Surrogate model for gravitational wave signals from nonspinning, comparable-to large-mass-ratio black hole binaries built on black hole perturbation theory waveforms calibrated to numerical relativity

Tousif Islam[,1,2,3,4,*] Scott E. Field,[2,3] Scott A. Hughes[,5] Gaurav Khanna[,6,1,3] Vijay Varma[,7]
Matthew Giesler[,8] Mark A. Scheel,[9] Lawrence E. Kidder,[8] and Harald P. Pfeiffer[7]

[1]*Department of Physics, University of Massachusetts, Dartmouth, Massachusetts 02747, USA*
[2]*Department of Mathematics, University of Massachusetts, Dartmouth, Massachusetts 02747, USA*
[3]*Center for Scientific Computing and Visualization Research, University of Massachusetts,
Dartmouth, Massachusetts 02747, USA*
[4]*Kavli Institute of Theoretical Physics, University of California Santa Barbara,
Santa Barbara, California 93106, USA*
[5]*Department of Physics and MIT Kavli Institute, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, USA*
[6]*Department of Physics, University of Rhode Island, Kingston, Rhode Island 02881, USA*
[7]*Max Planck Institute for Gravitational Physics (Albert Einstein Institute),
Am Mühlenberg 1, Potsdam 14476, Germany*
[8]*Cornell Center for Astrophysics and Planetary Science, Cornell University, Ithaca, New York 14853, USA*
[9]*Theoretical Astrophysics, Walter Burke Institute for Theoretical Physics,
California Institute of Technology, Pasadena, California 91125, USA*

We present a reduced-order surrogate model of gravitational waveforms from nonspinning binary black hole systems with comparable to large mass-ratio configurations. This surrogate model, `BHPTNRSur1dq1e4`, is trained on waveform data generated by point-particle black hole perturbation theory (ppBHPT) with mass ratios varying from 2.5 to 10,000. `BHPTNRSur1dq1e4` extends an earlier waveform model, `EMRISur1dq1e4`, by using an updated transition-to-plunge model, covering longer durations up to $30,500m_1$ (where m_1 is the mass of the primary black hole), includes several more spherical harmonic modes up to $\ell = 10$, and calibrates subdominant modes to numerical relativity (NR) data. In the comparable mass-ratio regime, including mass ratios as low as 2.5, the gravitational waveforms generated through ppBHPT agree surprisingly well with those from NR after this simple calibration step. We also compare our model to recent SXS and RIT NR simulations at mass ratios ranging from 15 to 32, and find the dominant quadrupolar modes agree to better than $\approx 10^{-3}$. We expect our model to be useful to study intermediate-mass-ratio binary systems in current and future gravitational-wave detectors.

I. INTRODUCTION

Detection of gravitational waves (GWs) [1,2] from the coalescence of binary compact objects offer a new window to study black holes and fundamental physics. Most of the GW signals detected so far by the LIGO/Virgo collaboration are consistent with binary black holes (BBHs), with mass ratio $q = m_1/m_2 \leq 10$.[1] Another interesting source of GWs are intermediate mass ratio inspirals (IMRIs) comprised of an intermediate-mass black hole (IMBH, mass $\sim 10^2-10^4 \, M_\odot$) [3–5] and a solar-mass black hole (mass $\sim 3-20 \, M_\odot$). The resulting binaries will have a

mass ratio in the range $10 \leq q \leq 10^4$. The existence of IMBHs on the low-end of this mass range has been confirmed by the detection of GW190521 [5], and electromagnetic evidence continues to mount indicating the likely existence of these objects across their possible mass range [5]. IMRIs are expected to form in dense globular clusters and galactic nuclei [6,7]. While these binaries are a prime source for future-generation detectors such as LISA [8], Cosmic Explorer (CE) or the Einstein Telescope (ET) [9,10], current detectors may also be able to detect IMRIs as their sensitivity improves. In particular, IMRIs with total mass $< 2000 \, M_\odot$ may be detected by the current generation of detectors [11] while future space-based missions, [12,13] such as LISA will observe heavier binaries. LISA is also expected to detect GW signals from extreme mass ratio inspirals (EMRIs) (having a mass ratio

---

*[*]tislam@umassd.edu

[1]We use the convention $q = m_1/m_2$, where m_1 and m_2 are the masses of the component black holes, with $m_1 \geq m_2$.

$q \gtrsim 10^5$) that consists of a stellar-mass black hole paired with a supermassive black hole [9,10].

Detection of GWs from IMRIs would shed light on many issues in both astrophysics and fundamental physics [9,10,13]. IMRIs may form in dense globular clusters or galactic nuclei through multiple possible formation channels. Detection and parameter inference with IMRIs [14–17] will help us probe formation channels and the evolutionary pathway to supermassive black holes [18]. As these systems form in dense environments, GW signals from IMRIs may carry an imprint of its surrounding environment and are ideal sources to investigate possible environmental effects [9,19–25]. IMRI signals could also be used to test the nature of gravity in an unexplored strong-field regime [26–32], complementing tests of general relativity (GR) performed with GW events detected so far [33].

Carrying out accurate parameter inference and performing fundamental physics analysis with IMRI GW signals will require multimodal waveform models that are fast and reliable [34]. While numerical relativity (NR) provides the most accurate waveforms from BBH mergers, it takes weeks to months to generate a single waveform, making them unfit to be directly used in multiquery studies. The availability of a large number of NR simulations in the comparable mass regime, however, has paved the way to build either NR-based reduced-order surrogate models [35–40] or calibrated phenomenological or effective-one-body (EOB) models [41–49] to NR. Only a few of these models [50], however, are tested or calibrated in the intermediate mass ratio regime, where only a few NR simulations are available [51–54]. In lieu of NR-based calibration, some EOB models [55] have been tuned to results from point particle black hole perturbation theory (ppBHPT), which provides an accurate waveform model as $q \to \infty$. Due to computational cost, ppBHPT waveform models also cannot be directly used in multiquery data analysis. This has been partly overcome by developing "kludge" models [56–59] that are fast and capture the qualitative features of an EMRI waveform in the inspiral regime using approximations for the amplitudes and phases. Recently Refs. [60,61] have introduced gravitational self-force-based waveform models that are as fast to compute as kludge models by using a combination of reduced order methods, deep-learning techniques and hardware acceleration. Also relevant is Ref. [62], which presents a fully relativistic second-order self-force model that can generate first-principles inspiral waveforms in milliseconds, at least for the case of quasicircular inspiral of nonspinning black holes.

To begin addressing these issues, Ref. [63] built a proof-of-principle ppBHPT surrogate model `EMRISur1dq1e4` for nonspinning binaries that extends from mass ratio $q = 3$ to $q = 10,000$ and covers $\sim 13,500m_1$ in duration. The model's dominant mode has been tuned to NR in the

comparable mass ratio regime ($q \leq 10$), and it was shown that after this simple calibration step the ppBHPT and NR waveforms agreed to better than $\approx 1\%$ at mass ratios $q \gtrsim 8$. These initial encouraging results suggest that suitably calibrated ppBHPT waveform data could provide for an accurate model of gravitational waves from IMRI systems. The agreement between NR and ppBHPT (with radiative corrections to the orbit) after a simple rescaling [63] is a surprising observation on its own.

In this paper, we describe more fully the methods we have used to build `EMRISur1dq1e4` as well as making numerous important improvements to the underlying model. The updated version of our surrogate model—which we call `BHPTNRSur1dq1e4`—is $\sim 30,500m_1$ in duration and covers all phases of the system's evolution from inspiral through plunge and ringdown—making it suitable to be used in a wider range of data analysis studies. It features a total of 50 important higher order modes up to $\ell = 10$ thereby permitting studies to quantify the effect of higher order modes in GW signals. Furthermore, by applying a simple calibration, we find the NR-calibrated ppBHPT waveforms agree remarkably well with NR for all of the higher order modes up to $\ell = 5$ in the comparable mass ratio regime.

The rest of the paper is organized as follows. Section II describes our method for computing ppBHPT waveforms by solving the Teukolsky equation. We describe the surrogate-modelling framework, calibration to NR, and assess model accuracy in Sec. III. Section IV provides a more detailed comparison between ppBHPT waveforms and NR data in the comparable and intermediate mass ratio regime with a focus on subdominant modes and new SXS and RIT simulations at mass ratios greater than 10. Finally, we outline future directions in Sec. V.

II. WAVEFORM DATA USING PERTURBATION THEORY

We generate the surrogate-model training data using point-particle black hole perturbation theory (ppBHPT). First, we compute the trajectory taken by the point-particle and then we use that trajectory to compute the gravitational wave emission. The next three subsections summarize the equations and algorithms for accomplishing this.

A. Numerically solving the Teukolsky equation

In the ppBHPT framework, the smaller black hole is modeled as a point-particle with no internal structure and a mass of m_2, moving in the spacetime of the larger Kerr black hole with mass m_1 and spin angular momentum per unit mass a. Here, we provide an executive summary of this framework and refer to Refs. [64–67] for additional details.

Gravitational radiation is computed by first numerically solving the Teukolsky equation

$$- \left[\frac{(r^2 + a^2)^2}{\Delta} - a^2 \sin^2\theta \right] \partial_{tt} \Psi - \frac{4 m_1 a r}{\Delta} \partial_{t\phi} \Psi$$

$$- 2s \left[r - \frac{m_1(r^2 - a^2)}{\Delta} + i a \cos\theta \right] \partial_t \Psi$$

$$+ \Delta^{-s} \partial_r (\Delta^{s+1} \partial_r \Psi) + \frac{1}{\sin\theta} \partial_\theta (\sin\theta \partial_\theta \Psi)$$

$$+ \left[\frac{1}{\sin^2\theta} - \frac{a^2}{\Delta} \right] \partial_{\phi\phi} \Psi + 2s \left[\frac{a(r - m_1)}{\Delta} + \frac{i \cos\theta}{\sin^2\theta} \right] \partial_\phi \Psi$$

$$- (s^2 \cot^2\theta - s)\Psi = -4\pi (r^2 + a^2 \cos^2\theta) T, \qquad (1)$$

sourced by the moving particle, where $\Delta = r^2 - 2m_1 r + a^2$ and s is the "spin weight" of the field. The $s = -2$ case for Ψ describes the radiative degrees of freedom of the gravitational field, the Weyl scalar ψ_4, in the radiation zone, and is directly related to the Weyl curvature scalar as $\Psi = (r - i a \cos\theta)^4 \psi_4$. The source term T in Eq. (1) for the smaller compact object m_2 is related to the energy-momentum tensor $T_{\alpha\beta}$ of a point particle. The Weyl scalar ψ_4 can then be integrated twice at future null infinity $\mathcal{I}^+$ to find the two polarization states h_+ and $h_\times$ of the transverse-traceless metric perturbations,

$$\psi_4 = \frac{1}{2} \left(\frac{\partial^2 h_+}{\partial t^2} - i \frac{\partial^2 h_\times}{\partial t^2} \right). \qquad (2)$$

The complex gravitational wave strain

$$h_+(t, \theta, \phi; q) - i h_\times(t, \theta, \phi; q)$$

$$= \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{\ell} h^{\ell m}(t; q)_{-2} Y_{\ell m}(\theta, \phi), \qquad (3)$$

can be formed from the two polarization states, which is subsequently decomposed into a basis of spin-weighted spherical harmonics $_{-2}Y_{\ell m}$. We build models for the harmonic modes $h^{\ell m}(t; q)$.

Once the trajectory of the perturbing compact body is fully specified (cf. Sec. II B), we solve the inhomogeneous Teukolsky equation in the time-domain while feeding the trajectory information into the particle source-term of the equation [64–68]. This involves a four step procedure: (i) rewriting the Teukolsky equation using compactified hyperboloidal coordinates (Eq. (1) is shown using standard Boyer-Lindquist coordinates) that allow us to extract the gravitational waveform directly at null infinity while also solving the issue of unphysical reflections from the artificial boundary of the finite computational domain; (ii) obtaining a set of $(2 + 1)$ dimensional PDEs by using the axisymmetry of the background Kerr space-time, and separating the dependence on azimuthal coordinate; (iii) recasting these equations into a first-order, hyperbolic PDE system; and finally (iv) implementing a high-order WENO (3,5) finite-difference scheme with Shu-Osher (3,3)

time-stepping [68]. The point-particle source term on the right-hand-side of the Teukolsky equation requires some specialized techniques for a finite-difference numerical implementation [64,65]. We set the spin of the central black hole to a value slightly away from zero, $a/m_1 = 10^{-8}$ for technical reasons.[2] Our numerical evolution scheme is implemented using OpenCL/CUDA-based GPGPU-computing which allows for very long duration and high-accuracy computations within a reasonable time-frame. Numerical errors in the phase and amplitude are typically on the scale of a small fraction of a percent [63,69].

B. Trajectory model

The particle's motion is characterized by three distinct regimes—an initial adiabatic inspiral, a late-stage geodesic plunge into the horizon, and a transition regime between those two.

During the initial adiabatic inspiral, the particle follows a sequence of geodesic orbits driven by radiative energy and angular momentum losses. The flux radiated to null infinity and through the event horizon are computed by solving the frequency-domain Teukolsky equation [70–73] using the open-source code GREMLINEQ [74,75] from the Black Hole Perturbation Toolkit [76]. The inspiral trajectory is then extended to include a plunge geodesic and a smooth transition region following a procedure similar to one proposed by Ori-Thorne [77]. We compute the transition between initial inspiral and the plunge using a generalized Ori-Thorne algorithm [78,79] (hereafter, the "GOT" algorithm). The GOT algorithm uses a parametrization of strong-field Kerr orbits based on *Mino time*, which separates the radial and polar motions of Kerr black hole orbits. It also introduces a correction that smooths a rather sharp discontinuity in the evolution of an inspiral's integrals of motion as presented in the original Ori-Thorne model. Detailed discussion of this point is given in Sec. IV A 2 of Ref. [79]. The use of Mino time is not so critical for our analysis since the separation of radial and polar motions is not an issue for equatorial orbits, but smoothing of the integrals of the motion is of great importance. Note that we use "Model 2" from Ref. [79] for this smoothing.

Our trajectory model does not include the effects of the conservative or second-order self-force [80], although once these post-adiabatic corrections are known they could be easily incorporated to improve the accuracy of the inspiral's phase.

C. Waveform smoothing

At low mass ratios, the GOT transition trajectory produces small nonphysical oscillations in high-order modes of the waveforms. Because the GOT algorithm is

---

[2]For example, to avoid a change in the definition of the coordinates from Kerr to Schwarzschild.
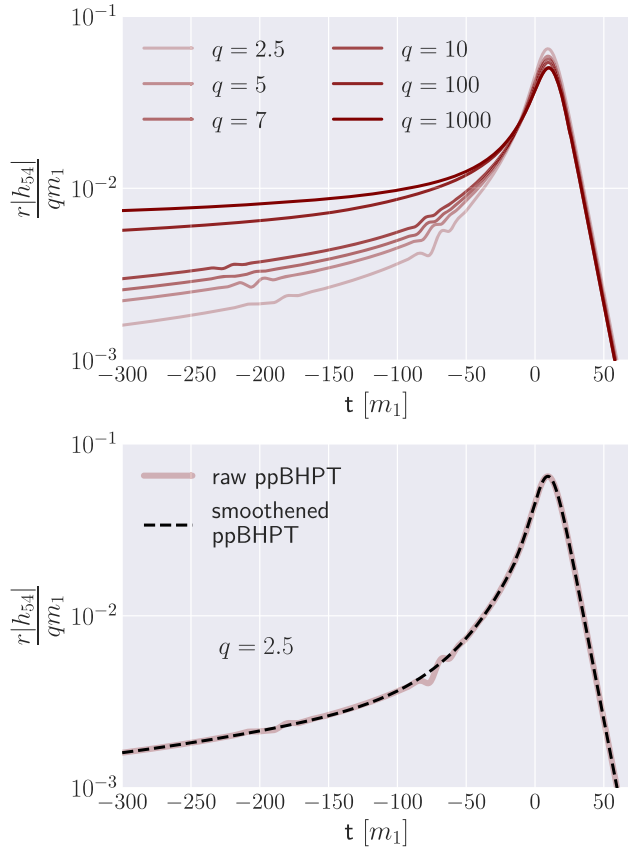
FIG. 1. At mass ratios $q \lesssim 10$ the generalized Ori-Thorne transition trajectory causes small, nonphysical oscillations in some of the $\ell \neq m$ modes. The raw ppBHPT waveform amplitudes are shown for one of the representative modes $(\ell, m) = (5, 4)$ for increasing value of the mass ratio q (upper panel) and one example waveform after performing the smoothing procedure (lower panel) described in Sec. II C.

designed for the regime $q \gg 1$, it is not surprising that some pathologies enter at low mass ratios. These nonphysical oscillations are results of a small jump in the acceleration of the point-particle as it exits the adiabatic inspiral and also when it begins the plunge. It is interesting, however, that these oscillations are more apparent in our data for modes with $\ell \neq m$. It is also worth noting that the oscillations are larger in amplitude when we use Model 1 from Ref. [79] for smoothing the evolution of the integrals of motion.

In Fig. 1 (upper panel), we show the unphysical oscillations in the scaled amplitudes in one of the representative modes $(\ell, m) = (5, 4)$ for increasing values of the mass ratios. It is clear that while $q \leq 10$ shows unphysical oscillations in the transition regime, these features vanish for high mass ratio simulations. At lower mass ratios we remove these unwanted oscillations by using a "smoothening" procedure [63]. To smooth the data, we (i) first remove the unphysical oscillatory portion of the waveform that we mention above, (ii) then use the rest of the

waveform data to construct a polynomial fit of degree 7, and finally (iii) evaluate the polynomial to obtain smoother data in the problematic regions. The lower panel of Fig. 1 shows the rescaled $(\ell, m) = (5, 4)$ amplitude for $q = 2.5$ before and after the 'smoothing'. A similar smoothing procedure is applied to the phase data. Our surrogate model is trained on—and for validation purposes, compared to—these smoothed waveform data.

To quantify the amount by which our smoothing procedure has modified the waveform, we compute a relative L_2-norm difference between the smoothed and original data. Normalized L_2-norm between two functions $h_1(t)$ and $h_2(t)$ is defined as a time-domain overlap integral with white-noise:

$$\mathcal{E}[h_1, h_2] = \frac{1}{2} \frac{\sum_{(\ell,m)} \int_{t_1}^{t_2} |h_1(t) - h_2(t)|^2 dt}{\sum_{(\ell,m)} \int_{t_1}^{t_2} |h_1(t)|^2 dt}, \quad (4)$$

Here, t_1 and t_2 denote the start and end of the waveform data respectively whereas h_1 and h_2 denote the smoothed and original data respectively. We find that the differences between the smoothed and original data for each mode is on average 8×10^{-5} with a maximum 5×10^{-4}. To compute errors for individual modes, we restrict the sum in Eq. (4) to only the mode of interest.

III. SURROGATE MODELING

In this section, we briefly describe the framework used to build the surrogate model. Our framework is constructed using a combination of methodologies proposed in earlier works [35,81,82].

A. Building the surrogate

a. Training data: To train the model, we collect a total of 41 ppBHPT waveforms by numerically solving the inhomogeneous Teukolsky equation at different values of the mass ratio q. These values of q have been chosen in such a way that it populates the mass ratio axis from $q = 2.5$ to $q = 10,000$ in a logarithmic scale (green circles in Fig. 2). For each value of q, we then extract the harmonic modes, $h^{\ell,m}(t; q)$. Note that we only model $m > 0$ modes as the negative m modes are computed from the positive m modes using the symmetry of the nonprecessing system under reflections about the orbital plane: $h^{\ell,-m} = (-1)^\ell h^{\ell,m*}$.

b. Data alignment: We first determine the peak of each waveform τ_{peak} to be the time when the quadrature sum,

$$A_{\text{tot}}(\tau) = \sqrt{\sum_{\ell,m} |h^{\ell,m}(\tau)|^2}, \quad (5)$$

reaches its maximum. Here the summation is taken over all the modes being modeled. In order to construct smooth parametric fits for the surrogate model, we align all the
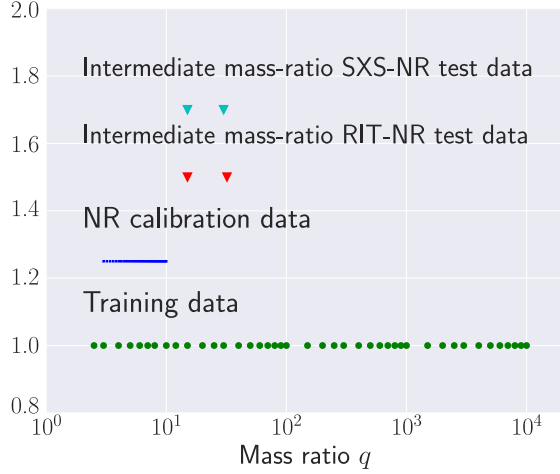
FIG. 2. ppBHPT waveform data points used in training `BHPTNRSur1dq1e4` (green circles). We also show points used in calibrating the ppBHPT waveforms to NR (blue squares) and data points used in validating NR calibration against high mass ratio NR simulation (red and cyan triangles).

waveforms such that their peaks occur at the same time. This is done by choosing a new time coordinate,

$$t = \tau - \tau_{\text{peak}}, \tag{6}$$

such that $A_{\text{tot}}(t)$ for each waveform peaks at $t = 0$. Next, we use cubic splines to interpolate the real and imaginary parts of the waveform modes onto a common time grid of $[-30500m_1, 115m_1]$ with a uniform time spacing of $dt = 0.1m_1$. Once all the waveforms are interpolated onto a common time grid, we perform a rotation about the z-axis such that at the start of the waveform $\phi_{22} = 0$ and $\phi_{21} \in [-\pi, 0]$, where ϕ_{22} and ϕ_{21} are the phases of the complex (2,2) and (2,2) modes, respectively. These pre-processing steps are necessary to ensure a smooth dependence of the training-set waveforms on mass ratio.

c. Data decomposition: After time and phase alignments, we decompose the inertial frame waveform modes into *waveform data pieces* that are slowly varying functions of time and are therefore simpler to model. We employ different decomposition strategy for the quadrupolar mode and the higher-order modes. The complex (2,2) waveform mode,

$$h_{22} = A_{22}e^{-i\phi_{22}}, \tag{7}$$

is decomposed into an amplitude, A_{22}, and phase, ϕ_{22}. For the higher order modes, we first apply a time-dependent rotation given by the instantaneous orbital phase ϕ_{orb} to transform the waveform into a co-orbital frame:

$$h^C_{\ell m} = h_{\ell m}e^{im\phi_{\text{orb}}}, \tag{8}$$

where h^C_{lm} represents the complex modes in the co-orbital frame and the orbital phase is taken to be

$$\phi^{\text{orb}} \equiv \frac{\phi_{22}}{2}. \tag{9}$$

We then use the real and imaginary parts of $h^C_{\ell m}$ as our waveform data pieces for the nonquadrupole modes. To summarize, the full set of waveform *data pieces* we model is as follows: A_{22}, ϕ_{22} for the (2,2) mode, and real and imaginary parts of $h^C_{\ell m}$ for the 24 higher order modes with $m > 0$.

d. Empirical interpolants: The next step is to construct an empirical interpolant (EI) in time using a greedy algorithm that picks the most representative time nodes [81,83–85]. The number of the time (or EI) nodes for each data piece is equal to the number of basis functions used. The empirical interpolant gives a compact representation for each data piece (and hence the full waveform) in the training set by permitting the full time-series to be reconstructed through a significantly sparser sampling defined by the EI nodes. We choose 7 basis functions for A_{22} and ϕ_{22}. For higher order modes, we use 13 basis functions for the real and imaginary parts if $\ell \leq 5$. Otherwise, 16 basis functions are used. We inspect the basis functions visually to ensure they are free from noise. Furthermore, unlike recently built surrogate models [38], we put no restriction on the location of EI nodes as we did not find this to improve our model.

e. Parametric fits: The final surrogate-building step is to construct parametric fits for each data piece at each of the EI nodes over the one-dimensional parameter space defined by q. Following Ref. [63], we fit the data-pieces using second degree interpolating splines (with smoothing factor

TABLE I. Overview of the `EMRISur1dq1e4` and `BHPTNRSur1dq1e4` models. Both models used the smoothing procedure described in Sec. II C.

Model	Plunge model	Available positive modes (ℓ, m)	Waveform length	NR-Calibration
`EMRISur1dq1e4`	Ori-Thorne [77]	$(2, \{1, 2\})$, $(3, \{1, 2, 3\})$ $(4, \{2, 3, 4\})$, $(5, \{3, 4, 5\})$	$13500m_1$	$(2,2)$
`BHPTNRSur1dq1e4`	Generalized Ori-Thorne [78,79]	$(2, \{1, 2\})$, $(3, \{1, 2, 3\})$ $(4, \{2, 3, 4\})$, $(5, \{3, 4, 5\})$ $(6, \{4, 5, 6\})$, $(7, \{5, 6, 7\})$ $(8, \{6, 7, 8\})$, $(9, \{7, 8, 9\})$ $(10, \{8, 9\})$	$30500m_1$	$(2, \{1, 2\})$ $(3, \{1, 2, 3\})$ $(4, \{2, 3, 4\})$ $(5, 5)$

$s = 0.0005$) after performing a logarithmic transformation of q [38,86].

B. Evaluating the surrogate

To generate the `BHPTNRSur1dq1e4` surrogate model waveforms, we provide mass ratio q as input. We then evaluate the parametric fits for each waveform data pieces at each EI node at the requested value of q. Next, the empirical interpolant is used to reconstruct the surrogate prediction of the waveform data pieces as a dense time-series. We evaluate the surrogate models for the amplitude and phase of the (2,2) mode and combine them to get the complex strain as $h_{22}^S = A_{22}^S e^{-i\phi_{22}^S}$. For the nonquadrupole modes, we first evaluate the surrogate models for the real and imaginary parts of the co-orbital frame waveform data pieces $h_{\ell m}^{C,S} \approx h_{\ell m}^C$ and treat it as $h_{\ell m}^C$. Finally, we use Eqs. (7), (8), and (9) to obtain the surrogate prediction for the inertial frame strain $h_{\ell,m}^S$ for these modes. The full surrogate, h_S, is then written

$$h^S(t, \theta, \phi; q) = \sum_{\ell,m} h_{\ell,m}^S(t; q)_{-2}Y_{\ell m}(\theta, \phi), \quad (10)$$

where $h_{\ell,m}^S$ is the surrogate model prediction for each harmonic mode.

C. Surrogate errors

In this section we assess the accuracy of `BHPTNRSur1dq1e4` by performing some of the tests described in Ref. [37] using the relative L_2-type norm defined in Eq. (4). In our case, ppBHPT waveforms used in training are already aligned in time and phase, and the surrogate is expected to reproduce this alignment. Therefore, we compute the time-domain error $\mathcal{E}$ without any further time/phase shifts.

To assess the surrogate model's error, we compute three different types of errors. First, we build the surrogate using all 41 ppBHPT training waveforms and calculate the *training error* between the training waveform and surrogate prediction. This checks whether the surrogate model can accurately reproduce the training waveforms.

Next we perform a *leave-one-out cross-validation* study. In this study, we hold out one ppBHPT waveform from the training set and build a trial surrogate from the remaining 40 ppBHPT waveforms. We then evaluate the trial surrogate at the mass ratio corresponding to the held out data, and compare its prediction with the held-out ppBHPT waveform. We refer to these errors as *validation errors*. Validation errors represent conservative error estimates for the surrogate model's generalization error against ppBHPT. Since we have 41 ppBHPT waveforms, we build 41 trial surrogates for each error study and assess the model's ability to predict new waveforms it was not trained on. For boundary cases (i.e., for $q = 2.5$ and $q = 10,000$), the test
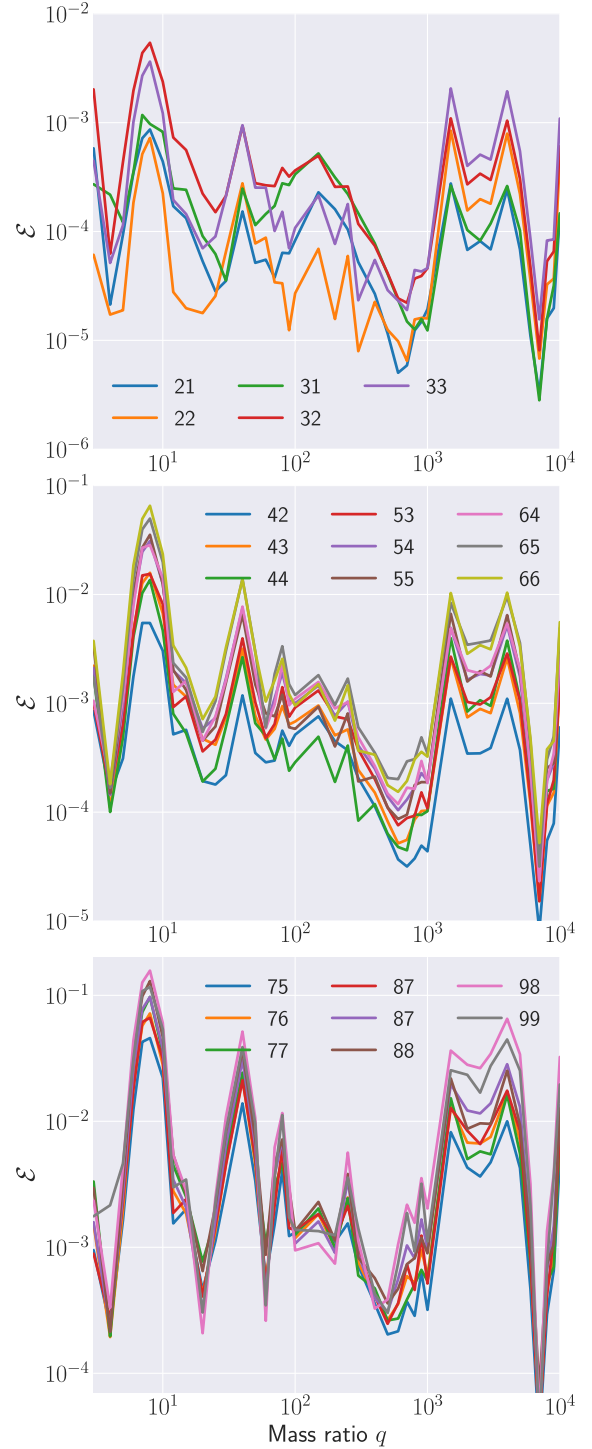

FIG. 3. Validation errors, computed as $\mathcal{E}[h^{\ell m}, h_S^{\ell m}]$, for the individual modes in our surrogate model as a function of mass ratio q.

surrogate predictions are effectively extrapolation and therefore yield uninformative errors. We exclude these points from Fig. 3.

We compare both of these errors to the numerical truncation error of the Teukolsky solver used to produce the ppBHPT training data. We refer to these errors as *ppBHPT numerical errors*.
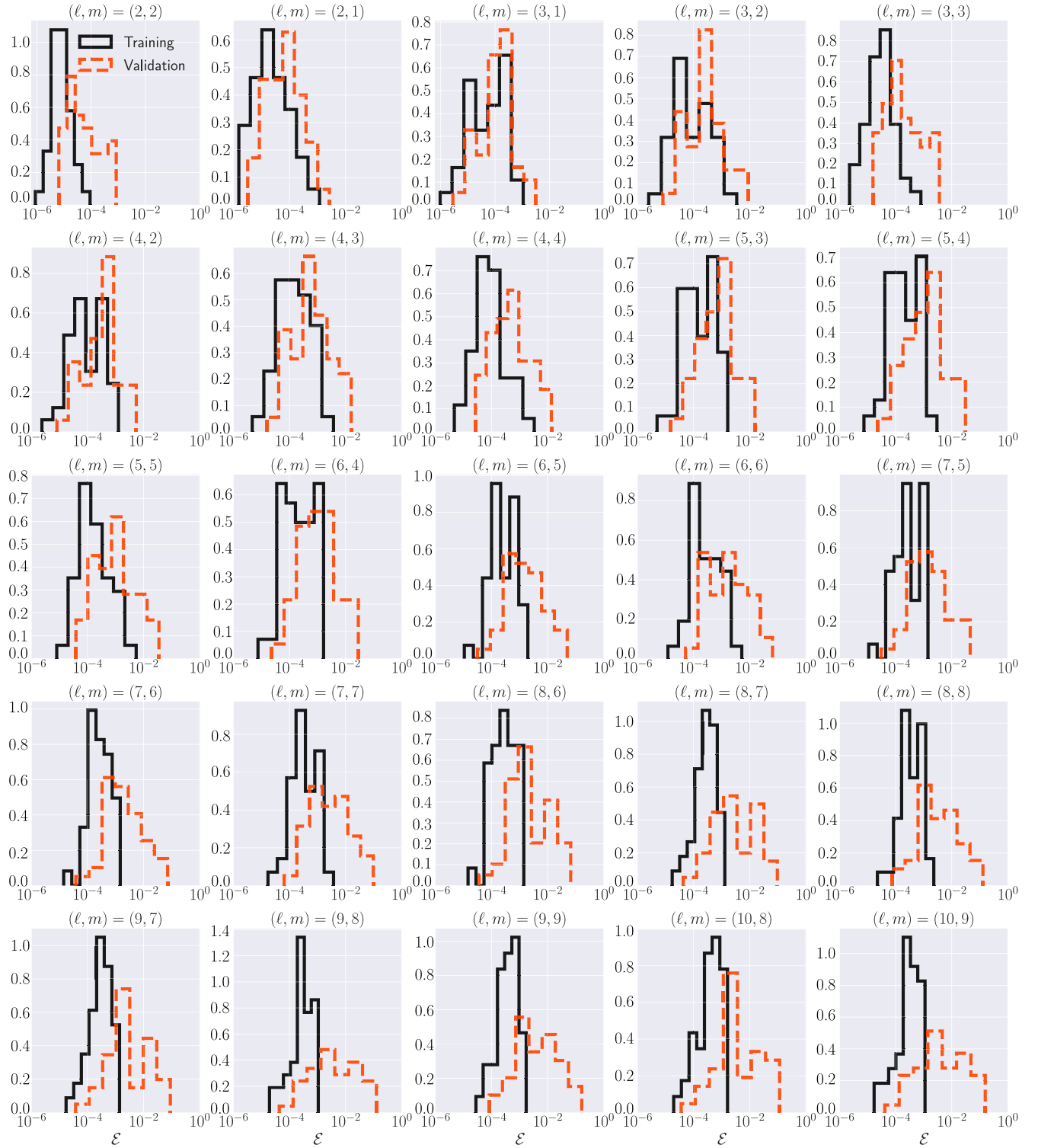
FIG. 4.   Time-domain errors $\mathcal{E}$, defined in Eq. (4), for individual modes considered in `BHPTNRSur1dq1e4`. For comparison, we show both training (black solid lines) and validation (orange-red dashed lines) errors.

In Fig. 3, we report the individual-mode validation errors for `BHPTNRSur1dq1e4` as a function of mass ratio q. We find that the errors are mostly $\leq 10^{-3}$ for modes up to $\ell = 3$, $\leq 10^{-2}$ for modes with $4 \leq \ell \leq 6$ and $\leq 10^{-1}$ for modes with $7 \leq \ell \leq 10$. We further note in Fig. 3 that the

highest errors in each mode corresponds to the same value of q. We also find that the zigzag structure of errors in Fig. 3 is a result of the chosen q values in the parameter space (blue circles in Fig. 2). We model the *data pieces* as a logarithmic function of q. However, the mass ratio values
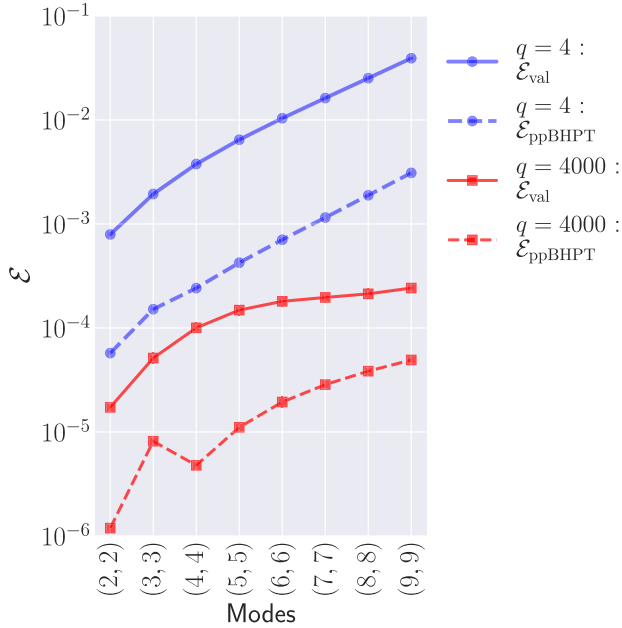
FIG. 5. Validation errors (solid lines) and ppBHPT numerical errors (dashed lines) of select modes for two representative cases: $q = 4$ (blue circles) and $q = 4000$ (red squares).

are not spaced uniformly in logarithmic scale. This results in repetitive patches of dense and sparsely spaced data points. Modeling errors are smaller (larger) around the dense patches (sparsely spaced patches).

Despite the problematic parameter-sampling strategy, the final model should be sufficiently accurate for many data analysis studies in the large-mass ratio regime (cf. Figs. 10 and 13 for more details). In Fig. 4 we provide a mode-by-mode comparison of the training and validation errors. For many of the modes considered in our model, both error measurements are consistent. For some of the modes, and especially for modes with $\ell \geq 8$, the larger validation errors indicate overfitting. Note that the high-error tails seen in Fig. 4 comes from the smallest and highest mass ratio boundaries.

Finally, in Fig. 5, we compare validation error and the ppBHPT Teukolsky solver's numerical error for a select number of modes for two representative cases: $q = 4$ (blue data) and $q = 4000$ (red data). We compute ppBHPT numerical errors by comparing ppBHPT waveforms of different resolution. We find that validation errors are around one order of magnitude larger than ppBHPT numerical errors across all modes.

D. Calibrating ppBHPT to numerical relativity in the comparable mass regime

In the previous subsections, we have built a surrogate model for gravitational waveforms computed within the ppBHPT framework. These waveforms faithfully approximate the physically correct ones only in the limit $q \to \infty$. In order to construct an accurate model at moderate to large

mass ratios, we introduce model calibration parameters and set their values by comparing to NR. These parameters modify each mode's amplitude and phase in a simple way and have the correct behavior in the limit of $q \to \infty$.

Before discussing the particulars of our calibration procedure, its important to consider how one should compare NR and ppBHPT waveforms. For example, both ppBHPT and NR frameworks express dimensioned quantities in terms of a freely specifiable mass scale, which is not the same in the two frameworks. For ppBHPT this scale is selected to be the background black hole spacetime's mass parameter, while in NR it is the sum of the Christodoulou masses of each black hole [87,88]. Our Teukolsky solver sets the background black hole's mass to $m_1 = 1$ (the ppBHPT's mass scale), while the corresponding NR simulation for nonspinning black holes sets the total mass to $m_1 + m_2 = 1$ (the NR simulation's mass scale). So before comparing, we should adjust the ppBHPT's mass-scale to use the NR convention of total mass, which in the ppBHPT's simulation would be $m_1 + m_2 = 1 + 1/q$. This line of reasoning suggests that the ppBHPT modes should be adjusted according to the formula $h^{\ell m}(t) \to \beta h^{\ell m}(t\beta)$ before comparing to NR, where $\beta = 1/(1 + 1/q)$. This straightforward identification works well when comparing post-Newtonian and NR waveforms [87] in the comparable mass ratio regime. In Ref. [63], it was found that (i) the naive value of β accounts for much of the discrepancy between NR and ppBHPT waveforms and (ii) additional model improvements can be obtained by solving an optimization problem for its value.

1. Previous calibration of EMRISur1dq1e4

Motivated by the mass scaling argument given above, Ref. [63] proposed modifying the ppBHPT waveforms according to the formula

$$h_\beta^{\ell,m}(t;q) = \beta h^{\ell,m}(t\beta;q), \qquad (11)$$

where β was set by minimizing the difference

$$\min_\beta \frac{\int |h_\beta^{22}(t;q) - h_{NR}^{22}(t;q)|^2 dt}{\int |h_{NR}^{22}(t;q)|^2 dt}, \qquad (12)$$

between ppBHPT waveforms and nonspinning NR surrogate model [35] trained on SXS simulation data [88–90] for the (2,2) harmonic mode and mass ratios $3 \leq q \leq 10$. The integral appearing in Eq. (12) was evaluated from $-2{,}750\,M$ to $100\,M$ (where M is the total mass of the binary), the duration of the NR surrogate model [35]. The data $\beta(\nu)$ was then fit to a degree 4 polynomial in the symmetric mass $\nu = q/(1 + q)^2$. The resulting function is shown as a dashed cyan line in Fig. 6.

While the calibration choices and techniques of Ref. [63] yielded surprisingly good agreement with NR, a number of deficiencies have been identified. These include (i) the NR surrogate model [35] was built before center-of-mass
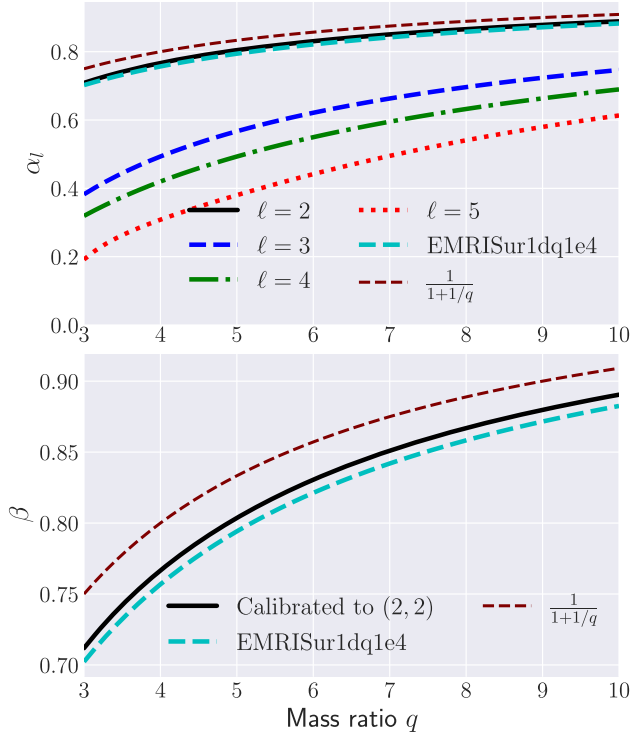
FIG. 6. Scaling parameters $\alpha_\ell(q)$ and $\beta(q)$ as a function of mass ratio q. These parameters are obtained by minimizing the errors between ppBHPT and hybridized NR waveforms. In both panels, the dashed maroon line refers to a naive value of $\alpha = \beta = 1/(1 + 1/q)$ set by including the mass the of smaller black hole as part of the background spacetime. We also show the scaling used in the previous model EMRISur1dq1e4 as a dashed cyan line in both panels.

(CoM) corrected waveform data was available and so the model inherited undesirable features due to CoM drifts, (ii) the NR surrogate model [35] only included about 15 orbits before merger, and it was later found that the calibration parameter deduced on this short interval does not work adequately well on longer time intervals, and (iii) it should be expected that, due to the point-particle approximation, the higher mode amplitudes computed within the ppBHPT framework will be overestimated as compared to NR.

2. Calibration of BHPTNRSur1dq1e4

To overcome the limitations of the previous calibration method discussed in Sec. III D 1, we present an updated set

of choices that provide improvements over the original method. Instead of calibrating BHPTNRSur1dq1e4 data directly to NR data, we use the NRHybSur3dq8 model [38]—a surrogate model for hybridized nonprecessing NR waveforms with the early inspiral waveform obtained using both PN and EOB waveforms. This model was trained on center-of-mass (CoM) corrected waveform data and is much longer in duration, thereby removing two of the three key limitations mentioned in Sec. III D 1.

To calibrate ppBHPT waveforms, we propose modifying the BHPTNRSur1dq1e4 model according to the formula

$$h_{\mathsf{S},\alpha_\ell,\beta}^{\ell,m}(t;q) = \alpha_\ell h_{\mathsf{S}}^{\ell,m}(t\beta;q), \qquad (13)$$

where α_ℓ and β are obtained by minimizing the difference between NRHybSur3dq8 and rescaled ppBHPT waveforms

$$\min_{\alpha_\ell,\beta} \frac{\int |h_{\mathsf{S},\alpha_\ell,\beta}^{\ell,m}(t;q) - h_{\mathrm{NRHyb}}^{\ell,m}(t;q)|^2 dt}{\int |h_{\mathrm{NRHyb}}^{\ell,m}(t;q)|^2 dt}, \qquad (14)$$

between our model BHPTNRSur1dq1e4 and hybridized NR surrogate waveform NRHybSur3dq8 in its nonspinning limit for individual modes over the time window. The integral appearing in Eq. (14) is evaluated from -5000 M to 115 M, which corresponds to the portion of the surrogate model described by NR simulations (i.e., after hybridization). The motivation for the new parameters α_ℓ can be seen as a correction to the point-particle approximation in the comparable mass regime i.e., it accounts for the larger relative size of the smaller black hole. We allow for ℓ-dependent values of α while keeping β fixed for all modes. By numerical computation we have checked that there is essentially no m-dependence $\alpha_{\ell m} \approx \alpha_\ell$ on these amplitude corrections, which can also be motivated by noting that under rotations the harmonic modes mix in m but not ℓ.

To obtain values for β we minimize the cost function (14) using the (2,2) mode, while to find best-fit α^l values we use

TABLE III. Fitting coefficients for β parameters as defined in Eq. (16).

A_β	B_β	C_β	D_β
-1.238 ± 0.003	1.596 ± 0.049	-1.776 ± 0.237	1.0577 ± 0.356

TABLE II. Fitting coefficients for α_ℓ parameters as defined in Eq. (15).

ℓ	A_α^ℓ	B_α^ℓ	C_α^ℓ	D_α^ℓ
2	-1.330 ± 0.007	2.720 ± 0.116	-5.904 ± 0.556	5.548 ± 0.833
3	-3.067 ± 0.017	6.244 ± 0.265	-9.944 ± 1.261	6.437 ± 1.894
4	-3.909 ± 0.032	9.431 ± 0.498	-14.734 ± 2.367	9.744 ± 3.556
5	-4.509 ± 0.102	4.751 ± 1.554	21.959 ± 7.381	-52.350 ± 11.085

$\ell = m$ modes (i.e. (2,2), (3,3), (4,4) and (5,5) modes). To discover each calibration parameters' q-dependence we sample from $q = 3.0$ to $q = 10.0$ with an increment of 0.2, giving a total of 36 data points. These data are then used to fit α_ℓ and β to polynomials in $1/q$:

$$\alpha_\ell(q) = 1 + \frac{A_\alpha^\ell}{q} + \frac{B_\alpha^\ell}{q^2} + \frac{C_\alpha^\ell}{q^3} + \frac{D_\alpha^\ell}{q^4}. \qquad (15)$$

$$\beta(q) = 1 + \frac{A_\beta^\ell}{q} + \frac{B_\beta^\ell}{q^2} + \frac{C_\beta^\ell}{q^3} + \frac{D_\beta^\ell}{q^4}. \qquad (16)$$

The order of the polynomial is chosen in the following way: We first build fits for α^ℓ and β using degrees of polynomial order in q^{-1} from one to four. Next, we select the fit order that minimizes the leave-one-out cross-validation error. A similar strategy has helped us identify that polynomials in q^{-1} lead to more stable fits to the data than polynomials in symmetric mass ratio ν. Values of the coefficients for the final fits are given in Tables II and III for α_ℓ and β, respectively.

Figure 6 shows the scaling parameters α_ℓ and β as a function of mass ratio. We find that the $\ell = 2$ calibration parameters α_2 and β closely match the analogous parameter used to calibrate the EMRISur1dq1e4 model. For higher order modes, α_ℓ becomes smaller, which we interpret as a correction in the point-particle framework to account for the extended size of the smaller black hole. Overall, α_ℓ shows a monotonically increasing behavior with q implying naive $\frac{1}{1+q}$ behavior will be recovered in the larger q limit. While we are using a simple β parameter for all modes, we have experimented with having β^ℓ for different modes. However, this did not appreciably change the final match between ppBHPT and NR hybrid waveforms. Furthermore, individual β^ℓ take almost the same values for different modes bolstering the claim that β, used to scale the time-axis, is related to the mass scaling and provides support for using one single β for all modes.

IV. COMPARISON BETWEEN THE CALIBRATED PPBHPT MODEL AND NR

A. Time domain error in the comparable mass regime

Figure 7 shows the time-domain error between the calibrated BHPTNRSur1dq1e4 and NRHybSur3dq8 models. We show the error using all available NRHybSur3dq8 modes[3] as well as individual mode errors. Our comparison includes both the entire waveform over all inspiral-merger-ringdown (IMR) regimes (solid

---

[3]The NRHybSur3dq8 model includes the following harmonic modes: $\{\ell, m\} = \{(2, 1), (2, 2), (3, 1), (3, 2), (3, 3), (4, 2), (4, 3), (4, 4), (5, 5)\}$.
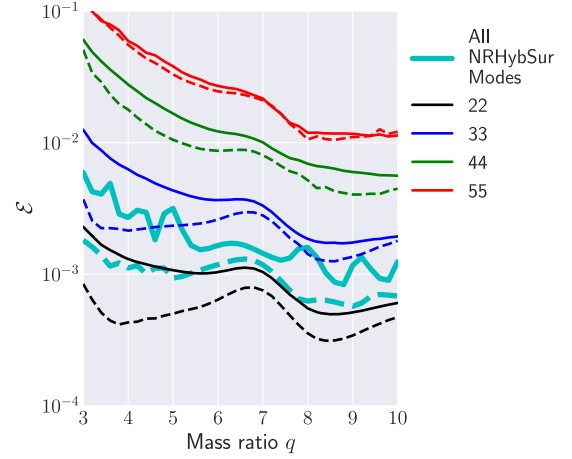

FIG. 7.   Time-domain error between the calibrated ppBHPT waveforms and the NR hybrid surrogate model as a function of mass ratio q. We show the errors computed over the full inspiral-merger-ringdown regimes (solid lines) and inspiral-only errors (dashed lines) restricting the waveform to $t \leq -50M$. We show errors for select modes [(2,2), (3,3), (4,4), and (5,5)] as well as for the case including all available NRHybSur3dq8 modes including $\ell \neq m$ (referred to as "All NRHybSur Modes").

lines) and inspiral-only errors (dashed lines), restricting the waveform to $t \leq -50M$.

In all cases, the differences before calibration are of order unity, while the agreement between the calibrated BHPTNRSur1dq1e4 and NR waveforms improves to $\mathcal{E} \sim 10^{-3}$ for $(\ell, m) = (2, 2)$ and $\mathcal{E} \sim 10^{-2}$ for $(\ell, m) = \{(3, 3), (4, 4), (5, 5)\}$. Many of the $\ell \neq m$ modes have errors as high as $\mathcal{E} \sim 10^{-1}$ even after calibration (not shown).

We expect that in the merger and ringdown regimes, the calibrated ppBHPT waveforms that match so well in the inspiral will no longer serve as a faithful physical ansatz. Instead, we expect the ringdown signal to be described by perturbations of the remnant black hole whose mass and spin only agree with the initial background solution in the limit $q \to \infty$. This expectation is confirmed in Fig. 7 as we see nearly an order-of-magnitude increase in the IMR error as compared to the inspiral-only error for mass ratios less than $q \approx 5$. By $q \approx 10$, however, the NR-calibrated BHPTNRSur1dq1e4 does a reasonably good job even in the ringdown regime, which is also apparent in the bottom panel of Fig. 9. This suggests that high-accuracy models based on calibrated ppBHPT waveforms may require special treatment in the merger-ringdown regime—which is commonly employed in other waveform modeling efforts —although at mass ratios beyond $q \approx 10$ the current approach already does a reasonably good job.

These results are shown in more detail for $q = 4$, where Fig. 8 shows four of the most important harmonic modes before and after calibration. Before calibration the BHPTNRSur1dq1e4 and NRHybSur3dq8 waveforms differ visibly in both amplitude and phase evolution. The calibrated ppBHPT and NRHybSur3dq8 waveforms,
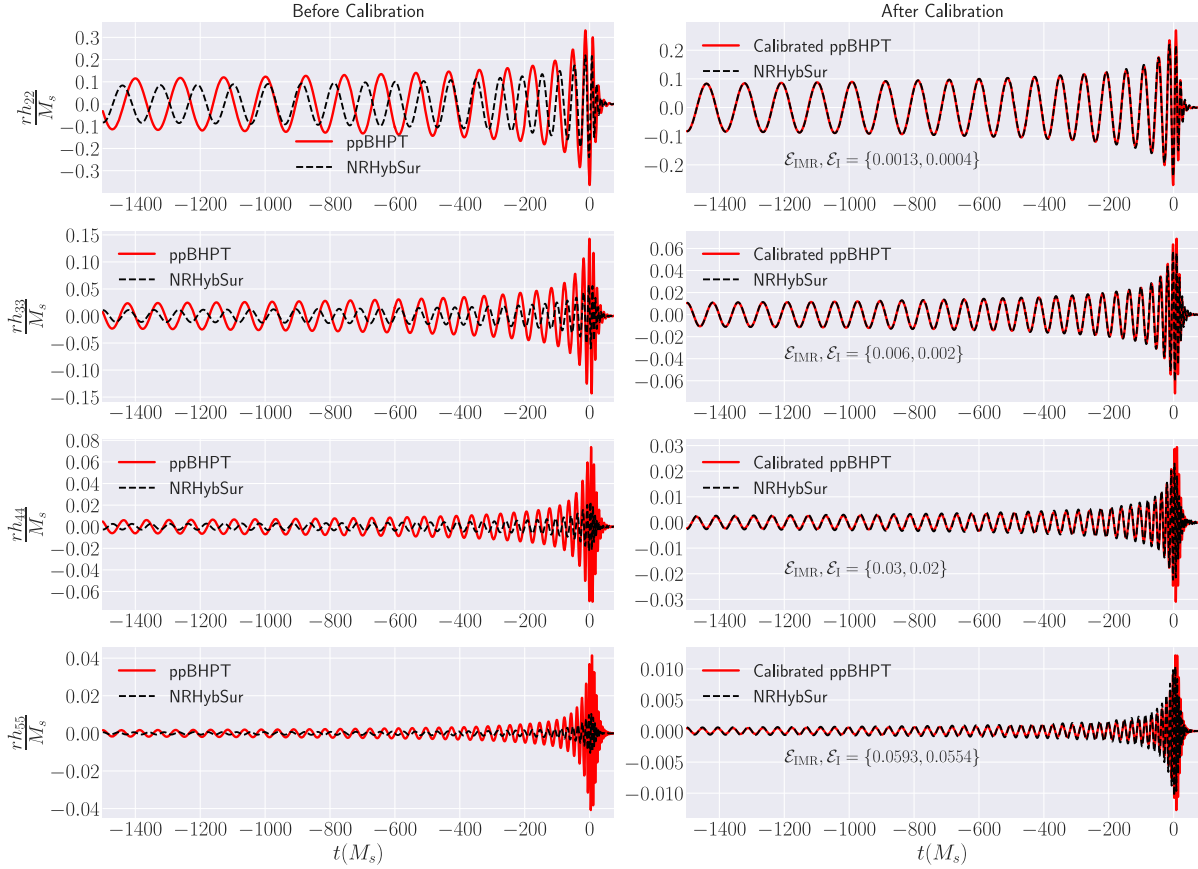
FIG. 8. Waveform difference between `NRHybSur3dq8` (*referred as "NRHybSur"*) and ppBHPT waveforms before and after calibration. We show the five different representative modes for $q = 4$ to demonstrate the efficacy of the calibration used. Errors for the full inspiral-merger-ringdown waveform and only inspiral part are denoted by $\mathcal{E}_{\mathrm{IMR}}$ and $\mathcal{E}_{\mathrm{I}}$ respectively. The mass scale is denoted by M_s, which is either m_1 (ppBHPT) or $m_1 + m_2$ (NR) on the left column or either $m_1\beta$ (calibrated ppBHPT) or $m_1 + m_2$ (NR) on the right column.

however, show surprisingly good agreement although some differences remain in the merger-ringdown part.

We note that these calibration parameters have been obtained by comparing the raw ppBHPT waveform to NR over a time window $[-5000, 115]M$ that characterizes the late inspiral through ringdown. To test whether the scaling works at earlier times too, we compare `NRHybSur3dq8` to the calibrated ppBHPT waveforms over the longest possible duration, which is $30,500m_1$ using the ppBHPT's mass scale. We show an example case in Fig. 9. We plot the dominant (2,2) mode for both `NRHybSur3dq8` (dashed black line) and rescaled ppBHPT (solid red line) waveforms at $q = 8$ in early inspiral as well as in the late inspiral and merger-ringdown parts. The waveforms are nearly indistinguishable for the entire duration, and we compute the error to be $\mathcal{E}_{\mathrm{IMR}} = 0.00086$.

B. Frequency domain mismatch between rescaled-ppBHPT surrogate and NR

In Fig. 10, we show the frequency-domain mismatch between the calibrated ppBHPT surrogate waveforms and `NRHybSur3dq8` waveforms as a function of total mass for

different mass ratios. Frequency domain mismatch between two waveforms $\mathfrak{h}_1$ and $\mathfrak{h}_2$ is defined as:

$$\langle \mathfrak{h}_1, \mathfrak{h}_2 \rangle = 4\mathrm{Re} \int_{f_{\min}}^{f_{\max}} \frac{\tilde{\mathfrak{h}}_1(f)\tilde{\mathfrak{h}}_2^*(f)}{S_n(f)}\,df, \qquad (17)$$

where $\tilde{\mathfrak{h}}(f)$ indicates the Fourier transform of the complex strain $\mathfrak{h}(t)$, $*$ indicates complex conjugation, Re indicates the real part, and $S_n(f)$ is the one-sided power spectral density of the Advanced LIGO detector at its design sensitivity. We set $f_{\min}$ to be 20 Hz while $f_{\max}$ is set to be 990 Hz. We note that `BHPTNRSur1dq1e4` waveforms are long enough for all of the mass ratio and total mass configurations considered in Fig. 10, even the initial instantaneous frequency of the (5,5) mode is below 20 Hz. Before transforming the time domain waveform to the frequency domain, we first taper the time domain waveform using a Planck window [91], and then zero-pad to the nearest power of two. The tapering at the start of the waveform is done over 1.5 cycles of the (2,2) mode. The tapering at the end is done over the last $20M$. The mismatches are always optimized over shifts in time,
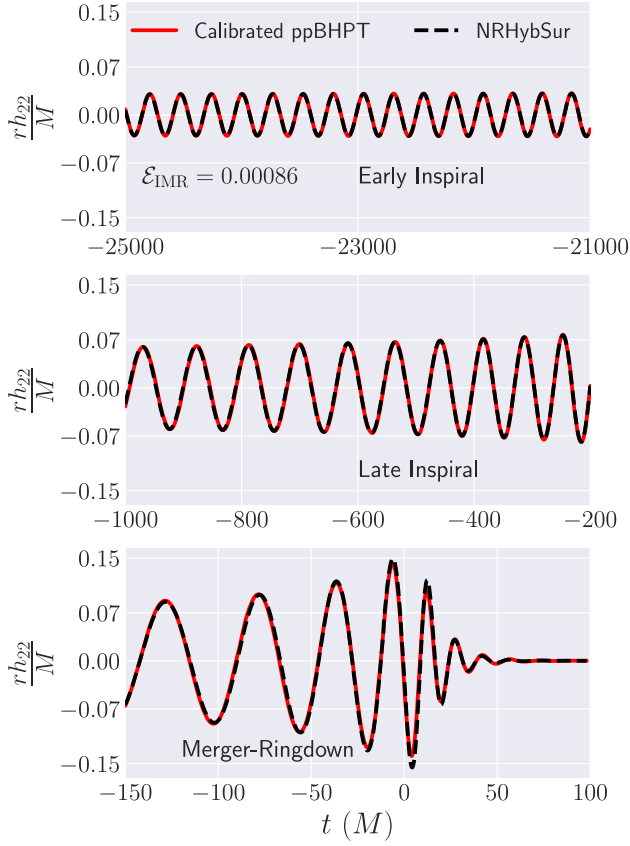
FIG. 9. Effectiveness of the calibration obtained from a restricted time window $[-5000, 115]M$ over the entire length of the waveform. Shown are the early inspiral (upper panel), late inspiral (middle panel), and merger-ringdown (lower panel) parts of the (2,2) mode for $q = 8$. The hybridized NR surrogate model NRHybSur3dq8 (*referred as "NRHybSur"*) and calibrated ppBHPT waveforms are shown in black dashed and solid red, respectively. $\mathcal{E}_{\mathrm{IMR}}$ is the L_2 norm error computed over the entire waveform.

polarization angle, and initial orbital phase. We find that as the mass ratio increases, agreement between the calibrated ppBHPT surrogate waveform and NRHybSur3dq8 model improves. For $q \geq 5$, mismatches are below 0.01 (Fig. 10) indicating at least an order of magnitude improvement over our precursor EMRISur1dq1e4 model.

C. Extrapolating the model to $q \to 1$

Even though BHPTNRSur1dq1e4 is trained for $q \geq 2.5$, we find that the model can be extrapolated to mass ratio $q = 1.2$, although we advise caution with any extrapolation. This is particularly exciting as our time-domain Teukolsky solver struggles to generate waveforms below $q \approx 2.5$. The model can therefore be used to simulate ppBHPT waveforms for binaries with mass ratios $1.2 \leq q \leq 2.5$. Although, as can be anticipated from Fig. 7, the calibrated BHPTNRSur1dq1e4 is not a faithful approximation to GR as $q \to 1$. Figure 11 shows
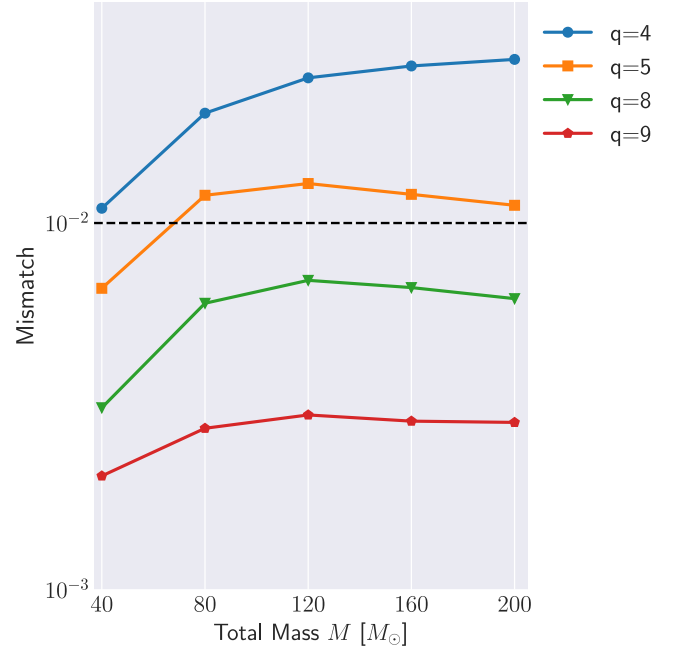

FIG. 10. Frequency-domain mismatches between the rescaled-ppBHPT surrogate waveforms and NRHybSur3dq8 model for different mass ratios. The mismatches are shown as a function of the binary total mass M at inclination $\iota = 0.0$ and orbital phase $\varphi = 0$, and are computed using the advanced LIGO design sensitivity noise curve. We set the minimum (maximum) frequency appearing in Eq. (17) to be 20 Hz (990 Hz). The dashed horizontal line demarcates a mismatch of 0.01, a commonly used threshold for sufficiently good model quality.

the dominant (2,2) mode of these two models evaluated at $q = 1.2$. At this mass ratio, we find the (2,2) mode error of our model to be ~ 0.026 for the full inspiral-merger-ringdown waveform and ~ 0.012 for inspiral-only part.
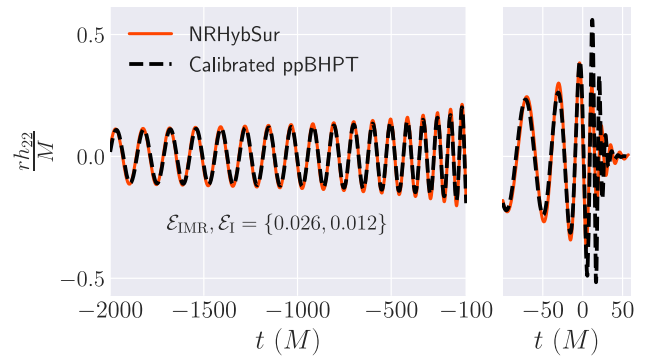

FIG. 11. Comparison between the calibrated ppBHPT waveform and the hybridized NR surrogate model NRHybSur3dq8 (*referred as "NRHybSur"*) at $q = 1.2$. We show the dominant $(\ell, m) = (2, 2)$ mode. Errors for the full inspiral-merger-ringdown waveform and inspiral-only part are denoted by $\mathcal{E}_{\mathrm{IMR}}$ and $\mathcal{E}_{\mathrm{I}}$, respectively.

D. Numerical relativity in the intermediate mass ratio regime

Due to a scarcity of NR data for intermediate-mass ratio ranges, say from $q = 10$ to $q = 10^4$, it is difficult to perform a detailed comparison between ppBHPT and NR data. Some recent breakthroughs, however, have made it possible to perform NR simulations for high mass ratio binaries. As we are interested in nonspinning systems, we will consider both SXS NR simulation data for $q = \{15, 30\}$ [53,54] and RIT NR simulations for $q = \{15, 32\}$ [51].

1. Comparison against SXS NR data at q = 15 and q = 30

We begin with a comparison between the calibrated ppBHPT waveforms, coming from `BHPTNRSur1dq1e4` model, and SXS NR data at mass ratio $q = 15$. In Fig. 12, we plot three representative modes. NR data is shown in red solid lines whereas NR-calibrated ppBHPT waveforms are plotted as black dashed lines. Note that we do not perform any on-the-fly re-scaling for the ppBHPT waveform but instead use results from Sec. III D 2. For comparison, we also include the `SEOBNRv4HM` model. We find both `SEOBNRv4HM` and `BHPTNRSur1dq1e4` match NR very well at $q = 15$. For `BHPTNRSur1dq1e4`, full [inspiral-only] waveform errors for (2,2), (3,3), and (4,4) modes are 0.00076 [0.00068], 0.0021 [0.0020], and 0.0054 [0.0047], respectively. The `SEOBNRv4HM` shows a noticeable offset from NR around merger and ringdown for the higher

modes, although both models deliver good accuracy overall, especially premerger.

When interpreting these errors it is important to note that, from Fig. 5, it is clear that our model cannot achieve an error $\mathcal{E} \lesssim 0.001$ due to numerical error in the un-calibrated `BHPTNRSur1dq1e4`. As this value is consistent with what we see in our comparisons with SXS data, it is not clear if even better agreement with NR could be achieved with higher-accuracy ppBHPT waveform training data. As an additional check, we also perform an $\{\alpha_\ell, \beta\}$ optimization between the raw ppBHPT and SXS NR data at $q = 15$. We find that the α_ℓ and β values obtained this way match closely to values obtained from Eqs. (15) and (16), and negligible improvement in errors observed. This implies that the calibration carried out in the range $3.0 \le q \le 10.0$ continues to work well at mass ratio 15.

Next, we compare calibrated-ppBHPT waveforms against the highest mass ratio SXS NR data at $q = 30$. Figure 13 shows three representative modes for $q = 30$ for both NR (red solid lines) and calibrated-ppBHPT waveforms (black dashed lines). We note that at $q = 30$, $\beta \approx 0.961$ while $1/(1 + 1/q) \approx 0.967$, suggesting even an uncalibrated ppBHPT waveform would work reasonably well. We find for the 22 mode, the errors between calibrated-ppBHPT and NR data are around $\sim 10^{-3}$ implying a good match. For higher order modes, errors are still around ~ 0.01. Furthermore, when we compare `SEOBNRv4HM` to NR data, we find the full waveform errors to be similar to the `BHPTNRSur1dq1e4` for higher order modes whereas
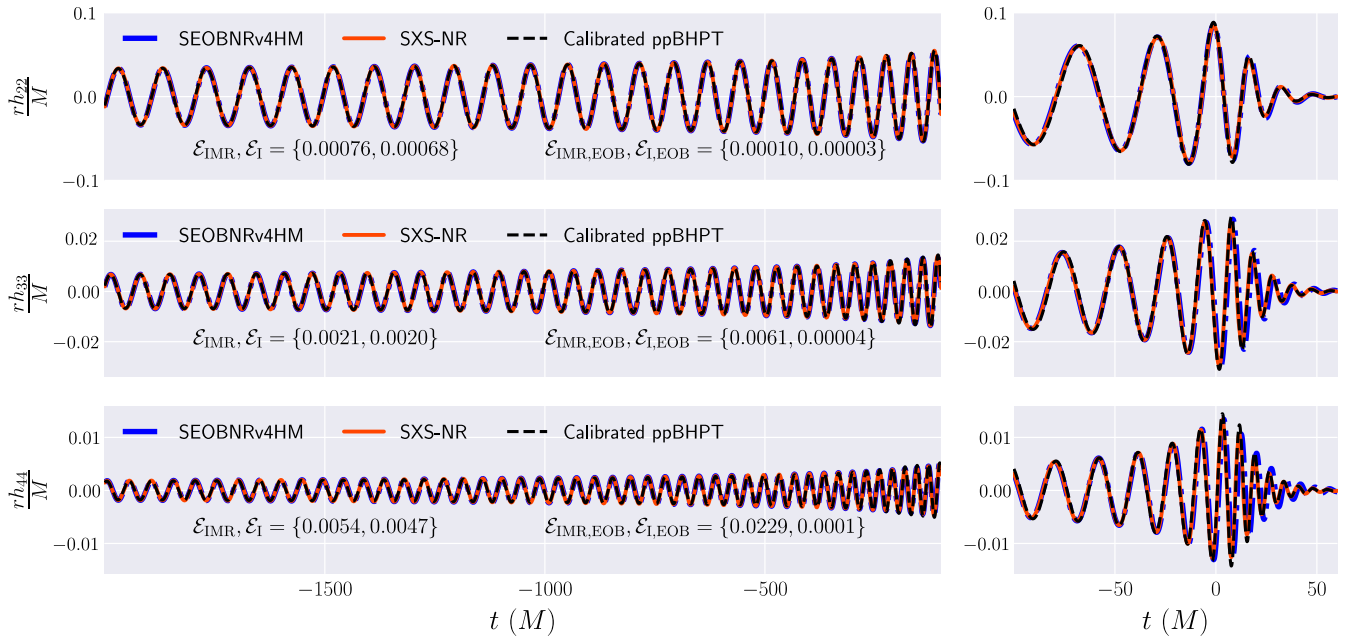

FIG. 12. Waveform difference between the calibrated ppBHPT waveform (dashed black) and NR data (solid red) from the SXS collaboration (simulation ID SXS:BBH:2304) for $q = 15$. We show three different representative modes to demonstrate the efficacy of our model. Errors for the full inspiral-merger-ringdown waveform and only inspiral part is denoted by $\mathcal{E}_{\mathrm{IMR}}$ and $\mathcal{E}_{\mathrm{I}}$ respectively. For comparison, we also show `SEOBNRv4HM` waveform modes (solid blue).
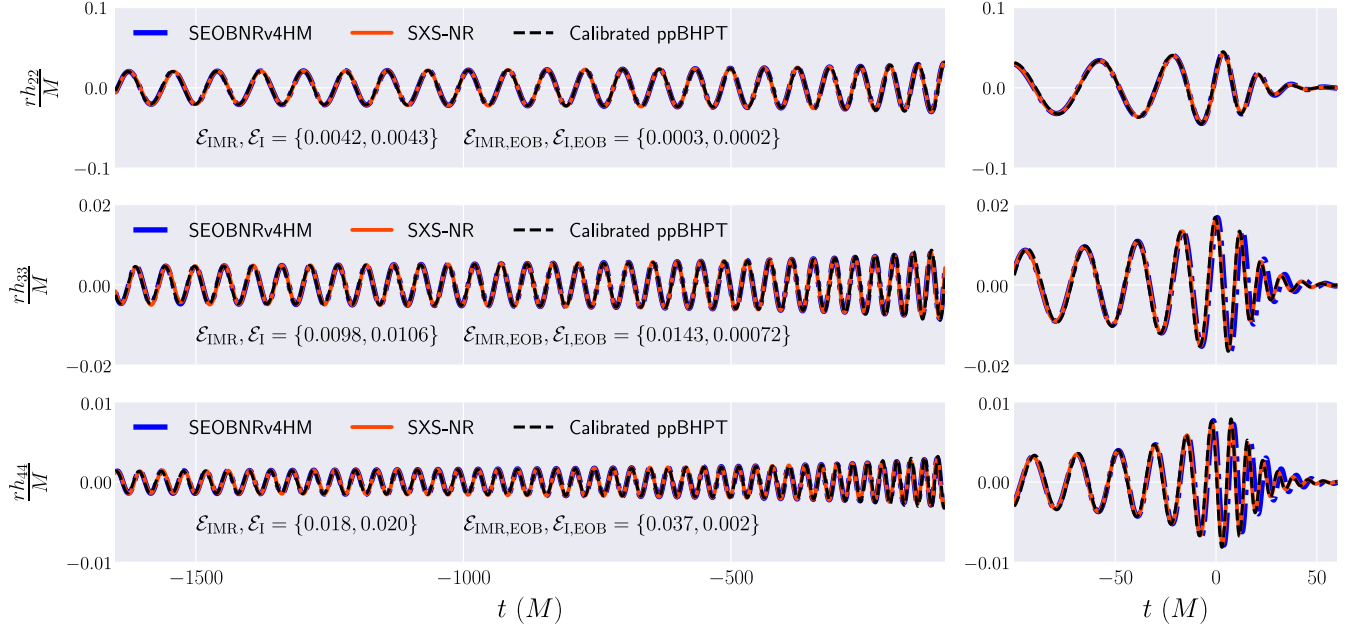
FIG. 13. Waveform difference between the calibrated ppBHPT waveform (dashed black) and NR data (solid red) from the SXS collaboration for $q = 30$ [54]. We show three different representative modes to demonstrate the efficacy of our model. Errors for the full inspiral-merger-ringdown waveform and only inspiral part is denoted by $\mathcal{E}_{\mathrm{IMR}}$ and $\mathcal{E}_{\mathrm{I}}$ respectively. For comparison, we also show `SEOBNRv4HM` waveform modes (solid blue).

around one order of magnitude better than the `BHPTNRSur1dq1e4` model for (2,2) mode.

Just as in the $q = 15$ analysis, as an additional check of our calibration parameters we perform a fresh $\{\alpha_\ell, \beta\}$

optimization between the raw ppBHPT and SXS NR data at $q = 30$. We find that while the α_ℓ and β values computed this way are quite close to values obtained from Eqs. (15) and (16), they provide a better match with NR data—with
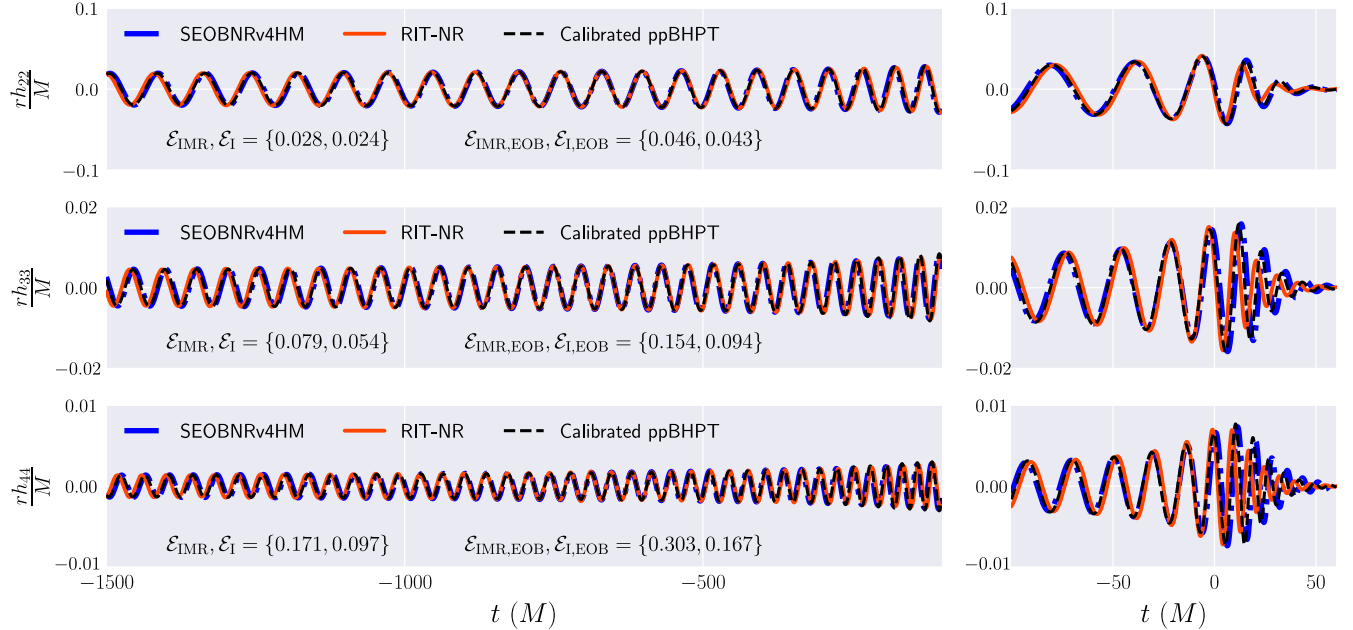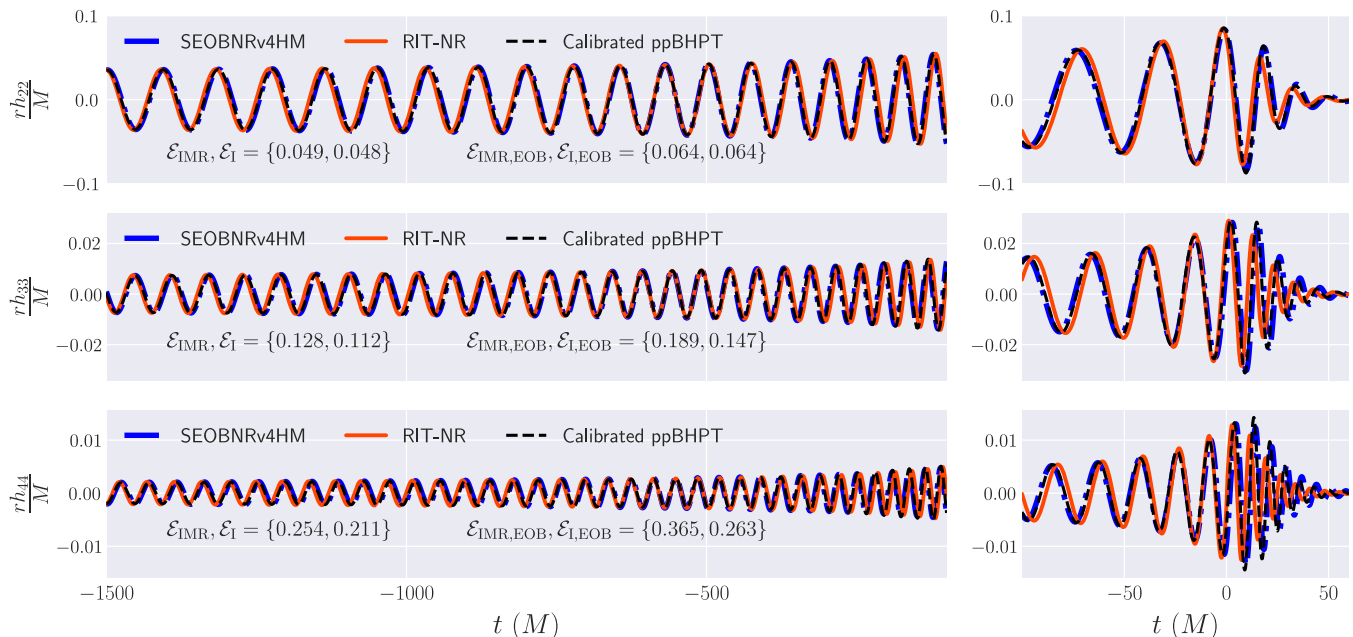

FIG. 14. Waveform difference between the calibrated ppBHPT waveform (black dashed) and NR data (solid red) from the RIT group (simulation ID RIT:BBH:0792) for $q = 32$. We show three different representative modes to demonstrate the efficacy of our model. Errors for the full inspiral-merger-ringdown waveform and only inspiral part is denoted by $\mathcal{E}_{\mathrm{IMR}}$ and $\mathcal{E}_{\mathrm{I}}$ respectively. For comparison, we also show `SEOBNRv4HM` waveform modes (solid blue).

FIG. 15. Waveform difference between the calibrated ppBHPT waveform (black dashed) and NR data (solid red) from the RIT group (simulation ID RIT:BBH:0373) for $q = 15$. We show three different representative modes to demonstrate the efficacy of our model. Errors for the full inspiral-merger-ringdown waveform and only inspiral part is denoted by $\mathcal{E}_{\mathrm{IMR}}$ and $\mathcal{E}_{\mathrm{I}}$, respectively. For comparison, we also show SEOBNRv4HM waveform modes (solid blue).

improving the (2,2) mode error by at least one order of magnitude making it comparable to the (2,2) mode error of SEOBNRv4HM. This suggests that relatively larger error for calibrated-ppBHPT at $q = 30$ when compared to NR data than at $q = 15$ is due to the error in extrapolating α and β scaling much beyond the mass ratio range ($3 \leq q \leq 10$) it is trained on. We expect this scaling to become more robust in the future as more NR simulations become available beyond $q = 10$. We further note that the NR data at $q = 30$ is shorter in length (only $\sim 1800M$ long). Careful comparison between ppBHPT and NR data in this regime needs longer NR data. We leave that study for the future when smoother and longer NR data will be available. Nonetheless, that a reasonable match between these two types of waveforms is obtained in this regime is a promising sign.

2. Comparison against RIT NR data at $q = 15$ and $q = 32$

We now compare calibrated BHPTNRSur1dq1e4 against the publicly available RIT NR data at mass ratio $q = 32$. Figure 14 shows three representative modes for both NR (red solid lines), calibrated BHPTNRSur1dq1e4 (black dashed lines), and SEOBNRv4HM (solid blue) at $q = 32$. We find that, for (2,2) mode, scaled ppBHPT and SEOBNRv4HM yields an error of ~ 0.01, and for subdominant modes the calibrated BHPTNRSur1dq1e4 matches the NR data a bit more closely than the SEOBNRv4HM model around merger and ringdown. It is interesting to note that at $q = 30$, both the calibrated

BHPTNRSur1dq1e4 and SEOBNRv4HM models provide at least one order of magnitude better match when tested against the SXS NR data. This may imply some systematic difference between SXS and RIT NR simulations at high mass ratios. RIT simulations also show evidence of residual eccentricity that introduces additional small modulations in the waveform—potentially increasing the disagreements. We have also compared the calibrated BHPTNRSur1dq1e4 ppBHPT and SEOBNRv4HM models with the RIT NR data at $q = 15$ (Fig. 15), but we encountered issues similar to the $q = 32$ case.

V. DISCUSSION AND CONCLUSION

In this paper, we have described more fully the methods used to build EMRISur1dq1e4 (which was introduced in a short letter [63]), a time-domain surrogate model of waveforms obtained through numerically solving the Teukolsky equation sourced by a test-particle with adiabatic-driven inspiral. We apply point-particle black hole perturbation theory (ppBHPT) framework to nonspinning systems with mass ratios from $q = 2.5$ to $q = 10,000$. While intermediate mass ratio systems ($q > 10$) are targets for our model, we use ppBHPT waveforms in the regime $q < 10$ to (i) carry out comparisons between numerical relativity and ppBHPT and (ii) calibrate the surrogate model to NR thereby vastly improving the model's accuracy.

We have also taken this opportunity to make numerous important improvements to the underlying model. The updated model, BHPTNRSur1dq1e4, is $30,500m_1$ in

duration, making it suitable to be used in building template banks for LIGO/Virgo data analysis at larger mass-ratios, and also serve as a useful tool for mock data analyses for future observatories. We also employ an improved transition trajectory algorithm between early inspiral and plunge [92]—thereby reducing nonphysical jumps/oscillations in the waveforms (cf. Sec. II C). The `BHPTNRSur1dq1e4` model includes a total of 50 modes up to $\ell = 10$, which is particularly important as subdominant modes are expected to play an important role in intermediate mass ratio systems [34,93–95]. Please see Table I for a complete summary of the `BHPTNRSur1dq1e4`. This model is publicly available as part of both the Black Hole Perturbation Toolkit [96] and GWSurrogate [97].

We also perform a detailed comparison between ppBHPT and NR waveforms in the comparable mass ratio regime for all modes up to $\ell \leq 5$. We find that after a simple calibration step the ppBHPT waveforms yield remarkable agreement with NR. The calibrated waveforms are also compared against available SXS NR data at $q = \{15, 30\}$ and against RIT NR data at $q = \{15, 32\}$, and are found to give good agreement for many of the subdominant modes even up through merger and ringdown. Furthermore, by construction, the calibration parameters "turn off" as $q \rightarrow \infty$, so that the correct test-particle behavior is recovered. Our results suggest that suitably calibrated ppBHPT

models may be used to generate accurate late inspiral, merger, and ringdown waveforms in the $q > 10$ regime that is especially challenging for NR. Future models should include obvious extensions such as spin, effects of eccentricity, and spin-orbit precession.

[1] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs, Phys. Rev. X **9**, 031040 (2019).

[2] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run, Phys. Rev. X **11**, 021053 (2021).

[3] Hua Feng and Roberto Soria, Ultraluminous X-ray Sources in the Chandra and XMM-Newton Era, New Astron. Rev. **55**, 166 (2011).

[4] Dheeraj R. Pasham, Tod E. Strohmayer, and Richard F. Mushotzky, A 400 solar mass black hole in the Ultra-luminous X-ray source M82 X-1 accreting close to its Eddington limit, Nature (London) **513**, 74 (2014).

[5] Mar Mezcua, Observational evidence for intermediate-mass black holes, Int. J. Mod. Phys. D **26**, 1730021 (2017).

[6] Nathan W. C. Leigh, Nora Lützgendorf, Aaron M. Geller, Thomas J. Maccarone, Craig Heinke, and Alberto Sesana, On the coexistence of stellar-mass and intermediate-mass black holes in globular clusters, Mon. Not. R. Astron. Soc. **444**, 29 (2014).

[7] Morgan MacLeod, Michele Trenti, and Enrico Ramirez-Ruiz, The close stellar companions to intermediate mass black holes, Astrophys. J. **819**, 70 (2016).

[8] Pau Amaro-Seoane *et al.*, Laser interferometer space antenna, arXiv:1702.00786.

[9] Pau Amaro-Seoane, Jonathan R. Gair, Marc Freitag, M. Coleman Miller, Ilya Mandel, Curt J. Cutler, and Stanislav Babak, Astrophysics, detection and science applications of intermediate- and extreme mass-ratio inspirals, Classical Quantum Gravity **24**, R113 (2007).

[10] Christopher P. L. Berry, Scott A. Hughes, Carlos F. Sopuerta, Alvin J. K. Chua, Anna Heffernan, Kelly Holley-Bockelmann, Deyan P. Mihaylov, M. Coleman Miller, and Alberto Sesana, The unique potential of extreme mass-ratio inspirals for gravitational-wave astronomy, arXiv:1903.03686.

[11] Pau Amaro-Seoane, Detecting intermediate-mass ratio inspirals from the ground and space, Phys. Rev. D **98**, 063018 (2018).

[12] Seiji Kawamura *et al.*, Current status of space gravitational wave antenna DECIGO and B-DECIGO, Prog. Theor. Exp. Phys. **2021**, 05A105 (2021).

[13] Manuel Arca Sedda *et al.*, The missing link in gravitational-wave astronomy: Discoveries waiting in the

decihertz range, Classical Quantum Gravity **37,** 215011 (2020).

[14] Jonathan R. Gair, Ilya Mandel, M. Coleman Miller, and Marta Volonteri, Exploring intermediate and massive black-hole binaries with the Einstein telescope, Gen. Relativ. Gravit. **43,** 485 (2011).

[15] E. A. Huerta and Jonathan R. Gair, Intermediate-mass-ratio-inspirals in the Einstein telescope: I. Signal-to-noise ratio calculations, Phys. Rev. D **83,** 044020 (2011).

[16] E. A. Huerta and Jonathan R. Gair, Intermediate-mass-ratio-inspirals in the Einstein telescope. II. Parameter estimation errors, Phys. Rev. D **83,** 044021 (2011).

[17] Carl-Johan Haster, Zhilu Wang, Christopher P. L. Berry, Simon Stevenson, John Veitch, and Ilya Mandel, Inference on gravitational waves from coalescences of stellar-mass compact objects and intermediate-mass black holes, Mon. Not. R. Astron. Soc. **457,** 4499 (2016).

[18] Jillian Bellovary, Alyson Brooks, Monica Colpi, Michael Eracleous, Kelly Holley-Bockelmann, Ann Hornschemeier, Lucio Mayer, Priya Natarajan, Jacob Slutsky, and Michael Tremmel, Where are the intermediate mass black holes?, arXiv:1903.08144.

[19] Enrico Barausse, Luciano Rezzolla, David Petroff, and Marcus Ansorg, Gravitational waves from Extreme Mass Ratio Inspirals in non-pure Kerr spacetimes, Phys. Rev. D **75,** 064026 (2007).

[20] Enrico Barausse and Luciano Rezzolla, The Influence of the hydrodynamic drag from an accretion torus on extreme mass-ratio inspirals, Phys. Rev. D **77,** 104027 (2008).

[21] Jonathan R. Gair, Eanna E. Flanagan, Steve Drasco, Tanja Hinderer, and Stanislav Babak, Forced motion near black holes, Phys. Rev. D **83,** 044037 (2011).

[22] Nicolas Yunes, Bence Kocsis, Abraham Loeb, and Zoltan Haiman, Imprint of Accretion Disk-Induced Migration on Gravitational Waves from Extreme Mass Ratio Inspirals, Phys. Rev. Lett. **107,** 171103 (2011).

[23] Enrico Barausse, Vitor Cardoso, and Paolo Pani, Can environmental effects spoil precision gravitational-wave astrophysics?, Phys. Rev. D **89,** 104059 (2014).

[24] Enrico Barausse, Vitor Cardoso, and Paolo Pani, Environmental effects for gravitational-wave astrophysics, J. Phys. Conf. Ser. **610,** 012044 (2015).

[25] A. Derdzinski, D. D'Orazio, P. Duffell, Z. Haiman, and A. Macfadyen, Evolution of gas disc-embedded intermediate mass ratio inspirals in the LISA band, Mon. Not. R. Astron. Soc. **501,** 3540 (2021).

[26] Jonathan R. Gair, Michele Vallisneri, Shane L. Larson, and John G. Baker, Testing general relativity with low-frequency, space-based gravitational-wave detectors, Living Rev. Relativity **16,** 7 (2013).

[27] Gabriel Andres Piovano, Andrea Maselli, and Paolo Pani, Model independent tests of the Kerr bound with extreme mass ratio inspirals, Phys. Lett. B **811,** 135860 (2020).

[28] Nicolas Yunes and C. F. Sopuerta, Testing effective quantum gravity with gravitational waves from extreme mass ratio inspirals, J. Phys. Conf. Ser. **228,** 012051 (2010).

[29] P. Canizares, J. R. Gair, and C. F. Sopuerta, Testing Chern-Simons modified gravity with observations of extreme-mass-ratio binaries, J. Phys. Conf. Ser. **363,** 012019 (2012).

[30] Priscilla Canizares, Jonathan R. Gair, and Carlos F. Sopuerta, Testing Chern-Simons modified gravity with gravitational-wave detections of extreme-mass-ratio binaries, Phys. Rev. D **86,** 044010 (2012).

[31] Carl L. Rodriguez, Ilya Mandel, and Jonathan R. Gair, Verifying the no-hair property of massive compact objects with intermediate-mass-ratio inspirals in advanced gravitational-wave detectors, Phys. Rev. D **85,** 062002 (2012).

[32] Alvin J. K. Chua, Sonke Hee, Will J. Handley, Edward Higson, Christopher J. Moore, Jonathan R. Gair, Michael P. Hobson, and Anthony N. Lasenby, Towards a framework for testing general relativity with extreme-mass-ratio-inspiral observations, Mon. Not. R. Astron. Soc. **478,** 28 (2018).

[33] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Tests of general relativity with binary black holes from the second LIGO-Virgo gravitational-wave transient catalog, Phys. Rev. D **103,** 122002 (2021).

[34] Tousif Islam, Scott E. Field, Carl-Johan Haster, and Rory Smith, High precision source characterization of intermediate mass-ratio black hole coalescences with gravitational waves: The importance of higher order multipoles, Phys. Rev. D **104,** 084068 (2021).

[35] Jonathan Blackman, Scott E. Field, Chad R. Galley, Béla Szilágyi, Mark A. Scheel, Manuel Tiglio, and Daniel A. Hemberger, Fast and Accurate Prediction of Numerical Relativity Waveforms from Binary Black Hole Coalescences Using Surrogate Models, Phys. Rev. Lett. **115,** 121102 (2015).

[36] Jonathan Blackman, Scott E. Field, Mark A. Scheel, Chad R. Galley, Christian D. Ott, Michael Boyle, Lawrence E. Kidder, Harald P. Pfeiffer, and Béla Szilágyi, Numerical relativity waveform surrogate model for generically precessing binary black hole mergers, Phys. Rev. D **96,** 024058 (2017).

[37] Jonathan Blackman, Scott E. Field, Mark A. Scheel, Chad R. Galley, Daniel A. Hemberger, Patricia Schmidt, and Rory Smith, A surrogate model of gravitational waveforms from numerical relativity simulations of precessing binary black hole mergers, Phys. Rev. D **95,** 104023 (2017).

[38] Vijay Varma, Scott E. Field, Mark A. Scheel, Jonathan Blackman, Lawrence E. Kidder, and Harald P. Pfeiffer, Surrogate model of hybridized numerical relativity binary black hole waveforms, Phys. Rev. D **99,** 064045 (2019).

[39] Vijay Varma, Scott E. Field, Mark A. Scheel, Jonathan Blackman, Davide Gerosa, Leo C. Stein, Lawrence E. Kidder, and Harald P. Pfeiffer, Surrogate models for precessing binary black hole simulations with unequal masses, Phys. Rev. Res. **1,** 033015 (2019).

[40] Tousif Islam, Vijay Varma, Jackie Lodman, Scott E. Field, Gaurav Khanna, Mark A. Scheel, Harald P. Pfeiffer, Davide Gerosa, and Lawrence E. Kidder, Eccentric binary black hole surrogate models for the gravitational waveform and remnant properties: comparable mass, nonspinning case, Phys. Rev. D **103,** 064022 (2021).

[41] Alejandro Bohé *et al.*, Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors, Phys. Rev. D **95,** 044028 (2017).

[42] Roberto Cotesta, Alessandra Buonanno, Alejandro Bohé, Andrea Taracchini, Ian Hinder, and Serguei Ossokine,

Enriching the symphony of gravitational waves from binary black holes by tuning higher harmonics, Phys. Rev. D **98,** 084028 (2018).

[43] Roberto Cotesta, Sylvain Marsat, and Michael Pürrer, Frequency domain reduced order model of aligned-spin effective-one-body waveforms with higher-order modes, Phys. Rev. D **101,** 124040 (2020).

[44] Yi Pan, Alessandra Buonanno, Andrea Taracchini, Lawrence E. Kidder, Abdul H. Mroué, Harald P. Pfeiffer, Mark A. Scheel, and Béla Szilágyi, Inspiral-merger-ring-down waveforms of spinning, precessing black-hole binaries in the effective-one-body formalism, Phys. Rev. D **89,** 084006 (2014).

[45] Stanislav Babak, Andrea Taracchini, and Alessandra Buonanno, Validating the effective-one-body model of spinning, precessing binary black holes against numerical relativity, Phys. Rev. D **95,** 024010 (2017).

[46] Sascha Husa, Sebastian Khan, Mark Hannam, Michael Pürrer, Frank Ohme, Xisco Jiménez Forteza, and Alejandro Bohé, Frequency-domain gravitational waves from nonprecessing black-hole binaries. I. New numerical waveforms and anatomy of the signal, Phys. Rev. D **93,** 044006 (2016).

[47] Sebastian Khan, Sascha Husa, Mark Hannam, Frank Ohme, Michael Pürrer, Xisco Jiménez Forteza, and Alejandro Bohé, Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era, Phys. Rev. D **93,** 044007 (2016).

[48] Lionel London, Sebastian Khan, Edward Fauchon-Jones, Cecilio García, Mark Hannam, Sascha Husa, Xisco Jiménez-Forteza, Chinmay Kalaghatgi, Frank Ohme, and Francesco Pannarale, First Higher-Multipole Model of Gravitational Waves from Spinning and Coalescing Black-Hole Binaries, Phys. Rev. Lett. **120,** 161102 (2018).

[49] Sebastian Khan, Katerina Chatziioannou, Mark Hannam, and Frank Ohme, Phenomenological model for the gravitational-wave signal from precessing binary black holes with two-spin effects, Phys. Rev. D **100,** 024059 (2019).

[50] Alessandro Nagar, James Healy, Carlos O. Lousto, Sebastiano Bernuzzi, and Angelica Albertini, Numerical-relativity validation of effective-one-body waveforms in the intermediate-mass-ratio regime, Phys. Rev. D **105,** 124061 (2022).

[51] Carlos O. Lousto and James Healy, Exploring the Small Mass Ratio Binary Black Hole Merger via Zeno's Dichotomy Approach, Phys. Rev. Lett. **125,** 191102 (2020).

[52] Carlos O. Lousto and James Healy, Study of the intermediate mass ratio black hole binary merger up to 1000∶1 with numerical relativity, arXiv:2203.08831.

[53] Jooheon Yoo, Vijay Varma, Matthew Giesler, Mark A. Scheel, Carl-Johan Haster, Harald P. Pfeiffer, Lawrence E. Kidder, and Michael Boyle, Targeted large mass ratio numerical relativity surrogate waveform model for GW190814, Phys. Rev. D **106,** 044001 (2022).

[54] Matthew Giesler, Mark A. Scheel, and Saul A. Teukolsky, Numerical simulations of extreme mass ratio binary black holes (to be published).

[55] Alejandro Bohé *et al.*, Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of

gravitational-wave astrophysics with advanced detectors, Phys. Rev. D **95,** 044028 (2017).

[56] Leor Barack and Curt Cutler, LISA capture sources: Approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy, Phys. Rev. D **69,** 082005 (2004).

[57] Stanislav Babak, Hua Fang, Jonathan R. Gair, Kostas Glampedakis, and Scott A. Hughes, "Kludge" gravitational waveforms for a test-body orbiting a Kerr black hole, Phys. Rev. D **75,** 024005 (2007); **77,** 04990(E) (2008).

[58] Alvin J. K. Chua, Christopher J. Moore, and Jonathan R. Gair, Augmented kludge waveforms for detecting extreme-mass-ratio inspirals, Phys. Rev. D **96,** 044005 (2017).

[59] Alvin J. K. Chua and Jonathan R. Gair, Improved analytic extreme-mass-ratio inspiral model for scoping out eLISA data analysis, Classical Quantum Gravity **32,** 232002 (2015).

[60] Alvin J. K. Chua, Michael L. Katz, Niels Warburton, and Scott A. Hughes, Rapid Generation of Fully Relativistic Extreme-Mass-Ratio-Inspiral Waveform Templates for LISA Data Analysis, Phys. Rev. Lett. **126,** 051102 (2021).

[61] Michael L. Katz, Alvin J. K. Chua, Lorenzo Speri, Niels Warburton, and Scott A. Hughes, Fast extreme-mass-ratio-inspiral waveforms: New tools for millihertz gravitational-wave data analysis, Phys. Rev. D **104,** 064047 (2021).

[62] Barry Wardell, Adam Pound, Niels Warburton, Jeremy Miller, Leanne Durkan, and Alexandre Le Tiec, Gravitational waveforms for compact binaries from second-order self-force theory, arXiv:2112.12265.

[63] Nur E. M. Rifat, Scott E. Field, Gaurav Khanna, and Vijay Varma, Surrogate model for gravitational wave signals from comparable and large-mass-ratio black hole binaries, Phys. Rev. D **101,** 081502 (2020).

[64] Pranesh A. Sundararajan, Gaurav Khanna, and Scott A. Hughes, Towards adiabatic waveforms for inspiral into Kerr black holes. I. A new model of the source for the time domain perturbation equation, Phys. Rev. D **76,** 104005 (2007).

[65] Pranesh A. Sundararajan, Gaurav Khanna, Scott A. Hughes, and Steve Drasco, Towards adiabatic waveforms for inspiral into Kerr black holes: II. Dynamical sources and generic orbits, Phys. Rev. D **78,** 024022 (2008).

[66] Pranesh A. Sundararajan, Gaurav Khanna, and Scott A. Hughes, Binary black hole merger gravitational waves and recoil in the large mass ratio limit, Phys. Rev. D **81,** 104009 (2010).

[67] Anil Zenginoglu and Gaurav Khanna, Null Infinity Waveforms from Extreme-Mass-Ratio Inspirals in Kerr Spacetime, Phys. Rev. X **1,** 021017 (2011).

[68] S. E. Field, S. Gottlieb, Z. J. Grant, L. F. Isherwood, and G. Khanna, A GPU-accelerated mixed-precision WENO method for extremal black hole and gravitational wave physics computations, Commun. Appl. Math. Comput. Sci. (2021).

[69] Justin McKennon, Gary Forrester, and Gaurav Khanna, High accuracy gravitational waveforms from black hole binary inspirals using openCL, arXiv:1206.0270.

[70] Ryuichi Fujita and Hideyuki Tagoshi, New numerical methods to evaluate homogeneous solutions of the Teukolsky equation, Prog. Theor. Phys. **112,** 415 (2004).

[71] Ryuichi Fujita and Hideyuki Tagoshi, New numerical methods to evaluate homogeneous solutions of the

Teukolsky equation II. Solutions of the continued fraction equation, Prog. Theor. Phys. **113,** 1165 (2005).

[72] Shuhei Mano, Hisao Suzuki, and Eiichi Takasugi, Analytic solutions of the Teukolsky equation and their low frequency expansions, Prog. Theor. Phys. **95,** 1079 (1996).

[73] William William Thomas Throwe, High precision calculation of generic extreme mass ratio inspirals, Ph.D. thesis, Massachusetts Institute of Technology, 2010.

[74] Stephen O'Sullivan and Scott A. Hughes, Strong-field tidal distortions of rotating black holes: Formalism and results for circular, equatorial orbits, Phys. Rev. D **90,** 124039 (2014); **91,** 109901(E) (2015).

[75] Steve Drasco and Scott A. Hughes, Gravitational wave snapshots of generic extreme mass ratio inspirals, Phys. Rev. D **73,** 024027 (2006); **88,** 109905(E) (2013); **90,** 109905(E) (2014).

[76] Black Hole Perturbation Toolkit, http://bhptoolkit.org/.

[77] Amos Ori and Kip S. Thorne, The transition from inspiral to plunge for a compact body in a circular equatorial orbit around a massive, spinning black hole, Phys. Rev. D **62,** 124022 (2000).

[78] Scott A. Hughes, Anuj Apte, Gaurav Khanna, and Halston Lim, Learning About Black Hole Binaries from Their Ringdown Spectra, Phys. Rev. Lett. **123,** 161101 (2019).

[79] Anuj Apte and Scott A. Hughes, Exciting black hole modes via misaligned coalescences: I. Inspiral, transition, and plunge trajectories using a generalized Ori-Thorne procedure, Phys. Rev. D **100,** 084031 (2019).

[80] Tanja Hinderer and Eanna E. Flanagan, Two timescale analysis of extreme mass ratio inspirals in Kerr. I. Orbital motion, Phys. Rev. D **78,** 064028 (2008).

[81] Scott E. Field, Chad R. Galley, Jan S. Hesthaven, Jason Kaye, and Manuel Tiglio, Fast Prediction and Evaluation of Gravitational Waveforms Using Surrogate Models, Phys. Rev. X **4,** 031006 (2014).

[82] Michael Pürrer, Frequency domain reduced order models for gravitational waves from aligned-spin compact binaries, Classical Quantum Gravity **31,** 195010 (2014).

[83] Y. Maday, N. C. Nguyen, A. T. Patera, and S. H. Pau, A general multipurpose interpolation procedure: the magic points, Commun. Pure Appl. Anal. **8,** 383 (2009).

[84] Saifon Chaturantabut and Danny C Sorensen, Nonlinear model reduction via discrete empirical interpolation, SIAM J. Sci. Comput. **32,** 2737 (2010).

[85] Priscilla Canizares, Scott E. Field, Jonathan Gair, Vivien Raymond, Rory Smith, and Manuel Tiglio, Accelerated Gravitational-Wave Parameter Estimation with Reduced Order Modeling, Phys. Rev. Lett. **114,** 071104 (2015).

[86] Vijay Varma, Davide Gerosa, Leo C. Stein, François Hébert, and Hao Zhang, High-Accuracy Mass, Spin, and Recoil Predictions of Generic Black-Hole Merger Remnants, Phys. Rev. Lett. **122,** 011101 (2019).

[87] Michael Boyle, Duncan A. Brown, Lawrence E. Kidder, Abdul H. Mroue, Harald P. Pfeiffer, Mark A. Scheel, Gregory B. Cook, and Saul A. Teukolsky, High-accuracy comparison of numerical relativity simulations with post-Newtonian expansions, Phys. Rev. D **76,** 124038 (2007).

[88] Michael Boyle *et al.*, The SXS Collaboration catalog of binary black hole simulations, Classical Quantum Gravity **36,** 195006 (2019).

[89] SXS Collaboration, The SXS Collaboration catalog of gravitational waveforms, http://www.black-holes.org/waveforms.

[90] Abdul H. Mroue *et al.*, Catalog of 174 Binary Black Hole Simulations for Gravitational Wave Astronomy, Phys. Rev. Lett. **111,** 241104 (2013).

[91] D. J. A. McKechan, C. Robinson, and B. S. Sathyaprakash, A tapering window for time-domain templates and simulated signals in the detection of gravitational waves from coalescing compact binaries, Classical Quantum Gravity **27,** 084020 (2010).

[92] Halston Lim, Gaurav Khanna, Anuj Apte, and Scott A. Hughes, Exciting black hole modes via misaligned coalescences: II. The mode content of late-time coalescence waveforms, Phys. Rev. D **100,** 084032 (2019).

[93] Vijay Varma and Parameswaran Ajith, Effects of nonquadrupole modes in the detection and parameter estimation of black hole binaries with nonprecessing spins, Phys. Rev. D **96,** 124024 (2017).

[94] Vijay Varma, Parameswaran Ajith, Sascha Husa, Juan Calderon Bustillo, Mark Hannam, and Michael Pürrer, Gravitational-wave observations of binary black holes: Effect of nonquadrupole modes, Phys. Rev. D **90,** 124004 (2014).

[95] Feroz H. Shaik, Jacob Lange, Scott E. Field, Richard O'Shaughnessy, Vijay Varma, Lawrence E. Kidder, Harald P. Pfeiffer, and Daniel Wysocki, Impact of subdominant modes on the interpretation of gravitational-wave signals from heavy binary black hole systems, Phys. Rev. D **101,** 124054 (2020).

[96] Scott Field, Tousif Islam, Gaurav Khanna, Nur Rifat, and Vijay Varma, BHPTNRSurrogate, http://bhptoolkit.org/BHPTNRSurrogate/.

[97] Jonathan Blackman, Scott Field, Chad Galley, and Vijay Varma, gwsurrogate, https://pypi.python.org/pypi/gwsurrogate/.