Self-forced evolutions of an implicit rotating source: A natural framework to model comparable and intermediate mass-ratio systems from inspiral through ringdown

E. A. Huerta
Prayush Kumar
Jonathan R. Gair
Sean T. McWilliams

Follow this and additional works at: https://researchrepository.wvu.edu/faculty_publications

Digital Commons Citation
Huerta, E. A.; Kumar, Prayush; Gair, Jonathan R.; and McWilliams, Sean T., "Self-forced evolutions of an implicit rotating source: A natural framework to model comparable and intermediate mass-ratio systems from inspiral through ringdown" (2014). Faculty Scholarship. 324.
https://researchrepository.wvu.edu/faculty_publications/324
Self-forced evolutions of an implicit rotating source: A natural framework to model comparable and intermediate mass–ratio systems from inspiral through ringdown

E. A. Huerta,1 Prayush Kumar,2 Jonathan R. Gair,3 and Sean T. McWilliams1

1Department of Physics, West Virginia University, White Hall, Morgantown, WV 26506, USA
2Department of Physics, Syracuse University, Syracuse, NY 13244, USA
3Institute of Astronomy, Madingley Road, CB3 0HA Cambridge, UK

(Dated: July 25, 2018)

We develop a waveform model to describe the inspiral, merger and ringdown of binary systems with comparable and intermediate mass-ratios. This model incorporates first-order conservative self-force corrections to the energy and angular momentum, which are valid in the strong-field regime [1]. We model the radiative part of the self-force by deriving second-order radiative corrections to the energy flux. These corrections are obtained by minimizing the phase discrepancy between our self-force model and the effective one body model [2,3] for a variety of mass-ratios. We show that our model performs substantially better than post-Newtonian approximants currently used to model neutron star-black hole mergers from early inspiral to the innermost stable circular orbit. In order to match the late inspiral evolution onto the plunge regime, we extend the “transition phase” developed by Ori and Thorne [4] by including finite mass–ratio corrections and modelling the orbital phase evolution using an implicit rotating source [5]. We explicitly show that the implicit rotating source approach provides a natural transition from late-time radiation to ringdown that is equivalent to ringdown waveform modelling based on a sum of quasinormal modes.
I. INTRODUCTION

The black hole (BH) mass function in the local Universe is a strongly bi-modal distribution that identifies two main families: stellar-mass BHs with typical masses $\sim 10M_{\odot}$ observed in Galactic X-ray binaries [6] and, more recently, in globular clusters [7], and supermassive BHs with masses $\gg 10^5M_{\odot}$ observed to be present in most galactic nuclei [8, 9]. However, a population of X-ray sources with luminosities in excess of $10^{39}$ erg s$^{-1}$ has recently been observed, and Chandra and XMM-Newton spectral observations of these ultra-luminous X-ray sources (ULXs) revealed cool disc signatures that were consistent with the presence of intermediate mass BHs (IMBHs) with masses $10^{2-4}M_{\odot}$ [10–12]. Subsequent observations have shown that these ULXs have spectral and temporal signatures that are not consistent with the sub-Eddington accretion regime that is expected for IMBHs at typical ULX luminosities. Rather, these later studies suggest that many ULXs are powered by super-Eddington accretion onto young stellar cluster, the features if its power spectrum, and some reported transitions between a hard state and a soft state seem to be inconsistent with the sub-Eddington accretion regime that is expected for IMBHs at typical ULX luminosities. Rather, these later studies suggest that many ULXs are powered by super-Eddington accretion onto massive BHs formed in low metallicity regions [14–16]. Several hyper-luminous X-ray sources, including M82 X-1, ESO 243-49 HLX-1, Cartwheel N 10 and CXO J122518.6+144545, present the best indirect evidence for the existence of IMBHs [17–20]. In particular, the colocation of M82 X-1 with a massive, young stellar cluster, the features if its power spectrum, and some reported transitions between a hard state and a thermal dominant state, make this object a strong IMBH candidate [21–24]. Recent searches of archival Chandra and XMM-Newton data sets have also uncovered two new hyper-luminous X-ray sources with luminosities in excess of $10^{39}$ erg s$^{-1}$. These sources are the most promising IMBH candidates currently known, although the highest possible super-Eddington accretion rate onto the largest permitted BH remnant cannot yet be ruled out [25]. This increasing body of observational evidence [26, 27], and the fact that the existence of IMBHs provides a compelling explanation for the initial seeding of supermassive BHs present in most galactic nuclei [28–30], has revived the quest for these elusive objects.

Since hyper-luminous X-ray sources are rare, and our knowledge about their astrophysical properties is still limited, we may have to use a different means to search for IMBHs in order to improve our knowledge about the channels that lead to the formation of these objects, and to shed light on their astrophysical properties, such as mass and spin distributions [31]. In this paper we explore the use of observations in the gravitational wave (GW) spectrum to gain insight into the properties of IMBHs.

The current upgrade of the Laser Interferometer Gravitational Wave Observatory (LIGO) and its international partners Virgo and Kagra [32–34], will enable the detection of GWs from coalescences involving IMBHs with masses $50M_{\odot} \lesssim M \lesssim 500M_{\odot}$, if these instruments achieve their target sensitivity down to the low-frequency cutoff at 10Hz (See Figure I.37). Advanced LIGO (aLIGO) and Advanced Virgo are expected to have greatest sensitivity in the 60Hz - 500Hz range, with a peak at $\sim 60$ Hz (see Figure I.37). Proposed third generation detectors, such as the Einstein Telescope [35], which aim to extend the frequency range of ground-based detectors down to 1Hz, while also maintaining high frequency sensitivity up to 10kHz, will enhance our ability to search for GWs emitted by sources that involve BHs with masses between $10^{2-4}M_{\odot}$ [37–39].

A promising channel for detection of IMBHs is through the emission of gravitational radiation during the coalescence of stellar-mass compact remnants — neutron stars (NSs) or BHs — with IMBHs in core-collapsed globular clusters. This expectation is backed up by numerical simulations of globular clusters [40–43] which suggest that IMBHs could undergo several collisions with stellar-mass compact remnants during the lifetime of the cluster through a variety of mechanisms, including gravitational radiation, Kozai resonances and binary exchange processes. As discussed in [44], the most likely mechanism for the formation of binaries involving a stellar-mass compact remnant and an IMBH is hardening via three body interactions, with an expected detection rate of $\sim 1 - 10$ yr$^{-1}$ with ground-based observatories [45–47].

The frequency of the dominant quadrupolar harmonic in the GWs emitted at the innermost stable circular orbit (ISCO) for a binary of non-spinning objects is

$$f_{\text{ISCO}} = 4.4\text{kHz} \left(\frac{M_{\odot}}{M}\right),$$

so for a typical intermediate mass-ratio coalescence (IMRC) with total mass $M \gtrsim 100M_{\odot}$, advanced detectors will observe the late inspiral, merger and ringdown. However, the heaviest IMRCs, with masses $\gtrsim 500M_{\odot}$, will coalesce out of band or at the low frequency limit of the bandwidth. Hence, merger and ringdown — which intrinsically generate a signal-to-noise ratio (SNR) suppressed by a factor of symmetric mass-ratio — see Table I.3 below — relative to the SNR generated during the inspiral phase — will significantly contribute to the SNR of IMRCs over a considerable portion of the detectable mass-range [50]. In order to get the most information from GW observations of IMRCs, it
is therefore necessary to develop waveform models that incorporate the inspiral, merger and ringdown in a physically consistent way. In order to make progress in this direction, we previously developed a waveform model that combined results from Black Hole Perturbation Theory (BHPT) and post-Newtonian (PN) theory to explore the information that could be obtained from observations of IMRCs with the Einstein Telescope \cite{38,39}. Although this model provided an important step in exploring the science that could be done with IMRC observations, the model was limited. In \cite{51,54–56} we explored using the self-force formalism \cite{52,53} to develop a waveform model with a robust description of the dynamical evolution of IMRCs during the inspiral phase. These were found to be effective when used to carry out matched-filter based searches for inspiral-only IMRCs \cite{50}, but searches for IMRCs in the advanced detector era will require waveform models that include not only the inspiral but also the merger and ringdown \cite{50}. The model described in \cite{38} included merger and ringdown but without the self-force driven inspiral. It would be ideal to develop a consistent inspiral, merger, ringdown waveform model by comparison to numerical relativity simulations, as described in \cite{2,3}. However, numerical relativity simulations for systems with mass-ratios \( q = m_1/m_2 \sim <1/10 \) are very computationally expensive at present \cite{57}. We can circumvent this problem by making use of recent breakthroughs in the self-force program that have shown that the conservative part of the self-force can reproduce, with good accuracy, results from numerical relativity simulations of comparable-mass binary systems \cite{58}. Furthermore, the recent computation of the self-force inside the ISCO equips us to now develop models that better reproduce the true dynamics of black hole binaries in the strong field regime \cite{1}. In addition to IMRCs, which may be detected by second generation detectors, and with more likelihood by third generation detectors, we have explored the suitability of using self-forced evolutions to model the mergers of systems involving NSs and stellar mass BHs \cite{59}. Since NSBH mergers are promising GW sources for second generation detectors, we need to develop accurate and computationally inexpensive templates appropriate for these events. Current efforts to model NSBH systems have generally used PN approximants, evaluated to the highest PN order available, but these approximants are not sufficiently reliable to model these events \cite{60,61}. In Figure 2 we show the phase difference between the PN approximant TaylorT4 \cite{62} and the EOB model introduced in \cite{2,3}. This exhibits a substantial discrepancy near the ISCO. These considerations have impelled us to develop a different approach to model events with the typical mass-ratios expected for NSBH binaries. In this paper we combine recent developments in the self-force program, in PN theory and in numerical relativity to develop a model that describes the inspiral, merger and ringdown of IMRCs and comparable mass-ratio systems that could be detected by second and third generation ground-based GW detectors. In Section 4 we discuss the modelling of the conservative part of the self-force. We show that using the linear in mass-ratio self force results for binaries with mass-ratios \( q \geq 1/6 \) gives a system without an ISCO. We then discuss the implications of this result for the modelling of comparable and intermediate mass-ratio binaries. Thereafter, we describe the approach we use to model the radiative part of the self-force for the inspiral evolution. Having constructed the inspiral part of the self-forced waveform model, we extend the transition scheme of Ori and Thorne \cite{63} by including finite mass-ratio corrections, and modelling the orbital phase evolution using the implicit rotating source (IRS) model. We adopt this description for the late-time radiation in order to provide a smooth progression from late inspiral to ringdown. We show that this approach provides the correct orbital frequency evolution in the vicinity of the light-ring. Finally, we construct the
FIG. 2. The phase discrepancy in radians between the PN approximant TaylorT4, and the Effective One Body model, shown as a function of time from $r = 30M$ to the point when the TaylorT4 model reaches the ISCO. The systems have mass-ratio, $q$, total mass, $M$, and final phase discrepancy, $\Delta \Phi$: $(q, M, \Delta \Phi) = (1/6, 7M_\odot, 21.5 \text{ rads})$ (top-left), $(1/8, 9M_\odot, 30.2 \text{ rads})$ (top-right), $(1/10, 11M_\odot, 70.1 \text{ rads})$ (bottom-left) and $(1/15, 16M_\odot, 83.2 \text{ rads})$ (bottom-right) respectively.

ringdown waveform using both a sum of quasinormal modes and the late-time radiation waveform evolution predicted by the IRS model, and show their equivalence. Section III presents a summary of our findings and future directions of work.

II. MODELING

A. Nomenclature

Throughout this paper we will use units with $G = c = 1$, unless otherwise stated. We will consider BH binaries on circular orbits with component masses $m_1, m_2$, such that $m_1 < m_2$. We assume that the binary components are non spinning. We will use several combinations of the masses $m_{1,2}$ in the following Sections, which are summarized in Table I.

Having defined the variables to be used in the subsequent sections, we shall now describe the construction of the self-forced waveform model. The model consists of four building blocks — the inspiral, the transition, the plunge and the ringdown phases. The next section describes the inspiral evolution.
### Binary masses

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_1$</td>
<td>mass of inspiralling compact object</td>
</tr>
<tr>
<td>$m_2$</td>
<td>mass of central compact object</td>
</tr>
<tr>
<td>$M = m_1 + m_2$</td>
<td>total mass of binary system</td>
</tr>
<tr>
<td>$q = \frac{m_1}{m_2}$</td>
<td>mass–ratio</td>
</tr>
<tr>
<td>$\mu = \frac{m_1 m_2}{m_1 + m_2}$</td>
<td>reduced mass</td>
</tr>
<tr>
<td>$\eta = \frac{\mu}{M}$</td>
<td>symmetric mass–ratio</td>
</tr>
</tbody>
</table>

**TABLE I.** The table summarizes the nomenclature we will use throughout our analysis.

### B. Inspiral evolution

We model the inspiral phase evolution in the context of the Effective One Body (EOB) formalism [64], i.e., we consider the scenario in which the dynamics of a binary system is mapped onto the motion of a test particle in a time-independent and spherically symmetric Schwarzschild space-time with total mass $M$:

$$ds_{\text{EOB}}^2 = -A(r)dt^2 + B(r)dt^2 + r^2d\Omega^2,$$

where the potentials $A, B$ are known to 3PN order [65, 66]. In the test-mass particle limit $\eta \to 0$, these potentials recover the Schwarzschild results, namely:

$$A(u, \eta \to 0) = B^{-1}(u, \eta \to 0) = 1 - 2u, \quad \text{with} \quad u = \frac{M}{r}.$$  

(3)

In the EOB formalism, the orbital frequency evolution can be derived from a Hamiltonian, $H_{\text{EOB}}$ [64], given by:

$$H_{\text{EOB}} = M \sqrt{1 + 2\eta (H_{\text{eff}} - 1)},$$  

(4)

using the general Hamiltonian equation:

$$d\phi \over dt = M\Omega = \frac{\partial H_{\text{EOB}}}{\partial L} = \frac{u^2 L(x) A(u)}{H(u) H_{\text{eff}}(u)},$$

(5)

where

$$H_{\text{eff}}(u) = \frac{A(u)}{\sqrt{A(u)}}, \quad \dot{A}(u) = A(u) + \frac{1}{2} u A'(u),$$

(6)

and

$$H(u) = \sqrt{1 + 2\eta (H_{\text{eff}} - 1)}.$$  

Recent work in the self-force formalism has enabled the derivation of gravitational self-force corrections to the EOB potential $A(u) \to 1 - 2u + \eta a(u) + \mathcal{O}(\eta^2)$ [67]. Deriving this gravitational self-force contribution, $a(u)$, is equivalent to including all PN corrections to the EOB potential $A(u)$ at linear order in $\eta$. We shall now briefly describe the construction of the gravitational self-force contribution $a(u)$, emphasizing the fact that this contribution encodes information about the strong-field regime of the gravitational field.

As shown by Detweiler and Whiting [68], the gravitational self-force corrected worldline can be interpreted as a geodesic in a smooth perturbed spacetime with metric

$$g_{\alpha\beta} = g^{00}_{\alpha\beta}(m_2) + h^{R}_{\alpha\beta},$$

(7)

where the regularized $R$ field is a smooth perturbation associated with $m_1$. Detweiler proposed a gauge invariant relation to handle the conservative effect of the gravitational self-force in circular motion [69, 70]:

$$z_1(\Omega) = \sqrt{1 - 3x} \left(1 - \frac{1}{2} h^{R,F}_{uu} + q \frac{x}{1 - 3x}\right),$$

(8)

where $x$ is the gauge-invariant dimensionless frequency parameter given by $x = (M\Omega)^{2/3}$, $h^{R,G}_{uu}$ is a double contraction of the regularised metric perturbation with the four-velocity, $u^\mu$, $h^{R,F}_{uu} = h^{R,G}_{\mu\nu} u^\mu u^\nu$, the label $G$ indicates the gauge
used to evaluate the metric perturbation and the label $F$ indicates that this expression is valid within the class of asymptotically flat gauges. In [1], $z_1(\Omega)$ was calculated in Lorenz gauge and the following gauge transformation can be used to link the asymptotically flat $h_{uu}^{R,F}$ metric perturbation to its Lorenz-gauge counterpart $h_{uu}^{R,L}$:

$$h_{uu}^{R,F} = h_{uu}^{R,L} + 2q \frac{x(1-2x)}{(1-3x)^{3/2}}.$$  \hfill (9)

Hence, inserting Eq. (9) into Eq. (5) leads to

$$z_1(\Omega) = \sqrt{1-3x} \left( 1 - \frac{1}{2} h_{uu}^{R,L} - 2q \frac{x(1-2x)}{(1-3x)^{3/2}} + q \frac{x}{1-3x} \right).$$  \hfill (10)

The numerical data obtained in [1] for $h_{uu}^{R,L}$ from $x > 1/5$ was new and the numerical accuracy for $x < 1/5$ was also much improved compared to previous results [69, 71]. The EoB potential $a(x)$ can be constructed from $h_{uu}^{R,L}$ via

$$a(x) = -\frac{1}{2} (1-3x) \tilde{h}_{uu}^{R,L} - 2\sqrt{1-3x},$$  \hfill (11)

with $\tilde{h}_{uu}^{R,L} = q^{-1} h_{uu}^{R,L}$. In [1], a global fit formula for $a(x)$ was given a dense sample of numerical values over the entire range $0 < x < \frac{1}{3}$. The numerical fit for $a(x)$ that we use in this study is taken from Eq. (54) of [1]. This global analytic fit for $a(x)$ can be recast using the relation

$$a(x) = 2x^3 \frac{(1-2x)}{\sqrt{1-3x}} a_E(x).$$  \hfill (12)

Using the above dictionary, the model for $a(x)$ in this paper reproduces the numerical data points for the function $a_E(x)$ to within a maximal absolute difference of $1.2 \times 10^{-5}$ over the domain $0 < x < \frac{1}{3}$. The corresponding self-force corrected energy and angular momentum are given by [1, 67]

$$E(u(x)) = E_0(x) + \eta \left( -\frac{1}{3} \frac{x}{\sqrt{1-3x}} a'(x) + \frac{1}{2} \frac{1-4x}{3(1-3x)^{3/2}} a(x) - E_0(x) \left( \frac{1}{2} E_0(x) + \frac{x}{3} \frac{1-6x}{(1-3x)^{3/2}} \right) \right),$$  \hfill (13)

$$L_z(u(x)) = L_0(x) + \eta \left( -\frac{1}{3} \frac{x}{\sqrt{x(1-3x)}} a'(x) - \frac{1}{2} \frac{1}{\sqrt{x(1-3x)^{3/2}}} a(x) - \frac{1}{3} \frac{1-6x}{\sqrt{3x(1-3x)^{3/2}}} (E_0(x) - 1) \right),$$  \hfill (14)

with

$$u(x) = x \left( 1 + \frac{1}{6} a'(x) + \frac{2}{3} \frac{1-2x}{\sqrt{1-3x}} - 1 \right),$$  \hfill (15)

where $E_0(x)$ and $L_0(x)$ are given by

$$E_0(x) = \frac{1-2x}{\sqrt{1-3x}} - 1,$$  \hfill (16)

$$L_0(x) = \frac{1}{\sqrt{x(1-3x)}}.$$  \hfill (17)

In Figure 8 we show the effect of these conservative corrections on the orbital parameters.

As discussed in [67], if one makes use of the self-force expression for the energy given by Eq. (13), and minimize it with respect to the orbital frequency, then one finds that binary systems with mass-ratios $q \in [1, 1/2, 1/3]$ do not have an ISCO in this model. It was argued in [67] that deriving self-force results in the strong field regime may alleviate this problem. We have explored this issue, and have found that using linear self-force corrections that are valid all the way to the light ring (last unstable circular orbit for massless particles) does not fix this problem for comparable mass-ratio systems. In Figure 8 we show that the existence of an ISCO is guaranteed for BH binaries with symmetric mass-ratio $\eta \leq 6/49$ (or $q \leq 1/6$), and its location may be approximated by

$$x_{\text{ISCO}} = \frac{1}{6} \left( 1 + 0.83401\eta + 4.59483\eta^2 \right).$$  \hfill (18)

It remains to be seen whether the inclusion of second-order conservative corrections gives an ISCO for binaries with mass-ratios $q \geq 1/6$.

In summary, the building blocks we use to construct the conservative dynamics are
FIG. 3. The panels show the energy and angular momentum given by Eqs. (13)-(14), respectively. We show the functional form of these parameters for binary systems with mass-ratio values, from top to bottom, $q \in [0, 1/100, 1/20, 1/10, 1/6, 1/5, 1]$.

FIG. 4. The location of the innermost stable circular orbit is determined by the condition $dE/dx = 0$. The panel shows $dE/dx$ as a function of the gauge invariant quantity $x = (M\Omega)^{2/3}$. The various curves represent binary systems with mass-ratios, from top to bottom, $q \in [0, 1/100, 1/20, 1/10, 1/6, 1/5, 1]$. Note that binaries with mass-ratios $q \gtrsim 1/6$ do not have an ISCO in this model.

- The orbital frequency evolution is computed using Eq. (5) with the gravitational self-force contribution included in the potential $A(u) = 1 - 2u + \eta a(u)$.

- Eq. (5) is evaluated using the self-force-corrected expression for the angular momentum, $L(x)$, given by Eq. (14). The self-force-corrected expression for the energy, given in Eq. (13) is only used to determine the point at which the inspiral ends and the transition region begins.

Eq. (5) provides a very accurate modeling of the orbital frequency from early inspiral through the ISCO. However, the post-ISCO time evolution of this prescription does not render an accurate representation of the orbital frequency as compared to numerical relativity simulations. This is a problem that has been addressed in the EOB formalism by introducing a phenomenological approach — the so-called non-quasi-circular coefficients — that enabled them to reproduce the orbital evolution extracted from numerical simulations. The approach we will follow to circumvent this problem is described in detail in Section II D and consists of embedding the self-force formalism in the context of the implicit rotating source model after the ISCO.

This completes the description of the conservative part of the self-force. We now describe how to couple this with the radiative part of the self-force to model the inspiral evolution.
C. Dissipative dynamics

A consistent self-force evolution model that incorporates first-order in mass-ratio conservative corrections should also include second-order radiative corrections. The model we described in the previous section was constructed including first order conservative corrections. However, second-order self-force radiative corrections are not known at present. Several studies have demonstrated the importance of including the missing second order corrections to the radiative part of the self-force, both for source detection and parameter estimation [51, 54, 55, 72, 73].

We use a new prescription for the energy flux that uses the first-order in mass-ratio terms derived in [74], including PN corrections up to 22nd PN order

\[
\left( \frac{\dot{E}}{E} \right)_{\text{PN}} = -\frac{32\mu^2}{5M} x^{7/2} \left[ 1 - \frac{1247}{336} x^2 + 4\pi x^{3/2} - \frac{44711}{9072} x^2 - \frac{8191}{672} \pi x^{5/2} \right] + x^3 \left\{ \frac{6643739519}{69854400} + \frac{16}{3} x^2 - \frac{1712}{105} \gamma \right\} - \frac{856}{105} \ln(16x) - \frac{16285}{504} \pi x^{7/2} + x^4 \left\{ -\frac{323105549467}{3178375200} + \frac{232597}{4410} \gamma - \frac{1369}{126} \pi + \frac{39931}{294} \ln(2) - \frac{47385}{1568} \ln(3) + \frac{232597}{4410} \ln(x) \right\} + x^9 \left\{ -\frac{265978667519}{745113600} - \frac{6848}{105} \gamma \pi - \frac{13696}{105} \pi \ln(2) - \frac{6848}{105} \pi \ln(x) \right\} + \text{higher order corrections up to 22PN order} \right].
\]

We include higher-order in mass-ratio terms using the exponential resummation approach described in [72]. In this approach, the energy flux is

\[
\left( \frac{dE}{dt} \right)_{\text{hybrid}} = \mathcal{L}_0 \exp \left( \mathcal{L}_\eta \right),
\]

where \( \mathcal{L}_0 \) denotes the leading-order in mass-ratio PN energy flux given in Eq. (19), and \( \mathcal{L}_\eta \) incorporates mass-ratio corrections to the highest PN order available [72, 77, 76], and additional corrections characterised by a set of unknown coefficients, \( b_i \)

\[
\mathcal{L}_0 = \left[ x \left[ -\frac{35}{12} \eta + b_1 \eta^2 \right] + 4\pi x^{3/2} \left[ b_2 \eta + b_3 \eta^2 \right] + x^2 \left[ \frac{9271}{504} \eta + \frac{65}{18} \eta^2 \right] + \pi x^{5/2} \left[ -\frac{583}{24} \eta + b_4 \eta^2 \right] \right] + x^3 \left[ \eta \left( \frac{134543}{7776} + \frac{41}{48} \pi^2 \right) - \frac{9403}{3024} \eta^2 - \frac{775}{324} \eta^3 \right] + \pi x^{7/2} \left[ \frac{214745}{1728} \eta + \frac{193385}{3024} \eta^2 \right].
\]

The coefficients \( b_i \) were taken to be constant in [72], but we found that a better match to the EOB phase evolution could be obtained by allowing an additional dependence on mass-ratio in these terms (see Eqs. (22)–(24) below). We constrain the \( b_i \) coefficients by ensuring that the phase evolution of this model reproduces the phase evolution predicted by the EOB model introduced in [2, 3], which was calibrated to the phase evolution of compact binaries observed in numerical relativity simulations. To do so, we implemented the EOB model [2] and performed a Monte Carlo simulation to optimize the values of the \( b_i \) coefficients (see Figure 5). The optimization was done in two stages. We started by considering the three coefficients \( b_1, b_2 \) and \( b_4 \), sampling a wide range of parameter space, namely \( b_i \in [-200, 200] \). We constrained the duration of the waveform from early inspiral to the light-ring to be similar to its EOB counterpart. Waveforms that differed from their EOB counterparts by more than \( 10^{-4} \) were discarded.

Once the region under consideration had been sparsely sampled, we focused on regions of parameter space where the orbital phase evolution was closest to the EOB evolution, and finely sampled these to obtain the optimal values for the coefficients. We found that this approach enabled us to reproduce the EOB phase evolution with a phase discrepancy of the order \( \sim 1 \) rad. After constraining \( b_1, b_2 \) and \( b_4 \), we explored whether including additional corrections could further improve the phase evolution, by adding \( \eta \) corrections beyond 3PN order. Such corrections were found to have a negligible impact on the actual phase evolution. This is not difficult to understand, since such corrections are of order \( O(\eta^4), O(\eta^5) \), at (3PN, 3.5PN) respectively. We found a similar behavior when we added leading order mass-ratios corrections beyond 4PN order. Thus, we took a different approach: having derived the optimal value for \( b_1, b_2 \) and \( b_4 \), we took these results as initial seeds for an additional MC simulation in which \( b_3 \) was also included in Eq. (21), and repeated the optimization procedure. The results of these simulations are shown in Figure 5.

We carried out several different Monte Carlo runs to find the ‘optimal’ optimization interval, meaning the range of radial separations over which we tried to best match the phase evolution relative to the EOB model. We found
that starting the optimization at \( r = 30M \) gave results that performed moderately well at early inspiral, but that underperformed at late inspiral, leading to phase discrepancies of order \( \sim 3 \) rads. Starting the optimization at \( r = 20M \) instead decreased the phase discrepancy with respect to the former case by a factor of 10 during early inspiral, and enabled us to reproduce the phase evolution in the EOB model for all the mass-ratios considered to within the accuracy of the numerical waveforms used to calibrate the EOB model in \([3, 4]\). Implementing these numerically optimized higher-order \( \eta \) corrections in Eq. (21) leads to:

\[
\mathcal{L}_\eta = \frac{x}{12} \left[ -35 \eta + B_1 \right] + 4\pi x^{3/2} B_2 + x^2 \left[ \frac{9271}{504} \eta + \frac{65}{18} \eta^2 \right] + \pi x^{5/2} \left[ -\frac{583}{24} \eta + B_3 \right] \\
+ x^3 \left[ \eta \left( \frac{134543}{7776} + \frac{41}{48} \eta^2 \right) - \frac{94403}{3024} \eta^2 - \frac{775}{324} \eta^3 \right] + \pi x^{7/2} \left[ \frac{214745}{1728} \eta + \frac{193385}{3024} \eta^2 \right],
\]

where:

\[
B_1 = \frac{1583.650 - 11760.507 \eta}{1 + 142.389 \eta - 981.723 \eta^2 \eta^2},
\]

\[
B_2 = \frac{-12.081 + 35.482 \eta}{1 - 4.678 \eta + 13.280 \eta^2 \eta^2} \eta + \frac{19.045 - 240.031 \eta}{1 - 18.461 \eta + 74.142 \eta^2 \eta^2},
\]

\[
B_3 = \frac{51.814 - 980.100 \eta}{1 - 13.912 \eta + 88.797 \eta^2 \eta^2} \eta^2.
\]

This improved prescription for the energy flux, which incorporates second-order mass-ratio corrections to the PN expansion up to 3.5PN order, is sufficient to generate a model whose phase evolution reproduces with excellent accuracy the phase evolution predicted by EOB throughout inspiral and merger (see Figure 6).

Given the energy flux defined by Eqs (19)–(21), we generate the inspiral trajectory using the simple prescription

\[
\frac{dx}{dt} = \frac{dE}{dt} \frac{dx}{dE},
\]

where we have used the mass-ratio corrected energy —Eq. (13)— to compute \( dE/dx \). Figure 6 shows that for binaries with mass-ratio \( q = 1/6 \), the phase discrepancy between our self-force model and EOB is \( \lesssim 0.5 \) rads at the light-ring, which is within the numerical accuracy of the simulations used to calibrate EOB. It has been shown recently that EOB remains accurate for mass-ratios up to \( q = 1/8 \) \([17]\). In that regime the phase discrepancy between this model and EOB at the light-ring is less than 1 rad, as shown in Figure 6. For binaries with \( q = 1/10 \), the phase discrepancy at the light-ring is \( \lesssim 1.2 \) rads, which is still within the numerical accuracy of available simulations \([78, 79]\).

It must be emphasized that even if we only use the inspiral evolution to model binaries with mass-ratios that typically describe NSBH binaries, our self-force evolution model performs better than TaylorT4, since we can reduce the phase discrepancy between TaylorT4 and EOB at the last stable circular orbit by a factor of \( (\sim 40, \sim 70) \) for binaries with \( q = (1/6, 1/10) \) and total mass \( M \in (7M_\odot, 11M_\odot) \) (see Figure 2).

Figure 6 conveys an important message — deriving the second order corrections to the radiative part of the self-force may well provide a robust framework to describe in a single model not only events that are naturally described by BHPT, such as the mergers of stellar mass compact objects with supermassive BHs in galactic nuclei \([51, 52, 53, 80–83]\), but also events that are more naturally described by PN or numerical methods, in particular the coalescences of compact object binaries with comparable or intermediate mass-ratios \([38, 39, 51, 84, 85]\).

To finish this Section, we describe the approach followed to construct the gravitational waveform from the inspiral trajectory. At leading PN order, a general inspiral waveform can be written as

\[
h(t) = -(h_+ - ih_x) = \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{\ell} h_{\ell m} Y_{\ell m}(\theta, \phi),
\]

If only the leading-order modes \( (\ell, m) = (2, \pm 2) \) and included, the inspiral waveform components are given by

\[
h_+(t) = \frac{4 \mu r^2 \dot{\phi}^2}{D} \left( 1 + \cos^2 \ell \right) \cos [2(\phi(t) + \Phi)],
\]

\[
h_\times(t) = \frac{4 \mu r^2 \dot{\phi}^2}{D} \cos \ell \sin [2(\phi(t) + \Phi)],
\]
FIG. 5. The (top, bottom) panels show the results of the optimization runs that were used to constrain the values of the \( b_i \) coefficients given in Eq. (21). The panels show the results for binaries of mass–ratio \( q \in [1/6, 1/8] \), and total mass \( M = [7M_\odot, 9M_\odot] \). The ‘optimal’ value for the coefficients has been chosen by ensuring that the flux prescription minimizes the phase discrepancy between the EOB model and our self-force model. The color bar shows the phase difference squared between both models, which is integrated from \( r = 20M \) all the way to the light-ring.

where \( D \) is the distance to the source. Since the orbital evolution will deviate from a circular trajectory during late inspiral (\( \dot{r} \neq 0 \)), we must consider more general orbits in which both \( \dot{r} \) and \( \dot{\phi} \) are non-negligible. For such orbits, the Newtonian GW polarizations are given by [58]:

\[
\begin{align*}
    h_+(t) &= \frac{2\mu}{D} \left\{ (1 + \cos^2 \iota) \left[ \cos [2(\phi(t) + \Phi)] \left( -\dot{r}^2 + r^2 \dot{\phi}^2 + \frac{1}{r} \right) \\
        &\quad + 2r \dot{r} \dot{\phi} \sin [2(\phi(t) + \Phi)] \right] + \left( -\dot{r}^2 - r^2 \dot{\phi}^2 + \frac{1}{r} \right) \sin^2 \iota \right\}, \\
    h_\times(t) &= \frac{4\mu}{D} \cos \iota \left\{ \sin [2(\phi(t) + \Phi)] \left( -\dot{r}^2 + r^2 \dot{\phi}^2 + \frac{1}{r} \right) \\
        &\quad - 2r \dot{r} \dot{\phi} \cos [2(\phi(t) + \Phi)] \right\},
\end{align*}
\]

(30)

(31)

where \( \dot{r} \) can be computed using

\[
\frac{dr}{dt} = -\frac{1}{u^2} \frac{du}{dx} \frac{dx}{dt}.
\]
Having described the construction of the inspiral evolution, we shall now describe the approach followed to smoothly connect the late inspiral evolution onto the plunge phase. The adiabatic prescription given by Eq. (26) breaks down when $dE/dx \to 0$. Hence, we need a scheme that enables us to match the late inspiral phase onto the plunge phase. We will do this by modifying the “transition” phase developed by Ori and Thorne [4] by including finite mass–ratio corrections.

### D. Transition and plunge phases

In this Section we describe an extension of the transition phase model introduced by Ori and Thorne [4]. The basic idea behind this approach can be understood by studying the motion of an inspiralling object in terms of the effective potential, $V(r, L)$, which takes the following simple form for a Schwarzschild BH [87]:

\[
V(r, L) = \left(1 - \frac{2}{r}\right) \left(1 + \frac{L^2}{r^2}\right). \tag{32}
\]

Throughout the inspiral, the body moves along a nearly circular orbit, and hence the ratio of the energy flux to the angular momentum flux is given by:

\[
\frac{dE}{d\tau} = \Omega \frac{dL}{d\tau}. \tag{33}
\]
Hence, near the ISCO, the energy and angular momentum of the body satisfy the following relations:

\[ E \rightarrow E_{\text{ISCO}} + \Omega_{\text{ISCO}} \xi, \tag{34} \]
\[ L \rightarrow L_{\text{ISCO}} + \xi. \tag{35} \]

Re-writing the effective potential, Eq. (32), in terms of \( \xi = L - L_{\text{ISCO}} \), one notices that during early inspiral, \( \xi \gg 0 \), the motion of the object is adiabatic, and the object sits at the minimum of the potential— as shown in the left panel of Figure 7. However, as the object nears the ISCO, the minimum of the potential moves inward due to radiation reaction. At some point, the body’s inertia prevents the body from staying at the minimum of the potential, and adiabatic inspiral breaks down [4]— illustrated in the right-hand panel of Figure 7.

\[ \frac{d{\xi}}{d\tau} = \kappa \eta, \quad \text{with} \quad \kappa = \left[ \frac{32}{5} \frac{\Omega_{\text{ISCO}}^{7/3}}{\sqrt{1 - 3u_{\text{ISCO}}}} \right], \tag{37} \]

where \( \dot{\xi} \) is the general relativistic correction to the Newtonian, quadrupole-moment formula [63]. We now extend these Eqs. by including finite mass-ratio corrections. Eq. (36) can be replaced by

\[ \frac{dx}{dt} = \frac{u^2(1 - 2u)}{E(x)} \left( \frac{du}{dx} \right)^{-1} \left[ E(x)^2 - V(u(x), L(x)) \right]^{1/2}, \tag{38} \]

where we have used

\[ \frac{d\tau}{dt} = \frac{1 - 2u(x)}{E(x)}, \tag{39} \]

and the expressions for the energy and angular momentum are given by Eqs. (13), (14). In order to linearize Eq. (38) we replace \( E(x) \) and \( L(x) \) by Eqs (34) and (35) respectively.

As discussed in [4], since these equations use the \( \eta \)-corrected values for \( E(x_{\text{ISCO}}), L(x_{\text{ISCO}}) \) and \( \Omega_{\text{ISCO}} \), then they remain valid even for finite mass-ratio \( \eta \). In Figure 8 we show the effect that these finite mass-ratio \( \eta \) corrections have on the effective potential \( V(x, L(x)) \). We determine the point at which the transition regime starts by carrying
FIG. 8. The left panel shows the effective potential for a Schwarzschild BH without including finite mass-ratio corrections. Note that the minimum of the potential takes place at the ISCO, which can be determined using Eq. (18). The right panel exhibits the influence of finite mass-ratio corrections on the effective potential used to modify Ori and Thorne transition regime [4]. The curves represent binaries, from top to bottom, with mass-ratios $q \in [0, 1/100, 1/20, 1/10, 1/6, 1/5, 1]$. 

out a stability analysis near the ISCO using $dE/dx$. As shown in Figure 4, the ISCO is determined by the relation $dE/dx = 0$. We have found that the relation

\[
\left( \frac{dE}{dx} \right)_{\text{transition}} = -0.054 + \frac{1.757 \times 10^{-4}}{\eta},
\]

(40)

provides a robust criterion to mark the start of the transition regime for binaries with mass-ratios $1/100 < q < 1/6$.

In [62], the authors only kept terms linear in $\xi$, but we have explored which higher order terms had a noticeable impact on the evolution by examining their impact on the length and phasing of the waveform. We found that terms $\propto \xi$ and $\propto (u - u_{\text{ISCO}})\xi$ were important, but corrections at order $\mathcal{O}(\xi^2)$ could be ignored even for comparable mass-ratio systems.

We model the evolution of the orbital frequency during the transition regime and thereafter in a different manner to that proposed by Ori and Thorne [4]. In order to ensure that the late-time evolution of the orbital frequency of our self-force model is as close as possible to the orbital evolution extracted from numerical relativity simulations, we incorporate the late-time frequency evolution that was derived by Baker et al [5] in their implicit rotating source (IRS) model, namely:

\[
\frac{d\phi}{dt} = \Omega_i + (\Omega_f - \Omega_i) \left( 1 + \tanh\left(\ln \sqrt{\kappa + \left( t - t_0 \right)/b} \right) \right)^\iota,
\]

(41)

where $\Omega_i$ is the value of the orbital frequency when the transition regime begins, and $\Omega_f$ is the value of the frequency at the light ring, which corresponds to $\omega_{\ell mn}/m$, where $\omega_{\ell mn}$ is the fundamental quasi-normal ringing frequency ($n = 0$) for the fundamental mode $(\ell, m) = (2, 2)$ of the post-merger black hole (see Eq. (45) below). The constant mass-dependent coefficient $t_0$ is computed by ensuring that $d\Omega/dt$ peaks at a time $t = t_0$. The parameter $\kappa$ is computed by enforcing continuity between the first order time derivative of the orbital frequency as predicted by the self-force evolution —Eq. (5)— and that given by the first order time derivative of Eq. (41).

At the end of the plunge phase, we match the plunge waveforms, which are generated using Eqs. (30), (31), onto the $l = m = 2, n = 0$ ringdown mode since this dominates the ringdown radiation. In the following Section we will describe in detail the procedure followed to attach the ringdown waveform.

After the transition regime, the equations of motion we use to model the plunge phase are: the second order time derivative of Eq. (38) which gives the radial evolution, and Eq. (41) which describes the orbital frequency evolution. We determine the point at which to attach the plunge phase by integrating Eq. (38) backwards in time, and finding the point at which the transition and plunge equations of motion smoothly match.

In Figures 9 and 10 we show the evolution obtained by combining Ori and Thorne’s [4] transition approach for the radial motion with the frequency evolution proposed by Baker et al [5]. In all the cases shown in Figures 9 and 10 the orbital frequency peaks and saturates at a value given by $\omega_{\ell mn}/m$. This can be understood if we analyze the
FIG. 9. (Top, bottom) panels: the left panel shows the inspiral, transition and plunge radial evolution for a BH binary of mass-ratio $q = (1/6, 1/8)$ — and total mass $M \in (7M_\odot, 9M_\odot)$ — using the coordinate transformation given by Eq. (15). The right panel shows the orbital frequency $M\Omega$ from late inspiral all the way to the light ring. The evolution starts from an initial radial value $r = 30M$.

asymptotic behavior of Eq. (41) near the light-ring, i.e.,

$$\frac{d\phi}{dt} \approx \Omega_t - (\Omega_t - \Omega_i) e^{-2(t-t_0)/b}. \quad (42)$$

Recasting Eq. (41) in this form, enables us to identify the constant coefficient $b$ with the e-folding rate for the decay of the fundamental quasinormal mode (QNM). As discussed previously, Eq. (41) predicts the expected orbital evolution during late inspiral and onward. To provide a unified description from late inspiral through to ringdown, we have decided to adopt the IRS approach, since this framework allows us to smoothly transition from late inspiral onto the plunge phase, and finally describe the ringdown waveform as a natural consequence of the IRS strain-rate amplitude decay relation $A^2(t) \propto \Omega \dot{\Omega}$. In other words, since Eq. (42) has the correct behavior near the light-ring as predicted by BHPT, the IRS model provides a natural framework to attach the ringdown waveform at the end of the plunge phase. We will discuss this feature in further detail in the following Section.

E. Ringdown Waveform

Numerical relativity simulations have shown that coalescing binary BHs in general relativity lead to the formation of a distorted rotating remnant, which radiates GWs while it settles down into a stationary Kerr BH [88, 89]. The GWs emitted during this intermediate phase resemble a ringing bell. Hence, this type of radiation is commonly known
in the literature as ringdown radiation, and consists of a superposition of QNMs — first discovered in numerical studies of the scattering of GWs in the Schwarzschild spacetime by Vishveshwara [90]. QNMs are damped oscillations whose frequencies and damping times are uniquely determined by the mass and spin of the post-merger Kerr BH. The frequency $\hat{\omega}$ of each QNM has two components: the real part represents the oscillation frequency, and the imaginary part corresponds to the inverse of the damping time:

$$\hat{\omega} = \omega_{\ell mn} - i/\tau_{\ell mn}. \quad (43)$$

As discussed above, the observables $\omega_{\ell mn}$, $\tau_{\ell mn}$ are uniquely determined by the final mass, $M_f$, and final spin, $q_f$, of the post-merger Kerr BH. To determine $M_f$, we use the analytic phenomenological expression for the final mass of the BH remnant that results from the merger of generic binary BHs on circular quasi-orbits introduced in [91], namely,

$$\frac{M_f}{M} = 1 - \left(1 - \frac{2\sqrt{2}}{3}\right) \eta - 0.543763 \eta^2. \quad (44)$$

This expression reproduces the expected result in the test mass particle limit, and also reproduces results from currently available numerical relativity simulations [91, 92]. We determine the final spin of the BH remnant $q_f$ using the fit proposed in [93]:

$$q_f = \sqrt{12} \eta + s_1 \eta^2 + s_2 \eta^3, \quad (45)$$
with:

\[ s_1 = -3.454 \pm 0.132, \quad s_1 = 2.353 \pm 0.548. \]  

(46)

This prescription is consistent with the numerical relativity simulations described in \([92, 93]\), and reproduces test mass limit predictions. This compact formula is also consistent with the prescriptions introduced in \([94, 95]\). The largest discrepancy between Eq. (45) and those derived in \([94, 95]\) is \(\lesssim 2.5\%\) for binaries with \(q \lesssim 1/6\). The ringdown waveform is given by

\[ h(t) = - (h_+ - i h_{\times}) = \frac{M}{D} \sum_{\ell m n} A_{\ell mn} e^{-i(\omega_{\ell mn} t + \phi_{\ell mn})} e^{-t/\tau_{\ell mn}}, \]  

(47)

where \(A_{\ell mn}\) and \(\phi_{\ell mn}\) are constants to be determined by smoothly matching the plunge waveform onto the subsequent ringdown. The ringdown portion of the self-force waveform model constructed in this paper includes the mode \(\ell = m = 2\) and the tones \(n = 0, 1, 2\). The approach we follow to attach the leading mode and overtones is the following:

- In order to ensure continuity between the plunge and ringdown waveforms, we use the end of the plunge waveform — Eqs. (30), (31) — to construct an interpolation function \(F(t)\). The interpolation method used to construct \(F(t)\) is a cubic spline.
- We match the plunge waveform onto the leading mode \(\ell = m = 2, n = 0\) of the ringdown waveform, Eq. (47), at the point where the amplitude of the plunge waveform peaks, \(t_{\text{max}}\). Attaching the mode requires \(F(t = t_{\text{max}})\) and \(F'(t = t_{\text{max}})\) which are computed from the interpolation function. These conditions fix two constants per polarisation.
- To attach the first overtone, \(\ell = m = 2, n = 1\), we insert into Eq. (47) the constants determined by attaching the leading mode as seeds to compute the amplitude and phase coefficients for the first overtone by enforcing continuity at \(t_{\text{max}} + dt\).
- Finally, we insert into Eq. (47) the value of the amplitude and phase coefficients previously determined for the leading mode and first overtone, and determine the four remaining constants by enforcing continuity at \(t_{\text{max}} + 2 dt\).

Having described the methodology followed to construct complete waveforms for comparable and intermediate mass-ratio systems, we finish this Section by putting together all these various pieces to construct sample waveforms for a few systems with mass-ratio \(q \in [1/6, 1/8, 1/10, 1/15]\), and total mass \(M \in (7M_\odot, 9M_\odot, 11M_\odot, 16M_\odot)\) in Figure 11.

**F. Ringdown waveform construction in the context of an Implicit Rotating Source**

Having described the ringdown waveform construction as a sum of quasinormal modes, we finish this Section by exhibiting the power of the IRS model to describe the ringdown evolution. In the IRS model, the strain-rate amplitude decay is given by \([3]\):

\[ A^2(t) = 16 \pi \dot{E} \approx 16 \pi \xi \Omega \dot{\Omega}. \]  

(48)

Using Eq. (42), in the limit \(\Omega \to \Omega_f\), the amplitude decay is given by

\[ A_0^2 e^{-2t/\tau} \approx \frac{32 \pi \xi \Omega_f}{b} (\Omega_f - \Omega_l) e^{-2(t-t_0)/\tau}. \]  

(49)

Hence, the late-time amplitude in the IRS model is given by

\[ A_{\ell m}^2 \approx 16 \pi \xi_{\ell m} \omega_{\ell m} \omega_{\ell m}, \]  

(50)

where the value of \(\xi_{\ell m}\) is set by ensuring amplitude continuity at the light-ring. In Figure 12 we explicitly show the equivalence of the ringdown waveform construction both in the IRS approach and using the sum of QNMs utilized in the previous Section. This detailed analysis shows that the IRS is a powerful tool to model the late time portions of the waveforms in a unified way.
FIG. 11. The panels show sample waveforms from inspiral to ringdown for systems with mass-ratios $q \in [1/6, 1/8]$ —and total mass $M \in (7M_\odot, 9M_\odot)$— (top panels—from left to right) and $q \in [1/10, 1/15]$ —and total mass $M \in (11M_\odot, 16M_\odot)$— (bottom panels —from left to right). The inspiral evolution for the [top, bottom] panels starts from $r = [30M, 25M]$.

G. Summary

In this Section we briefly summarize the key ingredients that were used to develop the waveform model described in this paper:

- **Inspiral evolution**
  - The building blocks of the inspiral evolution are the expressions for the energy, $E$, and angular momentum, $L$, — Eqs. (13) and (14) — that include gravitational self-force corrections and are valid over the domain $0 < x < \frac{1}{3}$.
  - The orbital frequency is modeled using Eq. (5). This prescription encapsulates gravitational self-force corrections that render a better phase evolution when calibrated against EOB.
  - The inspiral trajectory is modeled using the simple prescription (26). This scheme is no longer valid near ISCO, where $dE/dx = 0$ for binaries with $q \leq 6$.
  - We construct the inspiral waveform using Eqs. (30), (31).

When the inspiralling object nears the ISCO, we need to invoke the ‘transition scheme’ introduced by Ori and Thorne [4], which enables us to smoothly attach the late inspiral evolution onto the plunge phase.

- **Transition phase**
  - The transition regime starts at a point when $dE/dx$ satisfies Eq. (40).
FIG. 12. The panels show sample the late-time evolution of waveforms whose ringdown phase is modeled using the implicit rotating source (IRS) model and a sum of quasinormal modes (QNMs). The systems shown correspond to binaries with mass-ratios $q \in [1/6, 1/8]$ —and total mass $M \in (7M_\odot, 9M_\odot)$ — (top panels— from left to right) and $q \in [1/10, 1/15]$ —and total mass $M \in (11M_\odot, 16M_\odot)$ — (bottom panels —from left to right). The inspiral evolution for the [top, bottom] panels starts from $r = [30M, 25M]$.

- The equations of motion that govern the transition phase are Eq. (34), (35). These relations are valid, since the motion near the ISCO is nearly-circular.
- Using Eqs. (34), (35), we linearize the second order time derivative of Eq. (38).
- In order to reproduce the orbital phase evolution predicted by numerical simulations from the ISCO to the light-ring, we modify the original transition phase by smoothly matching the inspiral orbital phase evolution, Eq. (5), onto the IRS model, Eq. (41) at the start of the transition phase.

- Plunge phase

- The equations of motion that govern the plunge phase are given by the second order time derivative of Eq. (38), and Eq. (41).
- We integrate these relations backwards in time to find the point at which both the transition and plunge equations of motion smoothly match. The transition phase ends at this point.
- Near the light-ring Eq. (41) has the behavior predicted by BHPT, which enables us to smoothly match the plunge phase onto the ringdown.
- Both the transition and plunge waveforms are constructed using Eqs. (30) and (31).

Having constructed the waveform from inspiral to ringdown, we derived second-order radiative corrections to improve the radiative evolution of the waveform model. This was necessary in order to construct a waveform model...
that is internally consistent, i.e., since the orbital elements include first-order conservative corrections, then radiative corrections should enter the fluxes at second order. We have derived these corrections by enforcing a close agreement between our self-force model and EOB, and have shown that our model can reproduce the orbital phase evolution predicted by EOB within the numerical error of the simulations used to calibrate this model for a variety of mass-ratios.

- Ringdown phase
  - The ringdown waveform is constructed using Eq. (14).
  - We use the plunge waveform to construct an interpolation function $F(t)$, and then use this function to attach the leading mode $\ell = m = 2$, $n = 0$ at the point where the amplitude of the plunge waveform peaks, $t_{\text{max}}$. We enforce continuity by ensuring that $F(t_{\text{max}}) = h_{\text{RD}}^{n=0}$ and $F'(t_{\text{max}}) = h_{\text{RD}}''^{n=0}$.
  - We include the first and second overtone $n = 1$, $n = 2$ in the ringdown waveform.
  - Using the IRS model, we have shown that having knowledge of the time evolution of the orbital frequency provides sufficient information to construct the amplitude decay during ringdown. Hence, we can construct an alternative ringdown waveform using only Eq. (50), and ensuring smooth continuity with the plunge waveform.
  - Finally, we have explicitly shown that the implicit rotating source approach provides a natural transition from late-time radiation to ringdown that is equivalent to ringdown waveform modeling based on a sum of QNMs.

Throughout the paper we have emphasized the fact that our model provides a more reliable framework to model binaries whose components are non-spinning, and with mass-ratios $q \leq 1/6$, as compared to available PN approximants. It is worth emphasizing that our model is also computationally inexpensive. All the waveforms we generated to constrain the higher-order $\eta$ corrections in the energy flux —Eq. (22)— can be generated in fractions of a second. A direct comparison between our code and EOB shows that, averaged over 500 realizations, our code is $\sim 20\%$ faster than the optimized version of the EOB code currently available in the LIGO Scientific Collaboration LAL library. It should be emphasized, though, that our code at present has not been optimized, and hence, compared to EOB our model is expected to significantly reduce the cost of waveform generation, making it relatively more viable for parameter estimation efforts. This is a key feature in our model that enabled us to sample a wide region of parameter space to constrain the $B_\ell$ coefficients in Eq. (22). Furthermore, these self-forced waveforms do not need to be highly sampled near the light-ring, hence decreasing the speed with which they can be generated, because the prescription we have used to model the late-time orbital frequency evolution provides the correct evolution near the light-ring. This particular feature also enables us to match the plunge waveform onto its ringdown counterpart without having to interpolate the orbital frequency evolution using a phenomenological approach. The model is internally consistent and the only phenomenology invoked during its construction is related to currently unknown physics, i.e., higher order radiative corrections to the energy flux. Once these corrections are formally derived in the near future, the flexibility of our model will enable us to replace the radiative corrections that we have currently computed by numerical optimization. At that stage, we will be able to describe in a single unified model the dynamical evolution of binaries whose mass-ratios range from the extreme to the comparable regime.

III. CONCLUSIONS

In this paper we have developed a self-force waveform model that is capable of reproducing with good accuracy the inspiral, merger and ringdown of binaries with mass-ratios $q \leq 1/6$. This work suggests that a model that incorporates first order conservative self-force corrections in the orbital elements, and second-order radiative corrections in the dissipative piece of the self-force may suffice to describe in a unified manner the coalescence of binaries with mass-ratios that range from the comparable to the extreme. Our model includes conservative self-force corrections that have been derived in the strong-field regime [1]. Using these results, we find that binaries with mass-ratios $q \gtrsim 1/6$ do not have an ISCO. For systems with $q \leq 1/6$, we have derived a simple relation that provides the location of the ISCO in terms of the symmetric mass-ratio (see Eq. (13)). To describe the inspiral evolution, we have derived second-order corrections to the energy flux by minimizing the phase discrepancy between our self-force model and the EOB model [2, 3] for a variety of mass-ratios. We have shown that our model reproduces the phase evolution of the EOB model within the accuracy of available numerical simulations for a variety of mass-ratios.

This paper also presents an extension of the “transition regime” developed by Ori and Thorne [4] to smoothly match the late inspiral evolution onto the plunge phase. We have found that using the inspiral phase expression for the orbital frequency during the plunge phase does not render an accurate description of the actual orbital evolution as predicted by numerical simulations. Hence, we have embedded the self-force framework in the IRS model proposed
by Baker et al.\(^5\) to ensure that our model is as close as possible to the orbital evolution predicted by numerical relativity simulations. The implementation of this prescription ensures that the orbital frequency saturates near the light-ring, which facilitates matching onto the ringdown phase. We have shown that the IRS model provides a natural transition onto the ringdown phase that is equivalent to a ringdown waveform construction based on a sum of QNMs.

The motivation for constructing this model is two-fold: to exhibit the versatility of the self-force formalism to accurately describe the evolution of binaries beyond the extreme mass-ratio limit; and to provide a tool that can be used to explore the information that could be extracted by GW detectors that target binaries with comparable and intermediate mass-ratios. Current studies have only explored the use of PN approximants to model the merger of NSBH binaries, despite their inadequacy to capture the evolution of these systems.\(^6\) Comparing Figures 2 and 6 we conclude that even if second generation GW detectors were only capable of capturing the inspiral evolution of NSBH mergers, our self-force model would be better equipped to describe these events. The construction of this IRS self-force model constitutes an important step towards the construction of more reliable waveforms to describe IMRCs. In \(^50\), it was shown that Huerta–Gair (HG) waveforms — which are closely related to the model developed in this paper, and which were also based on a consistent combination of BHPT and self-force corrections — work in the relevant mass range for advanced detectors. The new model introduced in this article, improves upon these waveforms and should therefore not only work in the regime of interest to advanced detectors but over a wider parameter space that could be applicable to third generation detectors or later observations.

Having developed a strong foundation to model binaries on circular orbits whose components are not rotating, it is necessary to incorporate more ingredients to the waveform model to capture GW signals from binaries whose components have significant spin \(^2, 62, 82, 97, 101\), or systems that form in core-collapsed globular clusters, and hence are expected to have non-negligible eccentricity at merger.\(^102, 103\). In order to do so, we require input both from the self-force program — which is making substantial progress towards the computation of the self-force in a Kerr background \(^104, 100\) — and from numerical simulations \(^57\), in particular from binaries with typical mass-ratios \(1/20 < q < 1/10\) to probe the performance of our self-force model in this currently unconstrained regime. Undoubtedly, the development and implementation of improved numerical algorithms \(^107\) to carry out these simulations will facilitate the realization of these studies in the near future.

Recent observational discoveries \(^7\) suggest that NSBH mergers may also occur in globular clusters. Hence, in these dense stellar environments we may expect that multiple \(n\)-body interactions and binary exchange processes may lead to the formation of binaries on eccentric orbits. Detectors such as the Einstein Telescope, operating at low frequencies \(< 1\text{Hz}\), may be capable of detecting the signature of eccentricity during the early inspiral. In order to assess these effects, we aim to extend the model introduced in this paper by including eccentricity, making use of self-force corrections for generic orbits in a Schwarzschild background.\(^80\).

**ACKNOWLEDGMENTS**

EH thanks the hospitality of the TAPIR group at Caltech where part of this work was carried out, and the generosity of Prof Peter Saulson, who made this visit possible through his National Science Foundation (NSF) grant number PHY-120583. We are pleased to thank Chad Galley, Alex Huerta Gago, Anıl Zenginoğlu and Fan Zhang for useful interactions. JG’s work is supported by the Royal Society. PK acknowledges support from the NSF grant number PHY-0847611. Some calculations were performed on the Syracuse University Gravitation and Relativity cluster, which is supported by NSF grants PHY-1040231, PHY-1104371 and Syracuse University ITS.
