing now becomes 17.26 (45.49 rad), as shown in Table II. This is to be compared to the case when $q_2 = 0$, which returns a dephasing of 14.08 rad (45.29 rad). Thus, the effect of the second spin contributes roughly 3.3 rad in this case. We can also compare these results to those estimated by Barack and Cutler [18]. In their Appendix C, they estimate that for quasicircular inspirals similar to our Sys. II, the spin of the second body should induce a dephasing of roughly 1–10 rad. This is in good agreement with the results corresponding to Sys. II in our Table II. Our results are also in good agreement with the results of Wen-Biao [72], who estimated a dephasing of $\mathcal{O}(\pi/10)$ for our System II. Finally, these results are also in agreement with an upcoming and independent investigation of spin effects in the PN phasing [73].

Finally, taking into account the results of including finite mass-ratio effects, we can conclude that unless these are precisely modeled, it is not worth requiring an agreement better than 10–30 rad when calibrating the phase of the test-particle EOB waveforms against the Teukolsky-based waveforms.

VII. CONCLUSIONS

We have constructed an EOB model for EMRIs in quasicircular, equatorial orbits about spinning backgrounds. In the test-particle limit, this model consists of adiabatically evolving a test particle in the Kerr spacetime using the factorized energy flux of Refs. [46,47], augmented by 8 calibration coefficients. The latter are determined by comparing the factorized energy flux to a Teukolsky-based flux, built from solutions to the Teukolsky equation in the radiative approximation. In the adiabatic approximation, the EOB waveforms can be constructed in CPU seconds at a very low computational cost. When finite mass-ratio effects and the small object's spin are included, we build the EOB model by numerically solving the Hamilton equations with the spin EOB Hamiltonian of Ref. [18] and the Teukolsky-calibrated factorized energy flux augmented by finite mass-ratio effects [47].

For both EMRI systems considered, we find excellent phase and amplitude agreement, with dephasing less than 1 rad, over periods of months. The exact length of the agreement depends on how relativistic the EMRI system is. We also calculated the overlap between EOB and Teukolsky-based waveforms to find it higher than 97% over 4 to 9 months, depending on the EMRI system considered.

The EOB waveforms built here have higher overlaps and better phase agreements than all currently known EMRI models for spinning, equatorial systems, while requiring much fewer calibration parameters. In particular, the EOB model with 8 calibration coefficients outperforms by almost an order of magnitude the numerical kludge waveforms with the calibrated fluxes of Ref. [33] which use 45 calibration coefficients. This implies that EOB waveforms with 8 calibration coefficients can be used for longer coherent integrations, allowing us to obtain a 50% increase in SNR. In turn, this implies that EOB waveforms can see EMRIs that are farther out, increasing the accessible volume by at least a factor of 2 relative to numerical kludge waveforms. Furthermore, if we were using 16 calibration coefficients, we could improve the dephasing from 0.91 rad (8.7 rad) to 0.85 rad (4.2 rad) for System I (System II) after 6 months of evolution. In turn, this would decrease the mismatch from 0.2% (12%) to 0.19% (3.9%) for System I and II after 6 months of evolution.

Another possible avenue for future research is the calculation of high PN-order terms in the energy flux and waveforms. Our EOB model relies on the use of accurate fluxes, but for spinning systems, the flux to infinity is only known up to 4 PN order in the test-particle limit. This is in contrast to the nonspinning terms that are known to 5.5 PN order or the BH absorption terms that are known to 6.5 PN order. The calculation of the spin-dependent terms in the flux to infinity to 4.5, 5, and 5.5 PN order terms in the test-particle limit is not quixotic and would be invaluable. Once these coefficients are known, then presumably the EOB waveforms would be more accurate and might require less adjustable parameters.

Of course, the EOB waveforms we constructed here are less powerful than kludge waveforms [18,33,34] in their generality. Our waveforms cannot yet model inclined or eccentric inspirals. The inclusion of inclined orbit should be relatively straightforward, but the addition of

eccentricity might require some revamping of the EOB framework. Future work in this direction would be definitely worthwhile.

Ultimately, one would like to obtain a waveform model that is sufficiently fast, efficient, and accurate to do realistic EMRI data analysis for LISA. Such a model would need to include the correct self-force contributions to the conservative dynamics, the appropriate second-order terms in the radiation-reaction force and the correct terms that describe the small object's spin. The EOB model we developed here does agree with all known PN self-force calculation to date. However, since full self-force calculations are not yet completed, we do not have a way of assessing the error of including the currently known EOB finite ratio effects. In fact, so far the only comparisons between the PN/EOB and self-force results have been concerned with the nonspinning case, and have been limited to the ISCO shift [25,52] and other gauge-invariant quantities [74,75]. Quite interestingly, the calibration of the EOB model to comparable-mass numerical-relativity simulations improves the agreement of the model to self-force results [52,70]. Thus, we hope that future calibrations of the spin EOB Hamiltonian to comparable-mass simulations of spinning BHs will allow us to build an EOB model which includes finite mass-ratio effects in a more accurate way.

All that said, we have found that the presence of PN self-force terms in the spin EOB model of Ref. [41] leads to dephasing of 10–30 rad over 1 yr depending on the EMRI system and the spin of the background. By contrast, the inclusion of the small object's spin introduces dephasing on the order of a few radians. Taking into account those results, we can conclude that unless those finite ratio effects are precisely modeled, currently, it is not worth requiring an agreement better than 10–30 rad when calibrating the phase of the test-particle EOB waveforms against the Teukolsky-based waveforms. This in turn implies that calibrating 16 parameters instead of 8 to an EOB model is overkill as other systematics (induced by neglecting self-force effects) will be dominant.

ACKNOWLEDGMENTS

We are grateful to Leor Barack, Carlos Sopuerta and Frans Pretorius for useful suggestions and comments. N. Y., A. B., E. B., Y. P., and S. A. H. acknowledge support from the NSF Grants No. PHY-0745779, No. PHY-0903631, and No. PHY-0449884; A. B., S. A. H., and M. C. M. also acknowledge support from NASA Grants No. NNX09AI81G, No. NNX08AL42G, and No. NNX08AH29G. N. Y. acknowledges support from the National Aeronautics and Space Administration through the Einstein Postdoctoral Fellowship Award No. PF0-110080 issued by the Chandra X-ray Observatory Center, which is operated by the Smithsonian Astrophysical Observatory for and on behalf of the National Aeronautics Space Administration under Contract No. NAS8-03060.

APPENDIX A: MAPPING BETWEEN SPHEROIDAL AND SPHERICAL HARMONICS

1. Quantities at spatial infinity

Let us first consider the GW fluxes that describe radiation escaping to infinity, and later discuss the radiation absorbed by the BH. Frequency-domain Teukolsky equation codes expand the curvature scalar ψ_4 as

$$\psi_4 = \frac{1}{r}\sum_{\ell m} Z^H_{\ell m} S^-_{\ell m}(\theta, \phi) e^{-i\omega_m t}. \tag{A1}$$

We have here incorporated the $e^{im\phi}$ dependence into the spin-weighted spheroidal harmonics $S^-_{\ell m}$. These harmonics depend on the value of $m_1 q_1 \omega_m$, and the minus superscript is a reminder that we consider here harmonics of spin-weight -2. Throughout this appendix, the index ℓ refers to the spheroidal harmonic index, while l refers to the spherical harmonic index.

From ψ_4, we compute waveforms via $\psi_4 = (\ddot{h}_+ - i\ddot{h}_\times)/2$, and hence, for a frequency-domain application,

$$h_+ - ih_\times = -\frac{2}{r}\sum_{\ell m} \frac{Z^H_{\ell m}}{\omega_m^2} S^-_{\ell m}(\theta, \phi) e^{-i\omega_m t}. \tag{A2}$$

We have implicitly assumed that ω_m is time independent, or at least that its time dependence is subleading. A better expansion is to reexpress things in terms of the accumulated phase, i.e., the integral of the frequency $\omega_m \equiv \dot{\Phi}_m$, namely,

$$\psi_4 = \frac{1}{r}\sum_{\ell m} Z^H_{\ell m} S^-_{\ell m}(\theta, \phi) e^{-i\Phi_m}, \tag{A3}$$

$$h \equiv h_+ - ih_\times = -\frac{2}{r}\sum_{\ell m} \frac{Z^H_{\ell m}}{\omega_m^2} S^-_{\ell m}(\theta, \phi) e^{-i\Phi_m}. \tag{A4}$$

The EOB and numerical-relativity communities like to project these quantities onto a basis of spin-weighted spherical harmonics. For ψ_4, they define $C_{lm}(t, r)$ via

$$\psi_4 = \frac{1}{r}\sum_{lm} C_{lm}(t, r) Y^-_{lm}(\theta, \phi), \tag{A5}$$

and the harmonically decomposed waveforms $h_{lm}(t, r)$ via

$$h = \frac{1}{r}\sum_{lm} h_{lm}(t, r) Y^-_{lm}(\theta, \phi). \tag{A6}$$

The minus superscript again denotes spin-weight -2. Defining the inner product

$$\langle Y^-_{lm} | f \rangle = \int d\Omega Y^{-,*}_{lm}(\theta, \phi) f, \tag{A7}$$

the extraction of the C_{lm} and h_{lm} is simple:

$$C_{lm}(t, r) = r\langle Y^-_{lm} | \psi_4 \rangle, \qquad h_{lm}(t, r) = r\langle Y^-_{lm} | h \rangle. \tag{A8}$$

Let us now take the Schwarzschild limit to see whether these expressions simplify. In that limit

$$S^{-}_{\ell m}(\theta, \phi) = Y^{-}_{\ell m}(\theta, \phi), \tag{A9}$$

and thus, performing the necessary projections and taking advantage of the orthonormality of spherical harmonics, we find

$$C_{lm}(t, r) = Z^{H}_{lm} e^{-i\Phi_m}; \tag{A10}$$

$$h_{lm}(t, r) = -\frac{2Z^{H}_{lm}}{\omega_m^2} e^{-i\Phi_m}. \tag{A11}$$

In spinning backgrounds, however, the mapping is more complicated. We can use the fact that spheroidal harmonics can be expressed as a sum of spherical harmonics via

$$S^{-}_{\ell m}(\theta, \phi) = \sum_j b^{\ell}_j Y^{-}_{jm}(\theta, \phi). \tag{A12}$$

The dependence on $m_1 q_1 \omega_m$ enters through the expansion coefficients b^{ℓ}_j (see, e.g., Ref. [60]). Inserting this expansion into Eq. (A1) and using the inner-product definition, Eq. (A8), we find

$$C_{jm} = e^{-i\Phi_m} \sum_{\ell} b^{\ell}_j Z^{H}_{\ell m}; \tag{A13}$$

$$h_{jm} = -\frac{2e^{-i\Phi_m}}{\omega_m^2} \sum_{\ell} b^{\ell}_j Z^{H}_{\ell m} = -\frac{2C_{jm}}{\omega_m^2}. \tag{A14}$$

In the Schwarzschild limit, $b^{\ell}_j = \delta_{\ell j}$, so that Kerr simply limits as it should.

From the definition of the Isaacson stress-energy tensor, one can easily show that

$$\frac{d^2E^{\infty}}{dtd\Omega} = \sum_{\ell m} S^{-}_{\ell m}(\theta, \phi)^2 \frac{|Z^{H}_{\ell m}|^2}{4\pi\omega_m^2}, \tag{A15}$$

$$= \sum_{lm} Y^{-}_{lm}(\theta, \phi)^2 \frac{|C^{H}_{lm}|^2}{4\pi\omega_m^2}. \tag{A16}$$

We have here used the orthonormality of both the spheroidal and spherical harmonics to simplify the sums, as well as the fact that the time dependence of the C^{H}_{lm} disappears when its modulus is computed. Performing the angular integrals leaves us with familiar formulas:

$$\dot{E}^{\infty} = \sum_{\ell m} \frac{|Z^{H}_{\ell m}|^2}{4\pi\omega_m^2} = \sum_{lm} \frac{|C^{H}_{lm}|^2}{4\pi\omega_m^2}. \tag{A17}$$

This breaks down nicely enough that it is useful and sensible to define the modal contributions $\dot{E}_{\ell m}$.

2. Quantities at event horizons

As a general principle, computing quantities that are related to an event horizon are usually more complicated than computing the same quantities at spatial infinity. For the fluxes, for example, this is because there is no simple generalization of the Isaacson tensor on the horizon. Instead, one must examine the shear of the horizon's generators, look at how this shear generates entropy, and then apply the area theorem to compute fluxes [76]. The relevant quantities at the horizon depend on the Newman-Penrose scalar ψ_0, instead of ψ_4, the former of which is a quantity of spin-weight $+2$, rather than -2.

The GW energy flux per unit solid angle at the horizon is given by

$$\frac{d^2E^{H}}{dtd\Omega} = \frac{\omega_m m_1 r_+}{2\pi p_m} |\sigma^{\rm HH}|^2, \tag{A18}$$

where $p_m = \omega_m - m\omega_+$, and where $\omega_+ = q_1/(2r_+)$ is the angular velocity of observers corotating with the event horizon. The quantity $\sigma^{\rm HH}$ is the shear to the horizon's generators as found by Ref. [77]. This quantity is fairly simply related to ψ_0, so let us introduce an expansion of ψ_0 in spin-weight $+2$ spheroidal harmonics

$$\psi_0 = \Delta^{-2} \sum_{\ell m} W^{\infty}_{\ell m}(r) S^{+}_{lm}(\theta, \phi) e^{-i\Phi_m}, \tag{A19}$$

where Δ is given in Eq. (7). Notice that this quantity diverges on the event horizon r_+ because the Kinnersley tetrad, which is used to define the projection for ψ_0, is ill-behaved as $r \to r_+$. This can be corrected for by converting to $\sigma^{\rm HH}$ for any given (ℓ, m) mode [76]

$$\sigma^{\rm HH}_{\ell m} = \Delta^2 \gamma_m \psi_{0,\ell m} = \sum_{\ell m} \gamma_m W^{\infty}_{\ell m} S^{+}_{lm}(\theta, \phi) e^{-i\Phi_m}. \tag{A20}$$

The complex number γ_m is given by $\gamma_m = -[4(ip_m + 2\epsilon)(2m_1 r_+)^2]^{-1}$, where $\epsilon = \sqrt{m_1^2 - m_1^2 q_1^2}/(4m_1 r_+)$.

With this in hand, the GW energy flux at the horizon becomes

$$\frac{d^2E^{H}}{dtd\Omega} = \sum_{\ell m} \frac{\omega_m m_1 r_+}{2\pi p_m} |\gamma_m|^2 |W^{\infty}_{\ell m}|^2 (S^{+}_{\ell m})^2,$$
$$= \sum_{\ell m} (S^{+}_{\ell m})^2 \frac{\omega_m^3}{16 p_m (p_m^2 + 4\epsilon^2)(2m_1 r_+)^3} \frac{|W^{\infty}_{\ell m}|^2}{4\pi\omega_m^2}, \tag{A21}$$

which integrates to

$$\dot{E}^{H} = \sum_{\ell m} \frac{\omega_m^3}{16 p_m (p_m^2 + 4\epsilon^2)(2m_1 r_+)^3} \frac{|W^{\infty}_{lm}|^2}{4\pi\omega_m^2}. \tag{A22}$$

Implementing this equation is difficult because it requires that one computes both ψ_0 and ψ_4 when solving the Teukolsky equation. Since these quantities have different angular dependence and a different source function, this would be a nontrivial undertaking.

Instead, one can take advantage of a remarkable simplification, the so-called Starobinsky identities [78], that relates ψ_4 to ψ_0 and vice-versa via an algebraic relation.

We can use this to relate the coefficients $W^{\infty}_{\ell m}$ to the coefficients $Z^{\infty}_{\ell m}$, namely,

$$W^{\infty}_{\ell m} = \beta_{\ell m} Z^{\infty}_{\ell m}, \tag{A23}$$

where

$$\begin{aligned}
\beta_{\ell m} &= \frac{64(2m_1 r_+)^4 i p_m (p_m^2 + 4\epsilon^2)(-i p_m + 4\epsilon)}{c_{\ell m}} Z^{\infty}_{\ell m}, \\
|c_{\ell m}|^2 &= [(\lambda+2)^2 + 4q_1 m_1 \omega_m - 4q_1^2 m_1^2 \omega_m^2] \\
&\quad \times (\lambda^2 + 36 m q_1 m_1 \omega_m - 36 q_1^2 m_1^2 \omega_m^2) \\
&\quad + (2\lambda+3)(96 q_1^2 m_1^2 \omega_m^2 - 48 m q_1 m_1 \omega_m) \\
&\quad + 144 \omega_m^2 m_1^2 (1 - q_1^2), \\
\mathrm{Im}\, c_{\ell m} &= 12 m_1 \omega, \quad \mathrm{Re}\, c_{\ell m} = +\sqrt{|c_{\ell m}|^2 - 144 m_1^2 \omega^2}, \\
\lambda &= \mathcal{E}_{\ell m} - 2 q_1 m_1 m \omega + q_1^2 m_1^2 \omega^2 - s(s+1),
\end{aligned} \tag{A24}$$

and $\mathcal{E}_{\ell m}$ is the eigenvalue of $S^{-}_{\ell m}(\theta, \phi)$.

We can then finally write the energy flux formula as

$$\dot{E}^{H} = \sum_{\ell m} \alpha_{\ell m} \frac{|Z^{\infty}_{\ell m}|^2}{4\pi \omega_m^2}, \tag{A25}$$

where the coefficient $\alpha_{\ell m}$, given explicitly in Ref. [60], agglomerates the factors β, γ into one big expression. This expression reorganizes terms slightly in order for its structure to resemble the expression for the flux to infinity as much as possible. Using this, it is simple to write the modal contribution as decomposed into a spheroidal harmonic basis: $\dot{E}^{H}_{\ell m} = \alpha_{\ell m} |Z^{\infty}_{\ell m}|^2 / (4\pi \omega_m^2)$.

Decomposing the horizon energy flux formula into a spherical harmonic basis is slightly more difficult. The key confusing issue is that there are now two spherical harmonic bases to worry about: one for each spin weight. With respect to these two bases, we can define two spheroidal harmonic expansions:

$$S^{-}_{\ell m}(\theta, \phi) = \sum_j b^{\ell}_j Y^{-}_{jm}(\theta, \phi), \tag{A26}$$

$$S^{+}_{\ell m}(\theta, \phi) = \sum_j d^{\ell}_j Y^{+}_{jm}(\theta, \phi). \tag{A27}$$

(The coefficients d^{ℓ}_j expand the $+2$ spin-weight spheroidal harmonic in $+2$ spin-weight spherical harmonics, just as the coefficients b^{ℓ}_j do so for the -2 harmonics.) It is worth emphasizing that the different spherical harmonics are not simply related to one another.

The two quantities which can be put into a spherical harmonic basis are ψ_0 and ψ_4, both evaluated in the vicinity of the horizon:

$$\psi_0 = \Delta^{-2} \sum_{\ell j m} W^{\infty}_{\ell m} d^{\ell}_j Y^{+}_{jm}(\theta, \phi) e^{-i\Phi_m}, \tag{A28}$$

$$= \Delta^{-2} \sum_{jm} U^{\infty}_{jm} Y^{+}_{jm}(\theta, \phi), \tag{A29}$$

$$U^{\infty}_{jm} = \sum_{\ell} W^{\infty}_{\ell m} d^{\ell}_j e^{-i\Phi_m}; \tag{A30}$$

and

$$\psi_4 = \frac{\Delta^2}{(r - i q_1 m_1 \cos\theta)^4} \sum_{\ell j m} Z^{\infty}_{\ell m} b^{\ell}_j Y^{-}_{jm}(\theta, \phi) e^{-i\Phi_m}, \tag{A31}$$

$$= \frac{\Delta^2}{(r - i q_1 m_1 \cos\theta)^4} \sum_{jm} C^{\infty}_{jm} Y^{-}_{jm}(\theta, \phi), \tag{A32}$$

$$C^{\infty}_{jm} = \sum_{\ell} Z^{\infty}_{\ell m} b^{\ell}_j e^{-i\Phi_m}. \tag{A33}$$

Using the results presented in this appendix, one can easily find an expression for $\dot{E}^{H}$ in terms of the $+2$ harmonic coefficients U^{∞}_{lm}:

$$\dot{E}^{H} = \sum_{lm} \frac{\omega_m^3}{16 p_m (p_m^2 + 4\epsilon^2)(2m_1 r_+)^3} \frac{|U^{\infty}_{lm}|^2}{4\pi \omega_m^2}. \tag{A34}$$

Unfortunately, this is not that useful, as it requires knowledge of the ψ_0 expansion coefficients $W^{\infty}_{\ell m}$.

Taking advantage of the Starobinsky identity again, we can combine Eqs. (A23) and (A30) to find

$$U^{\infty}_{jm} = \sum_{\ell} \beta_{\ell m} Z^{\infty}_{\ell m} d^{\ell}_j e^{-i\Phi_m}. \tag{A35}$$

It is then straightforward to insert this into Eq. (A34) to obtain the down-horizon flux expanded into modes of $+2$ spherical harmonics.

3. Truncation issues

We have so far been rather schematic regarding the limits on all sums. In principle, all these sums should be carried out from some lower limit $l_{\min}$ to infinity, where the former is given by $l_{\min} = \min(|s|, |m|)$. In a numerical application, the upper limit must be truncated at some finite value $l_{\max}$. We typically find that the magnitude of terms falls off as a power of l. When decomposing into spheroidal harmonics, it is thus typically sufficient to pick some cutoff value and apply it uniformly.

Applying such a cutoff is slightly more complicated when we convert to spherical harmonics. The reason is that a given spheroidal harmonic ℓ has contributions from spherical harmonics at index $j > \ell$. Consider, as a concrete example, the spheroidal harmonic S^{-}_{54} for $a/m_1 = 0.99$, $\omega = 0.1$: the expansion coefficients for this harmonic are

$$b^5_3 = -0.011\,065\,7, \qquad b^5_4 = 0.999\,87,$$
$$b^5_5 = 0.011\,736\,8, \qquad b^5_6 = 0.000\,123\,221,$$
$$b^5_7 = 9.4336 \times 10^{-7}, \qquad b^5_8 = 6.4276 \times 10^{-8},$$
$$b^5_9 = 3.70511 \times 10^{-11}, \qquad b^5_{10} = 1.933\,17 \times 10^{-13},$$
$$b^5_{11} = 8.975\,58 \times 10^{-16}. \tag{A36}$$

Coefficients beyond b^5_{11} are small enough that our code does not compute them in this case. Notice that as we move away from the $j = \ell$ term (whose value is close to unity) the coefficients fall off by roughly powers of $\epsilon \simeq 0.01$. This behavior is typical, although the value of ϵ depends strongly on q_1 and ω (e.g., $\epsilon \simeq 10^{-3}$ for $q_1 = 0.1$, $\omega = 0.1$, but $\epsilon \simeq 0.4$ for $q_1 = 0.999$, $\omega = 5$).

When one converts from $Z^H_{\ell m}$ to C_{jm}, this behavior forces us to include a *buffer* of ℓ values beyond the maximum spherical harmonic that we want to compute. The size of the buffer depends (rather strongly) on the values of q_1 and ω. In the weak field, where $q_1 m_1 \omega \ll 1$ even for large m, it is enough to include a buffer of 2 (i.e., such that $\ell_{\max} = j_{\max} + 2$). For the results used in this paper, we have chosen a uniform buffer region of size 8.

APPENDIX B: GW ENERGY ABSORPTION

The $N/2$-th order Taylor expansion of the flux in PN theory is given by the series [19]

$$\mathcal{F}^{(N)}_{\text{Tay.}} = \mathcal{F}_{\text{Newt.}} \sum_{n=0}^{N} [a_n(\nu) + b_n(\nu)\log(v)]v^n, \tag{B1}$$

where $\mathcal{F}_{\text{Newt}} \equiv 32/5\nu^2 v^{10}$ is the leading-order (Newtonian) piece of the flux, v is the circular orbital frequency, log stands for the natural logarithm while a_n and b_n are PN parameters, with $b_{n<6} = 0$.

The PN parameters can be classified according to their physical origin and whether they include spin contributions or not. The flux pieces that account for GW emission to infinity are well known and, for example, are given in Ref. [79,80]. Those that correspond to radiation in-falling into the horizon will be labeled $(a^{\text{Hor}}_n, b^{\text{Hor}}_n)$, with a superscript S (NS) if they are further spin dependent (spin independent).

The coefficients associated with BH absorption can only be formally obtained employing BH perturbation theory, as PN theory treats BHs as effective test particles. The logarithm-independent terms associated with nonspinning contributions to the radiation flux through the horizon are

$$a^{\text{Hor,NS}}_8 = 1, \qquad a^{\text{Hor,NS}}_9 = 0, \tag{B2}$$

$$a^{\text{Hor,NS}}_{10} = 4, \qquad a^{\text{Hor,NS}}_{11} = 0, \tag{B3}$$

$$a^{\text{Hor,NS}}_{12} = \tfrac{172}{7}, \qquad a^{\text{Hor,NS}}_{13} = 0, \tag{B4}$$

where $a^{\text{Hor,NS}}_{<8} = 0$. All logarithm-dependent terms identically vanish here: $b^{\text{Hor,NS}}_n = 0$. Similarly, the logarithm-independent terms associated with spinning contributions to the radiation flux through the horizon are

$$a^{\text{Hor,S}}_5 = -\frac{\bar{q}}{4} - \frac{3\bar{q}^3}{4}, \tag{B5}$$

$$a^{\text{Hor,S}}_6 = 0, \tag{B6}$$

$$a^{\text{Hor,S}}_7 = -\bar{q} - \frac{33\bar{q}^3}{16}, \tag{B7}$$

$$\begin{aligned} a^{\text{Hor,S}}_8 = &-\frac{1}{2} + \frac{35}{6}\bar{q}^2 - \frac{3}{12}\bar{q}^4 \\ &+ \left(\frac{1}{2} + \frac{13}{2}\bar{q}^2 + 3\bar{q}^4\right)(1-\bar{q}^2)^{1/2} \\ &+ i\bar{q}(1+3\bar{q}^2))\{\psi^{(0)}[3 - 2i\bar{q}(1-\bar{q}^2)^{-1/2}] \\ &- \psi^{(0)}[3 + 2i\bar{q}(1-\bar{q}^2)^{-1/2}]\}, \end{aligned} \tag{B8}$$

$$a^{\text{Hor,S}}_9 = -\frac{43\bar{q}}{7} - \frac{4651\bar{q}^3}{336} - \frac{17\bar{q}^5}{56}, \tag{B9}$$

$$\begin{aligned} a^{\text{Hor,S}}_{10} = &-2 + \frac{433}{24}\bar{q}^2 - \frac{95}{24}\bar{q}^4 \\ &+ \left(2 + \frac{163}{8}\bar{q}^2 + \frac{33}{4}\bar{q}^4\right)(1-\bar{q}^2)^{1/2} \\ &- \frac{3}{24} i\bar{q}(4 - 3\bar{q}^2)\{\psi^{(0)}[3 + i\bar{q}(1-\bar{q}^2)^{-1/2}] \\ &- \psi^{(0)}[3 - i\bar{q}(1-\bar{q}^2)^{-1/2}]\} - 3i\bar{q}(1+3\bar{q}^2) \\ &\times \psi^{(0)}[3 + 2i\bar{q}(1-\bar{q}^2)^{-1/2}] + 3i\bar{q}(1+3\bar{q}^2) \\ &\times \psi^{(0)}[3 - 2i\bar{q}(1-\bar{q}^2)^{-1/2}], \end{aligned} \tag{B10}$$

where $a^{\text{Hor,S}}_{<5} = 0$ and where the polygamma function $\psi^{(n)}(z) \equiv (d^n\Gamma(z)/dz)\Gamma(z)^{-1}$ is the n-th derivative of the Gamma function. The coefficients $(a^{\text{Hor,S}}_{11}, a^{\text{Hor,S}}_{12})$ are also known, but we do not write them out here as they are lengthy and unilluminating (e.g., see Appendix J in Ref. [24]). Notice that the BH absorption coefficients in the spinning case are nonzero starting at 2.5 PN-order, which is to be contrasted with the nonspinning BH absorption terms that start at 4 PN-order.

An ambiguity exists when incorporating these BH absorption contributions into the flux. As one can observe, the spin-dependent coefficients $a^{\text{Hor,S}}_n$ depend on the spin parameter of the background $\bar{q}$, for which one could choose the real spin parameter $\bar{q} = q_1$ or the effective spin parameter $\bar{q} = q$, defined in Appendix C. Since $q = q_1 + \mathcal{O}(m_2/m_1)$, these choices are identical in the test-particle limit, when we calibrate to Teukolsky fluxes. For lack of better guidance, we here choose $\bar{q} = q$.

APPENDIX C: SPIN EOB HAMILTONIAN

In Sec. VI we have investigated how the analytical results calibrated to the Teukolsky-based waveforms change when we switch on the PN conservative self-force

and the second-order radiation-reaction effects, and when we include the spin of the small object. This study employed the spin EOB Hamiltonian of Ref. [41], which was derived from Ref. [58], building also on results of Ref. [40]. As we shall review below, the Hamiltonian of Ref. [41] reproduces the known spin-orbit (spin-spin) PN couplings through 2.5 PN (2 PN) order for comparable masses, and all PN couplings linear in the spin of the small object in the test-particle limit. We shall here restrict attention to circular, equatorial orbits, and assume that the spins are aligned with the orbital angular momentum.

The motion of a spinning test particle in a generic curved spacetime is described by the Papapetrou equation [81–83]. Reference [58] derived a Hamiltonian whose Hamilton equations are equivalent to the Papapetrou equation. This Hamiltonian therefore describes the motion of a spinning particle in a generic curved spacetime, and Ref. [41] computed it in the particular case of the Kerr spacetime in Boyer-Lindquist coordinates. Denoting with $\mathbf{S}_1$ and m_1 the spin and mass of the background BH, and with $\mathbf{S}_2$ and m_2 the spin and the mass of the smaller BH, the Hamiltonian of a spinning test particle in Kerr has the generic form [41]

$$H = H_{\rm NS} + H_{\rm S}, \tag{C1}$$

where $H_{\rm NS}$ is the Hamiltonian of a nonspinning test particle in Kerr, given by Eq. (2), while $H_{\rm S}$ depends on $\mathbf{S}_1$ and $\mathbf{S}_2$ and, if PN expanded, generates all PN terms linear in the small object's spin $\mathbf{S}_2$.

In Ref. [41] the authors constructed the spin EOB Hamiltonian by mapping the PN Hamiltonian of two BHs of masses $m_{1,2}$ and spins $\mathbf{S}_{1,2}$ into the effective Hamiltonian of a spinning test particle of mass $\mu = m_1 m_2/(m_1 + m_2)$ and spin $\mathbf{S}^*$ moving in a deformed-Kerr spacetime with mass $M = m_1 + m_2$ and spin $\mathbf{S}_{\rm Kerr}$, $\nu = \mu/M$ being the deformation parameter. Note that the deformed-Kerr spin parameter $q \equiv |\mathbf{S}_{\rm Kerr}|/M^2 \neq q_1$, but instead $q \approx q_1(1 - 2m_2/m_1 + \ldots)$ when $m_2/m_1 \ll 1$.

The effective Hamiltonian is [41]

$$H_{\rm eff} = H_{\rm NS} + H_{\rm S} - \frac{\mu}{2Mr^3} S_*^2, \tag{C2}$$

where $H_{\rm NS}$ is the Hamiltonian of a nonspinning effective particle in the deformed-Kerr background,

$$H_{\rm NS} = p_\phi \frac{\tilde{\omega}_{\rm fd}}{\Lambda_t} + \frac{\mu r \sqrt{\Delta_t}\sqrt{Q}}{\sqrt{\Lambda_t}}, \tag{C3}$$

which differs from Eq. (9) in that the Kerr potentials Δ and $\omega_{\rm fd}$ have been replaced by their deformed forms Δ_t and $\tilde{\omega}_{\rm fd}$ (and also $m_1 \to M$, $q_1 \to q$, $m_2 \to \mu$). Furthermore, $H_{\rm S}$ in Eq. (C2) is linearly proportional to the effective particle's spin S_* and reads

$$\begin{aligned} H_{\rm S} = \frac{S_*}{2\mu M\sqrt{\Delta_t}\Lambda_t^{5/2}(\sqrt{Q}+1)\sqrt{Q}r^2}\{&2\mu\sqrt{\Delta_t}\Lambda_t(\sqrt{Q}+1)\times(\sqrt{\Delta_t}p_\phi r^3 + \mu\sqrt{\Lambda_t Q}\tilde{\omega}_{\rm fd})r^2 \\ &+ \sqrt{\Delta_r}[\mu\Delta_{t,r}\sqrt{\Lambda_t}p_\phi((2\sqrt{Q}+1)(r^2+M^2q^2)^2 - \Lambda_t(\sqrt{Q}+1))r^3 \\ &+ 2\mu\Delta_t\sqrt{\Lambda_t}p_\phi(2\sqrt{Q}+1)(\Lambda_t - 2r^2(r^2+M^2q^2))r^2 \\ &+ \sqrt{\Delta_t}(p_\phi^2 r^2 + \mu^2\Lambda_t\sqrt{Q}(1+\sqrt{Q}))(\Lambda_t\tilde{\omega}_{\rm fd,r} - \Lambda_{t,r}\tilde{\omega}_{\rm fd})]\}, \end{aligned} \tag{C4}$$

where we denote with a comma the derivative with respect to r. The term proportional to S_*^2 in Eq. (C2) is added to reproduce known spin-spin results at 2 PN order. The quantities (C3) and (C4) depend on the Kerr-deformed potentials Δ_t, Δ_r, Λ_t, $\tilde{\omega}_{\rm fd}$, while

$$Q = 1 + \frac{p_\phi^2 r^2}{\mu^2 \Lambda_t}. \tag{C5}$$

In particular, we have [41]

$$\Lambda_t = (r^2 + M^2q^2)^2 - M^2q^2\Delta_t, \tag{C6}$$

and

$$\tilde{\omega}_{\rm fd} = 2qM^2 r + \omega_1^{\rm fd}\nu\frac{qM^4}{r} + \omega_2^{\rm fd}\nu\frac{q^3M^4}{r}, \tag{C7}$$

where $\omega_1^{\rm fd}$ and $\omega_2^{\rm fd}$ are two adjustable parameters regulating the frame dragging *strength*. Although precise values for $\omega_1^{\rm fd}$ and $\omega_2^{\rm fd}$ can only be determined by calibrating the model against NR simulations of comparable-mass spinning BHs, a preliminary comparison of the final spin predicted by the EOB model to NR results [84,85] suggests that $\omega_1^{\rm fd} \approx -10$ and $\omega_2^{\rm fd} \approx 20$.

The deformed-Kerr potential Δ_t is given at 3 PN order by

$$\Delta_t = r^2[A(u) + q^2u^2], \tag{C8}$$

$$A(u) = 1 - 2u + 2\nu u^3 + \nu\left(\frac{94}{3} - \frac{41}{32}\pi^2\right)u^4. \tag{C9}$$

When setting $\nu = 0$, Δ_t reduces to the Kerr expression (7) (with $m_1 \to M$, $q_1 \to q$) $\Delta = \Delta_t = r^2 - 2Mr + q^2M^2$. In order to guarantee the presence of deformed horizons (which correspond to the zeros of Δ_t), Ref. [41] suggested rewriting Eq. (C8) as ($u \equiv M/r$)

$$\Delta_t = r^2\left[\frac{1-2u(1-K\nu)}{(1-K\nu)^2}+q^2u^2\right]\times \tag{C10}$$

$$\left[1+\nu\Delta_0+\nu\log\left(1+\sum_{i=1}^{4}\Delta_i u^i\right)\right], \tag{C11}$$

with

$$K = K_0 + 4(K_1-K_0)\nu, \tag{C12}$$

and

$$\Delta_0 = K(\nu K-2), \tag{C13}$$

$$\Delta_1 = -2(\nu K-1)(K+\Delta_0), \tag{C14}$$

$$\Delta_2 = \tfrac{1}{2}\Delta_1(-4\nu K+\Delta_1+4) - q^2(\nu K-1)^2\Delta_0, \tag{C15}$$

$$\begin{aligned}\Delta_3 = {}& \tfrac{1}{3}[-\Delta_1^3+3(\nu K-1)\Delta_1^2+3\Delta_2\Delta_1-6(\nu K-1)\\ & \times(-\nu K+\Delta_2+1)-3q^2(\nu K-1)^2\Delta_1],\end{aligned} \tag{C16}$$

$$\begin{aligned}\Delta_4 = {}& \tfrac{1}{12}\{6q^2(\Delta_1^2-2\Delta_2)(\nu K-1)^2\\ & +3\Delta_1^4-8(\nu K-1)\Delta_1^3-12\Delta_2\Delta_1^2\\ & +12[2(\nu K-1)\Delta_2+\Delta_3]\Delta_1+12(\tfrac{94}{3}-\tfrac{41}{32}\pi^2)\\ & \times(\nu K-1)^2+6[\Delta_2^2-4\Delta_3(\nu K-1)]\}.\end{aligned} \tag{C17}$$

When expanding Eq. (C10) through 3 PN order, one recovers Eq. (C8). The quantity (C10) depends on two parameters K_0 and K_1. K_0 is fixed to the value 1.4467 in order to reproduce the results of Ref. [25] for the shift of the ISCO frequency due to the conservative part of the self force. Also, recent comparisons of the EOB model with numerical simulations of nonspinning comparable-mass BHs have suggested $K_1 \approx 3/4$.

The deformed-Kerr potential Δ_r is given by [41]

$$\Delta_r = \Delta_t\{1+\log[1+6\nu u^2+2(26-3\nu)\nu u^3]\}, \tag{C18}$$

which reduces to the Kerr-potential Δ in the limit $\nu = 0$ (with $m_1 \to M$, $q_1 \to q$).

Finally, the spins $\mathbf{S}_{\rm Kerr}$ and $\mathbf{S}^*$ in the effective description are not equal to S_1 and S_2, but are instead given by

$$S^* = S_1\frac{m_2}{m_1}+S_2\frac{m_1}{m_2}+\frac{1}{c^2}\Delta_{S^*}, \tag{C19}$$

$$S_{\rm Kerr} = S_1+S_2, \tag{C20}$$

where

$$\begin{aligned}\Delta_{S^*} = {}& \frac{\nu}{12}\left\{\frac{2M}{r}\left[7\left(S_1\frac{m_2}{m_1}+S_2\frac{m_1}{m_2}\right)-4(S_1+S_2)\right]\right.\\ & \left.+(Q-1)\left[3(S_1+S_2)+4\left(S_1\frac{m_2}{m_1}+S_2\frac{m_1}{m_2}\right)\right]\right\}.\end{aligned} \tag{C21}$$

With all of this at hand, the EOB Hamiltonian used in Sec. VI is

$$H_{\rm EOB} = M\sqrt{1+2\nu\left(\frac{H_{\rm eff}}{\mu}-1\right)}. \tag{C22}$$

A few final observations are due at this point. When the smaller BH has zero spin $S_2 = 0$ and mass $m_2 \ll m_1$, at lowest order in m_2/m_1 the EOB Hamiltonian of Eq. (C22) reduces to the Hamiltonian of a nonspinning test particle in Kerr. This is because both S_* and the deformations of the Kerr potentials are $\mathcal{O}(m_2/m_1)$. However, at the next-to-leading order in the mass ratio, the EOB Hamiltonian presents corrections with respect to the Hamiltonian of a nonspinning test particle in Kerr, (i) because of the deformations of the Kerr potentials; (ii) because of the effective spin S_*, which is not zero; (iii) because of the higher-order terms in ν that one obtains expanding Eq. (C22). These corrections encode the conservative part of the self-force in the EOB framework [41].

[1] P. Bender *et al.*, Max-Planck-Institut für QuantenoptikReport No. MPQ 233 (1998).
[2] K. Danzmann and A. Rüdiger, Classical Quantum Gravity **20**, S1 (2003).
[3] T. Prince, Bulletin of the American Astronomical Society **35**, 751 (2003).
[4] J. R. Gair, L. Barack, T. Creighton, C. Cutler, S. L. Larson, E. S. Phinney, and M. Vallisneri, Classical Quantum Gravity **21**, S1595 (2004).
[5] P. Amaro-Seoane *et al.*, Classical Quantum Gravity **24**, R113 (2007).
[6] C. F. Sopuerta and N. Yunes, Phys. Rev. D **80**, 064006 (2009).
[7] N. A. Collins and S. A. Hughes, Phys. Rev. D **69**, 124022 (2004).
[8] S. J. Vigeland and S. A. Hughes, Phys. Rev. D **81**, 024030 (2010).
[9] S. A. Hughes, AIP Conf. Proc. **873**, 233 (2006).
[10] T. A. Apostolatos, G. Lukes-Gerakopoulos, and G. Contopoulos, Phys. Rev. Lett. **103**, 111101 (2009).
[11] G. Lukes-Gerakopoulos, T. A. Apostolatos, and G. Contopoulos, Phys. Rev. D **81**, 124005 (2010).
[12] K. Glampedakis and S. Babak, Classical Quantum Gravity **23**, 4167 (2006).
[13] J. R. Gair, C. Li, and I. Mandel, Phys. Rev. D **77**, 024035 (2008).

[14] E. Barausse and L. Rezzolla, Phys. Rev. D **77**, 104027 (2008).
[15] E. Barausse, L. Rezzolla, D. Petroff, and M. Ansorg, Phys. Rev. D **75**, 064026 (2007).
[16] M. Kesden, J. Gair, and M. Kamionkowski, Phys. Rev. D **71**, 044015 (2005).
[17] L. Barack and C. Cutler, Phys. Rev. D **75**, 042003 (2007).
[18] L. Barack and C. Cutler, Phys. Rev. D **69**, 082005 (2004).
[19] P. J. Andrzej Krlak, Living Rev. Relativity **8**, 3 (2005), http://www.livingreviews.org/lrr-2005-3.
[20] C. Cutler and M. Vallisneri, Phys. Rev. D **76**, 104018 (2007).
[21] L. Lindblom, B. J. Owen, and D. A. Brown, Phys. Rev. D **78**, 124020 (2008).
[22] I. Mandel and J. R. Gair, Classical Quantum Gravity **26**, 094036 (2009).
[23] L. Blanchet, Living Rev. Relativity **9**, 4 (2006), and references therein.
[24] Y. Mino *et al.*, Prog. Theor. Phys. Suppl. **128**, 1 (1997).
[25] L. Barack and N. Sago, Phys. Rev. Lett. **102**, 191101 (2009).
[26] N. Sago, Classical Quantum Gravity **26**, 094025 (2009).
[27] L. Barack, T. Damour, and N. Sago, Phys. Rev. D **82**, 084036 (2010).
[28] L. Barack, Classical Quantum Gravity **26**, 213001 (2009).
[29] S. A. Hughes, Phys. Rev. D **61**, 084004 (2000).
[30] S. A. Hughes, Phys. Rev. D **64**, 064004 (2001).
[31] E. Poisson, Phys. Rev. D **47**, 1497 (1993).
[32] A. Pound, E. Poisson, and B. G. Nickel, Phys. Rev. D **72**, 124001 (2005).
[33] J. R. Gair and K. Glampedakis, Phys. Rev. D **73**, 064037 (2006).
[34] S. Babak, H. Fang, J. R. Gair, K. Glampedakis, and S. A. Hughes, Phys. Rev. D **75**, 024005 (2007).
[35] A. Buonanno and T. Damour, Phys. Rev. D **59**, 084006 (1999).
[36] A. Buonanno and T. Damour, Phys. Rev. D **62**, 064015 (2000).
[37] T. Damour, P. Jaranowski, and G. Schaefer, Phys. Rev. D **62**, 084011 (2000).
[38] T. Damour, Phys. Rev. D **64**, 124013 (2001).
[39] A. Buonanno, Y. Chen, and T. Damour, Phys. Rev. D **74**, 104005 (2006).
[40] T. Damour, P. Jaranowski, and G. Schaefer, Phys. Rev. D **78**, 024009 (2008).
[41] E. Barausse and A. Buonanno, Phys. Rev. D **81**, 084024 (2010).
[42] A. Nagar, T. Damour, and A. Tartaglia, Classical Quantum Gravity **24**, S109 (2007).
[43] T. Damour and A. Nagar, Phys. Rev. D **76**, 064028 (2007).
[44] S. Bernuzzi and A. Nagar, Phys. Rev. D **81**, 084056 (2010).
[45] T. Damour, B. R. Iyer, and B. S. Sathyaprakash, Phys. Rev. D **57**, 885 (1998).
[46] T. Damour, B. R. Iyer, and A. Nagar, Phys. Rev. D **79**, 064004 (2009).
[47] Y. Pan, A. Buonanno, R. Fujita, E. Racine, and H. Tagoshi, arXiv:1006.0431.
[48] R. Fujita and B. R. Iyer, Phys. Rev. D **82**, 044051 (2010).
[49] T. Damour and A. Nagar, Phys. Rev. D **79**, 081503 (2009).
[50] A. Buonanno *et al.*, Phys. Rev. D **79**, 124028 (2009).
[51] Y. Pan *et al.*, Phys. Rev. D **81**, 084041 (2010).
[52] T. Damour, Phys. Rev. D **81**, 024017 (2010).
[53] N. Yunes, A. Buonanno, S. A. Hughes, M. Coleman Miller, and Y. Pan, Phys. Rev. Lett. **104**, 091102 (2010).
[54] N. Yunes, GW Notes **2**, 3 (2009).
[55] Y. Levin, Mon. Not. R. Astron. Soc. **374**, 515 (2007).
[56] J. M. Bardeen and J. A. Petterson, Astrophys. J. **195**, L65 (1975).
[57] J. M. Bardeen, W. H. Press, and S. A. Teukolsky, Astrophys. J. **178**, 347 (1972).
[58] E. Barausse, E. Racine, and A. Buonanno, Phys. Rev. D **80**, 104025 (2009).
[59] S. A. Teukolsky and A. Teukolsky, Astrophys. J. **185**, 635 (1973).
[60] S. A. Hughes, Classical Quantum Gravity **18**, 4067 (2001).
[61] S. Drasco and S. A. Hughes, Phys. Rev. D **73**, 024027 (2006).
[62] R. Fujita and H. Tagoshi, Prog. Theor. Phys. **112**, 415 (2004).
[63] R. Fujita and H. Tagoshi, Prog. Theor. Phys. **113**, 1165 (2005).
[64] W. Throwe and S. A. Hughes (unpublished); also W. Throwe, B.S. thesis, Massachusetts Institute of Technology, 2010 (unpublished).
[65] M. Shibata, M. Sasaki, H. Tagoshi, and T. Tanaka, Phys. Rev. D **51**, 1646 (1995).
[66] T. Tanaka, H. Tagoshi, and M. Sasaki, Prog. Theor. Phys. **96**, 1087 (1996).
[67] L. Barack and C. Cutler, Phys. Rev. D **70**, 122002 (2004).
[68] E. Berti, A. Buonanno, and C. M. Will, Phys. Rev. D **71**, 084025 (2005).
[69] A. Pound and E. Poisson, Phys. Rev. D **77**, 044013 (2008).
[70] M. Favata, Phys. Rev. D **83**, 024027 (2011).
[71] E. A. Huerta and J. R. Gair, Phys. Rev. D **79**, 084021 (2009).
[72] W.-B. Han, Phys. Rev. D **82**, 084013 (2010).
[73] M. Favata (unpublished).
[74] L. Blanchet, S. Detweiler, A. Le Tiec, and B. F. Whiting, Phys. Rev. D **81**, 064004 (2010).
[75] L. Blanchet, S. Detweiler, A. L. Tiec, and B. F. Whiting, Phys. Rev. D **81**, 084033 (2010).
[76] S. A. Teukolsky and W. H. Press, Astrophys. J. **193**, 443 (1974).
[77] S. W. Hawking and J. B. Hartle, Commun. Math. Phys. **27**, 283 (1972).
[78] A. A. Starobinskij and S. M. Churilov, Zh. Eksp. Teor. Fiz. **65**, 3 (1973).
[79] M. Boyle *et al.*, Phys. Rev. D **78**, 104020 (2008).
[80] K. G. Arun, A. Buonanno, G. Faye, and E. Ochsner, Phys. Rev. D **79**, 104023 (2009).
[81] A. Papapetrou, Proc. Phys. Soc. London Sect. A **64**, 57 (1951).
[82] A. Papapetrou, Proc. R. Soc. A **209**, 248 (1951).
[83] E. Corinaldesi and A. Papapetrou, Proc. R. Soc. A **209**, 259 (1951).
[84] L. Rezzolla *et al.*, Phys. Rev. D **78**, 044002 (2008).
[85] E. Barausse and L. Rezzolla, Astrophys. J. Lett. **704**, L40 (2009).