Surrogate models for precessing binary black hole simulations with unequal masses

Vijay Varma [1,*] Scott E. Field,[2] Mark A. Scheel,[1] Jonathan Blackman,[1] Davide Gerosa [3] Leo C. Stein,[4]
Lawrence E. Kidder [5] and Harald P. Pfeiffer [6]

[1] *TAPIR 350-17, California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, USA*
[2] *Department of Mathematics, Center for Scientific Computing and Visualization Research,*
University of Massachusetts, Dartmouth, Massachusetts 02747, USA
[3] *School of Physics and Astronomy and Institute for Gravitational Wave Astronomy, University of Birmingham,*
Birmingham, B15 2TT, United Kingdom
[4] *Department of Physics and Astronomy, University of Mississippi, University, Mississippi 38677, USA*
[5] *Center for Radiophysics and Space Research, Cornell University, Ithaca, New York 14853, USA*
[6] *Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlenberg 1, Potsdam 14476, Germany*

Only numerical relativity simulations can capture the full complexities of binary black hole mergers. These simulations, however, are prohibitively expensive for direct data analysis applications such as parameter estimation. We present two fast and accurate surrogate models for the outputs of these simulations: the first model, NRSur7dq4, predicts the gravitational waveform and the second model, NRSur7dq4Remnant, predicts the properties of the remnant black hole. These models extend previous seven-dimensional, noneccentric precessing models to higher mass ratios and have been trained against 1528 simulations with mass ratios $q \leqslant 4$ and spin magnitudes $\chi_1, \chi_2 \leqslant 0.8$, with generic spin directions. The waveform model, NRSur7dq4, which begins about 20 orbits before merger, includes all $\ell \leqslant 4$ spin-weighted spherical harmonic modes, as well as the precession frame dynamics and spin evolution of the black holes. The final black hole model, NRSur7dq4Remnant, models the mass, spin, and recoil kick velocity of the remnant black hole. In their training parameter range, both models are shown to be more accurate than existing models by at least an order of magnitude, with errors comparable to the estimated errors in the numerical relativity simulations. We also show that the surrogate models work well even when extrapolated outside their training parameter space range, up to mass ratios $q = 6$.

I. INTRODUCTION

As the LIGO [1] and Virgo [2] detectors reach their design sensitivity, gravitational wave (GW) detections [3–9] are becoming routine [10,11]. To maximize the science output of the data collected by the network of detectors, it is crucial to accurately model the source of the GWs. Among the most important sources for these detectors are binary black hole (BBH) systems, in which two black holes (BHs) lose energy through GWs, causing them to inspiral and eventually merge.

Numerical relativity (NR) simulations are necessary to accurately model the late inspiral and merger stages of the BBH evolution. These simulations accurately solve Einstein's equations to predict the evolution of the BBH spacetime. The most important outputs of NR simulations are the gravitational waveform and the mass, spin, and recoil kick velocity of the remnant BH left after the merger.

For interpreting detected signals, model waveforms are used in comparison with detector data so we can infer the properties of the source [12–14]. The mass and spin of the remnant determine the black hole ringdown frequencies, which are used in testing general relativity [15–17]. In addition, the recoil kick is astrophysically important because it can cause the remnant BH to be ejected from its host galaxy [18–20].

Unfortunately, NR simulations are too expensive to be directly used in data analysis applications and incorporated into astrophysical models. As a result, several approximate models that are much faster to evaluate have been developed for both waveforms [21–31] and remnant properties [18,19,32–50]. These models typically assume an underlying phenomenology based on physical motivations and calibrate any remaining free parameters to NR simulations.

Among BBHs, systems with BH spins that are misaligned with respect to the orbital angular momentum are complicated to model analytically or semianalytically. For these systems, the spins interact with both the orbital angular momentum and each other, causing the system to precess about the direction of the total angular momentum [51]. This precession is imprinted on the waveform as characteristic modulations in the amplitude and frequency of the GWs and can be used to extract information about the spins of the source. One

---

*vvarma@caltech.edu

important application of the extracted spins is to distinguish between formation channels of BBHs [52–55].

The precessing BBH problem for quasicircular orbits is parametrized by seven parameters: the mass ratio $q = m_1/m_2 \geqslant 1$ and two spin vectors $\boldsymbol{\chi}_{1,2}$, where the index 1 (2) refers to the heavier (lighter) BH. The total mass scales out of the problem and does not constitute an additional parameter for modeling. The surrogate models of Ref. [56] for the gravitational waveform and Ref. [57] for the remnant properties were the first to model the seven-dimensional space of generically precessing BBH systems, although restricted to mass ratios $q \leqslant 2$ and dimensionless spin magnitudes $\chi_{1,2} \leqslant 0.8$. Trained directly against numerical simulations, these models do not need to introduce additional assumptions about the underlying phenomenology of the waveform or remnant properties that necessarily introduce some systematic error. Through cross-validation studies, it was shown that both these models achieve accuracies comparable to the numerical simulations themselves [56,57], and as a result, are the most accurate models currently available for precessing systems, within their parameter space of validity.

In this paper, we present extensions of the above surrogate models to larger mass ratios. Our new surrogate models are called NRSur7dq4 and NRSur7dq4Remnant, for the gravitational waveform and remnant properties, respectively. They are trained against 1528 precessing NR simulations with mass ratios $q \leqslant 4$, spin magnitudes $\chi_1, \chi_2 \leqslant 0.8$, and generic spin directions. Both models are made publicly available through the gwsurrogate [58] and surfinBH [59] PYTHON packages; example evaluation codes are provided at Refs. [60] and [59], respectively, for NRSur7dq4 and NRSur7dq4Remnant.

The rest of the paper is organized as follows. Section II covers some preliminaries to set up the modeling problem for precessing BBH systems. Section III describes the training simulations. Section IV describes the NRSur7dq4 waveform surrogate model. Section V describes the NR-Sur7dq4Remnant remnant properties surrogate model. Section VI compares these models against NR simulations to assess their accuracy. Finally, Sec. VII presents some concluding remarks. In Appendix A, we examine how accurate these models are when extrapolated beyond mass ratio $q = 4$, and in Appendix B, we investigate some features in the error distribution of the NR simulations.

II. PRELIMINARIES AND NOTATION

It is convenient to combine the two polarizations of the waveform into a single complex, dimensionless strain $h = h_+ - ih_\times$ and to represent the waveform on a sphere as a sum of spin-weighted spherical harmonic modes:

$$h(t, \iota, \varphi_0) = \sum_{\ell=2}^{\infty} \sum_{m=-l}^{l} h_{\ell m}(t) \, _{-2}Y_{\ell m}(\iota, \varphi_0). \tag{1}$$

Here, $_{-2}Y_{\ell m}$ are the spin $= -2$ weighted spherical harmonics, and ι and φ_0 are the polar and azimuthal angles on the sky in the source frame.

For nonprecessing systems, the direction of orbital angular momentum ($\boldsymbol{L}$) is fixed and the $\hat{z}$ direction of the source frame is chosen to be along $\hat{\boldsymbol{L}}$ by convention. The gravitational
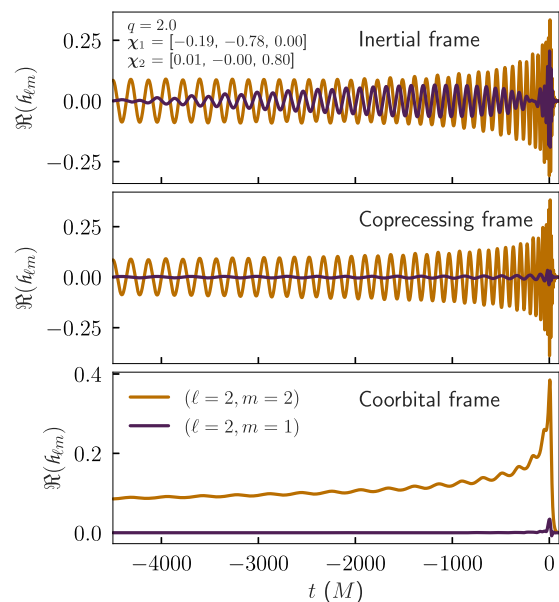

FIG. 1. The real part of the (2,2) and (2,1) modes of the gravitational waveform in the inertial (top), coprecessing (middle), and coorbital (bottom) frames. In the inertial frame, the amplitude of the (2,1) mode can be comparable to that of the (2,2) mode. In the coprecessing frame, on the other hand, the (2,2) mode always dominates. In addition, most effects of precession are removed by the rotation and the waveform in the coprecessing frame resembles that of a nonprecessing system. In the co-orbital frame, finally, the waveform is further simplified and does not oscillate about zero. Mass ratio and initial spins used to produce this figure are indicated in the text within the figure.

radiation is strongest along the directions parallel and antiparallel to $\hat{\boldsymbol{L}}$. Therefore, for nonprecessing systems, the quadrupole modes ($\ell = 2, m = \pm 2$) dominate the sum in Eq. (1), but the nonquadrupole modes can become important at large mass ratios or ι close to $\pi/2$ [61–70].

By contrast, for precessing systems the direction of $\boldsymbol{L}$ varies due to precession [51] and so there is not a fixed axis along which the radiation is dominant. The standard practice is to choose $\hat{z}$ of the source frame along the direction of $\boldsymbol{L}$ (or the total angular momentum) at a reference time or frequency.

Heuristically, one can think of a precessing system as a nonprecessing system with time-dependent frame rotations applied to it. In this noninertial frame, the rotation causes mixing of power between modes of fixed ℓ. For example, the power of the $(2, \pm 2)$ modes leaks into the $(2, \pm 1)$ and $(2,0)$ modes. This means that all $\ell = 2$ modes can be dominant in Eq. (1). While this rotating-frame picture ignores some dynamical features such as nutation, it accounts for most of the effects of precession in the waveform.

By the same logic, one could apply a time-dependent rotation to a precessing system such that $\hat{z}$ always lies along $\hat{\boldsymbol{L}}(t)$. In this noninertial frame, referred to as the coprecessing frame [71–73], the radiation is always strongest along $\hat{z}$, and the $(\ell = 2, m = \pm 2)$ modes are dominant. In fact, since most precessional effects are accounted for by the frame rotation, the waveform in the coprecessing frame is qualitatively similar to that of a nonprecessing system (cf. Fig. 1). This

observation has been exploited in the literature [21,24,27,56,74] to simplify the modeling of precessing systems. Here, we proceed similarly, using the coprecessing frame described in Ref. [73] and denoting the strain in this frame as $h_{\ell m}^{\mathrm{copr}}$.

The waveform can be made even simpler, and therefore easier to model, by applying an additional rotation about the z axis of the coprecessing frame by an amount equal to the instantaneous orbital phase:

$$h_{\ell m}^{\mathrm{coorb}}(t) = h_{\ell m}^{\mathrm{copr}}(t)\, e^{im\phi(t)}. \qquad (2)$$

Here, we define the orbital phase,

$$\phi(t) = \frac{\arg\left[h_{2,-2}^{\mathrm{copr}}(t)\right] - \arg\left[h_{2,2}^{\mathrm{copr}}(t)\right]}{4}, \qquad (3)$$

using the coprecessing frame strain. The waveform $h_{\ell m}^{\mathrm{coorb}}(t)$ corresponds to a new frame, called the co-orbital frame, in which the BHs are always on the x axis, with the heavier BH on the positive x axis.[1] More importantly, the waveform in the co-orbital frame is nearly nonoscillatory, simplifying the modeling problem greatly. Figure 1 shows an example of a waveform in the inertial, coprecessing, and co-orbital frames.

III. NR SIMULATIONS

Our NR simulations are performed using the spectral Einstein code (SpEC) [75–80] developed by the SXS [81] Collaboration.

A. Parameter space coverage

We use 890 precessing NR simulations used in the construction of the surrogate models of Refs. [56,57], which provide coverage in the $q \leqslant 2$ and $\chi_1, \chi_2 \leqslant 0.8$ regions of the parameter space. We also make use of 64 aligned-spin simulations with $q \leqslant 4$ and $\chi_1, \chi_2 \leqslant 0.8$ used in the construction of the surrogate model presented in Ref. [82]. Finally, we performed 574 new simulations with $2 < q \leqslant 4$, $\chi_1, \chi_2 \leqslant 0.8$, and generic spin directions. The parameters for the first 204 of these are chosen based on sparse grids as detailed in Appendix A of Ref. [56]. The remaining parameters are chosen as follows. We randomly sample 1000 points uniformly in mass ratio, spin magnitude, and spin direction on the sphere. We compute the distance between points a and b using the metric

$$ds^2 = \left(\frac{q^a - q^b}{\Delta q}\right)^2 + \sum_{i \in \{1,2\}} \left(\frac{|\chi_i^a - \chi_i^b|}{\Delta \chi}\right)^2, \qquad (4)$$

where $\Delta q = 4 - 1 = 3$ and $\Delta \chi = 0.8$ are the ranges of these parameters. These normalization factors are somewhat arbitrary, although any choice of order unity should provide a reasonable criteria for point selection. For each sampled parameter, we compute the minimum distance to all previously chosen parameters. We then add the sampled parameter maximizing this minimum distance to the set

of chosen parameters. This is done iteratively for 370 additional parameters. The new simulations have identifiers SXS:BBH:1346-1350 and SXS:BBH:1514-2082 and are made publicly available through the SXS public catalog [83]. The parameter space covered by the $890 + 64 + 574 = 1528$ NR simulations used in this work is shown in Fig. 2. Note that not all of these are independent simulations: For 154 of these cases, we have $q = 1$, with $\chi_1 \neq \chi_2$; for each of these cases, we effectively obtain an additional simulation by exchanging the labels of the two BHs.

The start time of these simulations varies between $4693M$ and $5234M$ before the peak of the waveform amplitude, where $M = m_1 + m_2$ is the total Christodoulou mass measured close to the beginning of the simulation at the "relaxation time" [84]. The initial orbital parameters are chosen through an iterative procedure [85] such that the orbits are quasicircular; the largest eccentricity for these simulations is 9.8×10^{-4}, while the median value is 3.8×10^{-4}.

B. Data extracted from simulations

We make use of the following quantities extracted from the NR simulations: the waveform modes $h_{\ell m}(t)$, component spins $\chi(t)$, mass ratio q, remnant mass m_f, spin χ_f, and kick velocity $\boldsymbol{v}_f$.

The waveform is extracted at several extraction spheres at varying finite radii from the origin and then extrapolated to future null infinity [84,86]. Then, the extrapolated waveforms are corrected to account for the initial drift of the center of mass [87,88]. The time steps during the simulations are chosen nonuniformly using an adaptive time stepper [84]. Using cubic splines, we interpolate the real and imaginary parts of the waveform modes to a uniform time step of $0.1M$; this is dense enough to capture all frequencies of interest, including near merger. The interpolated waveform at future null infinity, scaled to unit mass and unit distance, is denoted as $h_{\ell m}(t)$ in this paper.

The component spins $\chi_{1,2}(t)$ and masses $m_{1,2}$ are evaluated on the apparent horizons [77] of the BHs. The masses at the relaxation time [84] are used to define the mass ratio $q = m_1/m_2$. Unless otherwise specified, all masses in this paper are given in units of the total mass $M = m_1 + m_2$ at relaxation. The spins are interpolated onto the same time array[2] as used for the waveform, using cubic splines.

The remnant mass m_f and spin χ_f are determined from the common apparent horizon long after ringdown, as detailed in Ref. [84]. The remnant kick velocity is derived from conservation of momentum, $\boldsymbol{v}_f = -\boldsymbol{P}^{\mathrm{rad}}/m_f$ [90]. The radiated momentum flux $\boldsymbol{P}^{\mathrm{rad}}$ is integrated [91] from the strain $h_{\ell m}$.

---

[1]Here, the BH positions are defined from the waveform at future null infinity and do not necessarily correspond to the (gauge-dependent) coordinate BH positions in the NR simulation.

[2]The waveforms at future null infinity use a time coordinate t that is different from the simulation time $\tilde{t}$ at which the spins are measured in the near zone [84]. In this paper, we identify t with $\tilde{t}$. While this identification is gauge dependent, the spin directions are already gauge dependent. We, however, note that the spin and orbital angular momentum vectors in the damped harmonic gauge used by SpEC agree quite well with the corresponding vectors in post-Newtonian (PN) theory [89].
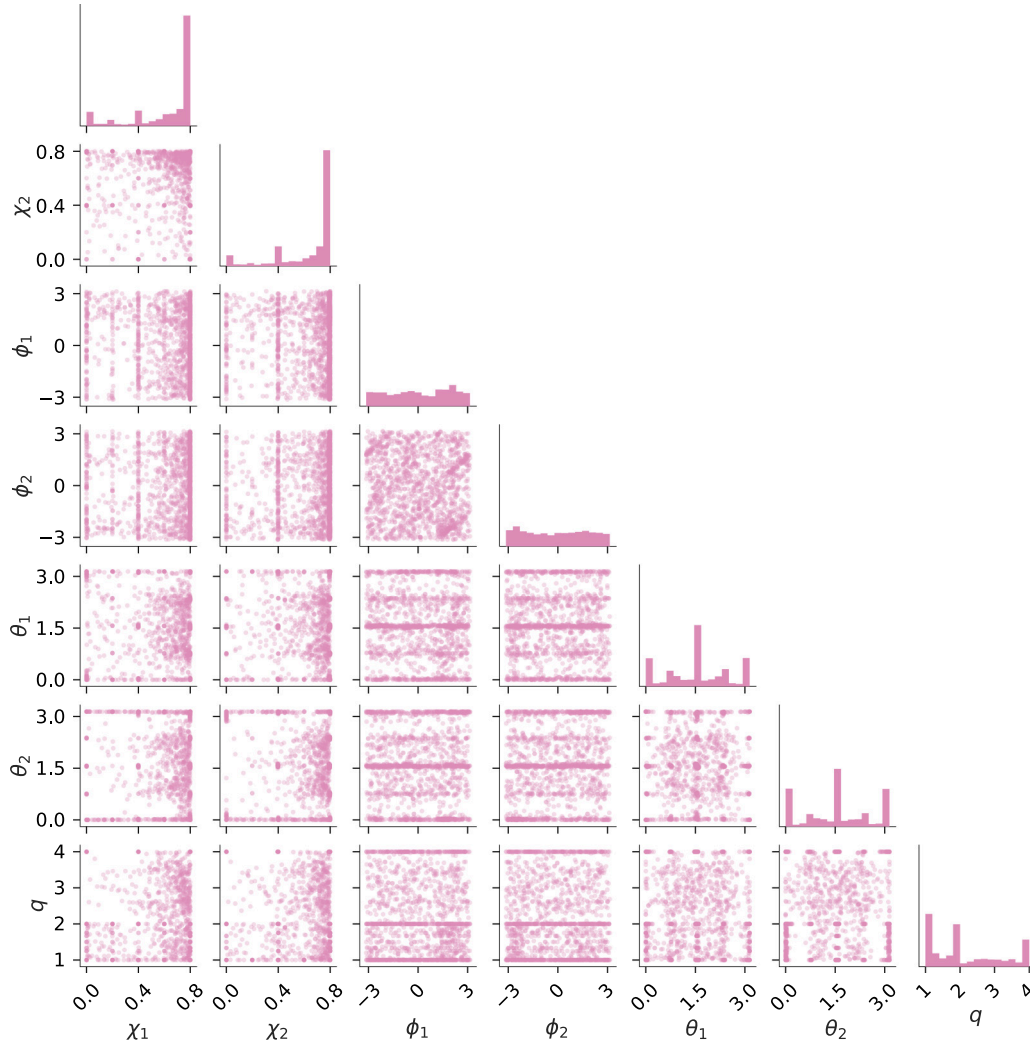
FIG. 2. Parameters of the 1528 NR simulations used in the construction of the surrogate models in this paper. We show the distribution of mass ratio q and the spin components in standard spherical polar coordinates (χ, θ, ϕ) at $-4300M$ from the waveform amplitude peak. The index 1 (2) refers to the heavier (lighter) BH.

C. Postprocessing the output of NR simulations

After extracting the strain and spins from the simulations, we apply the following postprocessing steps before building the surrogate models.

First, we shift the time arrays of all waveforms such that $t = 0$ occurs at the peak (see Ref. [56] for how the peak is determined) of the total waveform amplitude, defined as

$$A(t) = \sqrt{\sum_{\ell m} |\hbar_{\ell m}(t)|^2}. \tag{5}$$

Then we rotate the waveform modes such that at a reference time $t_0 = -4300M$, the inertial frame coincides with the co-orbital frame. This means that the $\hat{z}$ direction of the inertial frame is along the principal eigenvector of the angular momentum operator [73] at the reference time. In addition, the $\hat{x}$ direction of the inertial frame is along the line of separation from the lighter BH to the heavier BH (in other words, the orbital phase is zero). The spin vectors $\boldsymbol{\chi}_{1,2}(t)$ are also transformed into the same inertial frame.

We then truncate the waveform and spin time series by dropping all times $t < -4300M$ to exclude the initial transients known as "junk radiation." After the truncation, the reference time $t = -4300M$ is also the start time of the data.

For $t > -100M$, the spin measurements from the apparent horizons start to become unreliable as the horizons become highly distorted. Following Ref. [56], starting at $t = -100M$, we extend the spins to later times using PN spin evolution equations. This evolution is done even past the merger stage, into the ringdown. We stress that the extended spins are unphysical but are a useful parametrization to construct fits at late times.

Finally, we apply a smoothing filter (see Eq. (6) of Ref. [56]) on the spin time series to remove fast oscillations taking place on the orbital time scale. This smoothing helps improve the numerical stability of the ordinary differential equation (ODE) integrations described in Sec. IV B. Note that we use the filtered spins for the waveform surrogate (Sec. IV) but not for the remnant surrogate (Sec. V), for which we just
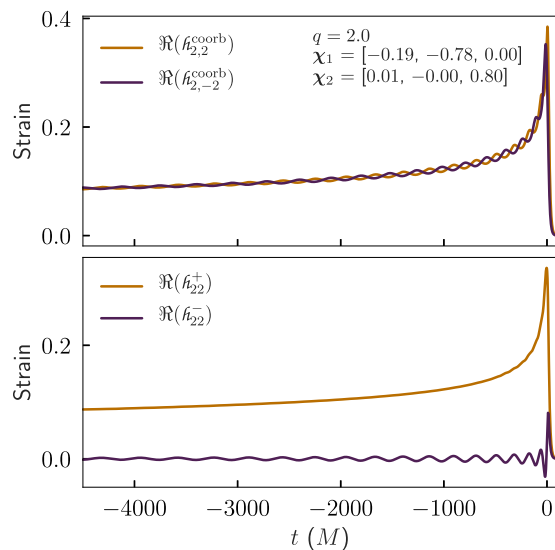
FIG. 3. The top panel shows the real part of the (2,2) and $(2, -2)$ modes of the waveform in the co-orbital frame. Notice that the orbital timescale oscillations of these two modes have opposite signs. The bottom panel shows the real parts of $\hbar_{2,2}^+$ and $\hbar_{2,2}^-$ [cf. Eq. (6)], we take advantage of the above fact to move most of the oscillations from the larger to the smaller data piece.

use the unfiltered spins since there are no ODE integrations involved.

IV. WAVEFORM SURROGATE

To construct the waveform surrogate, we closely follow the model of Ref. [56], with some modifications to adapt it to higher mass ratios. We refer to the new waveform model as NRSur7dq4.

A. Co-orbital frame surrogate

We find that the surrogate accuracy improves when working with slowly varying functions, rather than oscillatory ones. Therefore, we first decompose the strain into several "data pieces," each of which is a slowly varying function of time, and build a surrogate for each of them. At evaluation time, we combine the various data pieces to reconstruct the inertial frame strain. To reduce the cost of these transformations, we first downsample the inertial frame strain onto a set of 2000 time values t_i^{coorb} that are approximately uniformly spaced in the orbital phase (using the method described in Appendix B of Ref. [56]).

As described in Sec. II, the waveform is simpler in the co-orbital frame. A further simplification is possible by considering combinations of $m > 0$ and $m < 0$ counterparts of a fixed ℓ mode:

$$\hbar_{\ell m}^{\pm} = \frac{h_{\ell,m}^{\text{coorb}} \pm h_{\ell,-m}^{\text{coorb}\,*}}{2}. \quad (6)$$

Figure 3 shows an example of the simplification obtained with this combination. For all $m \neq 0$ modes, we model the real and imaginary parts of $\hbar_{\ell m}^{\pm}$. For $m = 0$ modes, we directly model the real and imaginary parts of the co-orbital frame

strain $\hbar_{\ell,m}^{\text{coorb}}$. We construct an independent surrogate model for each of these data pieces and refer to the combination of these models as the co-orbital frame surrogate.

As described in Ref. [56], for each waveform data piece, we construct a linear basis using singular value decomposition with an rms tolerance of 3×10^{-4}. We then construct an empirical time interpolant with the same number of empirical time nodes as basis functions for that data piece [92–94]. The empirical time nodes are chosen as a subset of the 2000 co-orbital time values (t_i^{coorb}) and are specific to each data piece. Finally, for each empirical time node, we construct a parametric fit for the waveform data piece. The fits are parametrized as functions of the mass ratio and the spins in the co-orbital frame at that time. We describe our fitting procedure in Sec. IV C. At evaluation time, the co-orbital frame spins at any time are obtained using the dynamics surrogate described in Sec. IV B.

B. Dynamics surrogate

The surrogate described in Sec. IV A only models the strain in the co-orbital frame. We also need to model the following quantities:

(1) the orbital phase in the coprecessing frame, which is required to transform the strain from the co-orbital frame to the coprecessing frame [cf. Eq. (2)];

(2) the quaternions describing the coprecessing frame, which are required to transform the strain from the coprecessing frame to the inertial frame;

(3) the spins as a function of time, which are used in the evaluation of the parametric fits described in Sec. IV C.

We refer to the model for these quantities as the dynamics surrogate. Using the fitting method of Sec. IV C, we first construct parametric fits for $\omega(t)$, $\Omega_{x,y}^{\text{coorb}}(t)$, and $\dot{\chi}_{1,2}^{\text{coorb}}(t)$ at selected time nodes referred to as the dynamical time nodes t_i^{dyn}. Here, $\dot{\chi}_{1,2}^{\text{coorb}}(t)$ are the time derivatives of the coprecessing frame spins transformed to the co-orbital frame, $\omega(t)$ is $d\phi/dt$ [cf. Eq. (3)], and $\Omega_{x,y}^{\text{coorb}}(t)$ is the angular velocity of the coprecessing frame, transformed to the co-orbital frame. These quantities are described in more detail in Sec. III of Ref. [56]. Note that $\Omega_z^{\text{coorb}}(t) \sim 0$. For the dynamical time nodes t_i^{dyn}, we chose 238 time values such that there are approximately 10 nodes per orbit (see Appendix B of Ref. [56] for details).

We use a fourth-order Adams-Bashforth scheme to integrate $\omega(t_i^{\text{dyn}})$, $\Omega_{x,y}^{\text{coorb}}(t_i^{\text{dyn}})$, and $\dot{\chi}_{1,2}^{\text{coorb}}(t_i^{\text{dyn}})$ over the set of dynamical time nodes t_i^{dyn} providing the time evolution of the orbital phase $\phi(t_i^{\text{dyn}})$, the coprecessing frame quaternions $\hat{Q}(t_i^{\text{dyn}})$, and the component spins in the co-orbital frame $\chi_{1,2}^{\text{coorb}}(t_i^{\text{dyn}})$. This involves solving a coupled ODE as described in Sec. V of Ref. [56]. At each step of the ODE integration, the co-orbital frame spins at the current node t_i^{dyn} are first obtained. These are then used to evaluate the parametric fits for the derivative quantities mentioned above. Note that the spins used in the dynamics surrogate are the filtered spins mentioned in Sec. III C; this improves the accuracy of the ODE integration by making the spin time derivatives easier to model.

C. Parametric fits

For the co-orbital frame surrogate of Sec. IV A, we need to construct parametric fits at various empirical time nodes for the different data pieces. Similarly, for the dynamics surrogate of Sec. IV B, we need to construct fits for various time derivatives at the dynamical time nodes t_i^{dyn}. We use the same procedure for each of these fits. Let us refer to the data to be fitted as $y(\mathbf{\Lambda})$, where $\mathbf{\Lambda}$ is a seven-dimensional set of parameters.

For each of these fits, the seven parameters $\mathbf{\Lambda}$ must contain information on mass ratio q and co-orbital frame spins $\chi_{1,2}^{\mathrm{coorb}}(t_i)$ at the time corresponding to the fit. Following Ref. [57], we parametrize the fits using

$$\mathbf{\Lambda} = \left[\log(q), \chi_{1x}^{\mathrm{coorb}}, \chi_{1y}^{\mathrm{coorb}}, \hat{\chi}^{\mathrm{coorb}}, \chi_{2x}^{\mathrm{coorb}}, \chi_{2y}^{\mathrm{coorb}}, \chi_{a}^{\mathrm{coorb}} \right], \tag{7}$$

where $\hat{\chi}^{\mathrm{coorb}}$ is the spin parameter entering the GW phase at leading order [12,26,95,96] in the PN expansion

$$\hat{\chi}^{\mathrm{coorb}} = \frac{\chi_{\mathrm{eff}}^{\mathrm{coorb}} - 38\eta\left(\chi_{1z}^{\mathrm{coorb}} + \chi_{2z}^{\mathrm{coorb}}\right)/113}{1 - 76\eta/113}, \tag{8}$$

$$\chi_{\mathrm{eff}}^{\mathrm{coorb}} = \frac{q\chi_{1z}^{\mathrm{coorb}} + \chi_{2z}^{\mathrm{coorb}}}{1 + q}, \tag{9}$$

$$\eta = \frac{q}{(1+q)^2}, \tag{10}$$

and χ_a^{coorb} is the "antisymmetric spin"

$$\chi_a^{\mathrm{coorb}} = \tfrac{1}{2}\left(\chi_{1z}^{\mathrm{coorb}} - \chi_{2z}^{\mathrm{coorb}}\right). \tag{11}$$

We empirically found this parametrization to perform more accurately than the more intuitive choice $\mathbf{\Lambda}_{\mathrm{ref56}} = [q, \chi_{1x}^{\mathrm{coorb}}, \chi_{1y}^{\mathrm{coorb}}, \chi_{1z}^{\mathrm{coorb}}, \chi_{2x}^{\mathrm{coorb}}, \chi_{2y}^{\mathrm{coorb}}, \chi_{2z}^{\mathrm{coorb}}]$ used in Ref. [56].

Fits are constructed using the forward-stepwise greedy fitting method described in Appendix A of Ref. [74]. We choose the basis functions to be a tensor product of one-dimensional (1D) monomials in the components of $\mathbf{\Lambda}$. The components of $\mathbf{\Lambda}$ are first affine mapped to the interval $[-1, 1]$ before constructing the tensor product. We consider up to cubic powers in $\log(q)$ and up to quadratic powers in the spin parameters. We find that going to higher powers does not significantly improve the fit accuracy within the training region, but the mass ratio extrapolation errors estimated in Appendix A become much larger.

It is always possible to improve the accuracy of a fit by adding more basis functions. However, this can lead to overfitting when the data contain some noise. Our source of noise is mostly due to NR truncation error, but also systematic errors such as waveform extrapolation and residual eccentricity. In order to safeguard against overfitting, we perform 10 trial fits, leaving a random 10% of the dataset out as validation points in each trial, to determine the set of basis functions used in constructing the final fit. We allow a maximum of 100 basis functions for each fit. See Appendix A of Ref. [74] for more details.

D. Surrogate evaluation

To evaluate the surrogate, we begin with a user-specified mass ratio q and spins $\chi_{1,2}^{\mathrm{coorb}}$ at the initial time $t = -4300M$.

Note that at this time, the inertial frame coincides with the co-orbital frame. These values are used to initialize the dynamics surrogate described in Sec. IV B, which predicts the coprecessing frame quaternions $\hat{Q}(t_i^{\mathrm{dyn}})$, the orbital phase $\phi(t_i^{\mathrm{dyn}})$ in the coprecessing frame, and the coorbital frame spins $\chi_{1,2}^{\mathrm{coorb}}(t_i^{\mathrm{dyn}})$ at the dynamic time nodes t_i^{dyn}. We then use cubic splines to interpolate these quantities on to the time array for the co-orbital frame surrogate t_i^{coorb}, giving us $\hat{Q}(t_i^{\mathrm{coorb}})$, $\phi(t_i^{\mathrm{coorb}})$, and $\chi_{1,2}^{\mathrm{coorb}}(t_i^{\mathrm{coorb}})$.

The co-orbital frame surrogate described in Sec. IV A is used to predict the strain in the co-orbital frame. This involves evaluating the fits at the empirical time nodes for this surrogate using $\chi_{1,2}^{\mathrm{coorb}}(t_i^{\mathrm{coorb}})$ and q. Then, the orbital phase $\phi(t_i^{\mathrm{coorb}})$ is used to transform the strain from the co-orbital frame to the coprecessing frame [cf. Eq. (2)]. Finally, the coprecessing frame quaternions $\hat{Q}(t_i^{\mathrm{coorb}})$ are used to transform the strain from the coprecessing frame to the inertial frame (this involves Wigner matrices; see Appendix A of Ref. [73]). This gives us $\hat{h}_{\ell m}(t_i^{\mathrm{coorb}})$, which is interpolated onto any required time array t using cubic splines to get $\hat{h}_{\ell m}(t)$.

V. REMNANT SURROGATE

To construct the remnant properties surrogate, we closely follow the model of Ref. [57]. We refer to the new model presented here as NRSur7dq4Remnant.

We model the remnant mass m_f, spin $\mathbf{\chi}_f$, and kick velocity $\mathbf{v}_f$. Before constructing the fits, $\mathbf{\chi}_f$ and $\mathbf{v}_f$ are transformed into the co-orbital frame at $t = -100M$. We model each component of the vectors independently. The fits are parametrized by the same $\mathbf{\Lambda}$ of Eq. (7), but using the component spins at $t = -100M$. Unlike the waveform surrogate case, we do not filter out orbital-timescale oscillations. The filtered spins were found to be necessary for the accuracy of the time integration in Sec. IV B, which is not necessary here because the remnant properties can evaluated from the BBH parameters at a single time $t = -100M$.

All fits are performed using Gaussian process regression (GPR), as described in the supplementary materials of Ref. [57]. We find that GPR fitting is, in most cases, more accurate but also significantly more expensive than the polynomial fitting method described in Sec. IV C. GPR becomes impractical to use for the waveform surrogate as there are hundreds of fits that need to be evaluated to generate the waveform. For the remnant fits, however, the additional cost of GPR is acceptable because one is only fitting seven quantities $(m, \mathbf{\chi}_f, \mathbf{v}_f)$. In addition, GPR naturally provides error estimates which can be useful in data analysis applications. The efficacy of the GPR error estimate in reproducing the underlying error of the surrogate models was investigated thoroughly in the supplementary materials of Ref. [57].

Although NRSur7dq4Remnant is parameterized internally by input spins specified in the co-orbital frame at $t = -100M$, we allow the user to specify input spins at earlier times and in the inertial frame; this case is handled by two additional levels of spin evolution. Given the inertial-frame input spins at an initial orbital frequency f_0, we first evolve the spins using a post-Newtonian (PN) approximant—3.5PN SpinTaylorT4 [89,97,98]—until we reach the domain of validity of the
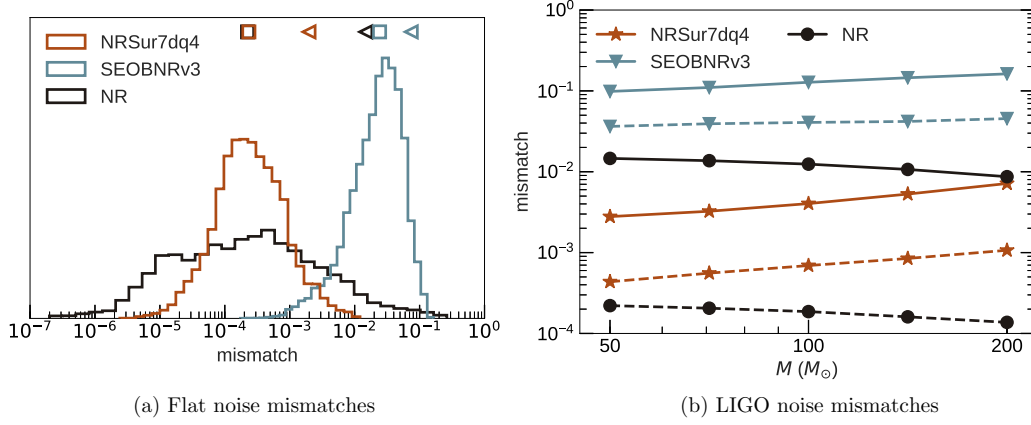
(a) Flat noise mismatches

(b) LIGO noise mismatches

FIG. 4. Mismatches for NRSur7dq4 and SEOBNRv3 models, when compared against precessing NR simulations using all $\ell \leqslant 5$ modes with mass ratios $q \leqslant 4$ and spin magnitudes $\chi_1, \chi_2 \leqslant 0.8$. The NRSur7dq4 errors shown are out-of-sample errors. Also shown are the NR resolution errors. Mismatches are computed at several sky locations using all available modes for each model: $\ell \leqslant 4$ for NRSur7dq4, and $\ell = 2$ for SEOBNRv3. The NR error is computed using all $\ell \leqslant 5$ modes from the two highest available resolutions. Left panel: Mismatches computed using a flat noise curve. The square (triangle) markers at the top indicate the median (95th percentile) values. Right panel: Mismatches computed using the Advanced LIGO design sensitivity noise curve, as a function of total mass. The dashed (solid) lines indicate the median (95th percentile) values over different NR simulations and points in the sky.

more accurate NRSur7dq4 ($t = -4300M$ from the peak). We then use the dynamics surrogate of NRSur7dq4 to evolve the spins until $t = -100M$. These spins are then transformed to the co-orbital frame and used to evaluate the remnant fits. Thus, spins can be specified at any given orbital frequency and are evolved consistently before estimating the final BH properties. Note that NRSur7dq4 uses the filtered spins, while NRSur7dq4Remnant expects unfiltered spins at $t = -100M$, but we find that the errors introduced by this discrepancy are negligible compared to the errors due to PN spin evolution.

VI. RESULTS

We evaluate the accuracy of our surrogate models by comparing them against the waveform and remnant properties from the NR simulations used in this work. For this, we perform a 20-fold cross-validation study to compute "out-of-sample" errors as follows. We first randomly divide the 1528 training simulations into 20 groups of ≈ 76 simulations each. For each group, we build a trial surrogate using the ≈ 1452 remaining training simulations and test against these ≈ 76 validation ones, which may include points on the boundary of the training set.

A. Waveform surrogate errors

To estimate the difference between two waveforms, h_1 and h_2, we use the mismatch

$$\mathcal{MM} = 1 - \frac{\langle h_1, h_2 \rangle}{\sqrt{\langle h_1, h_1 \rangle \langle h_2, h_2 \rangle}}, \quad (12)$$

$$\langle h_1, h_2 \rangle = 4\mathrm{Re} \int_{f_{\min}}^{f_{\max}} \frac{\tilde{h}_1(f)\tilde{h}_2^*(f)}{S_n(f)} df, \quad (13)$$

where $\tilde{h}(f)$ indicates the Fourier transform of the complex strain $h(t)$, * indicates a complex conjugation, Re indicates

the real part, and $S_n(f)$ is the one-sided power spectral density of a GW detector. We taper the time domain waveform using a Planck window [99] and then zero pad to the nearest power of 2. We further zero pad the waveform to increase the length by a factor of 8 before performing the Fourier transform. 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 $30M$. Note that our model contains times up to $100M$ after the peak of the waveform amplitude, and the signal has essentially died down by the last $30M$. We take $f_{\min}$ to be twice the waveform angular velocity (as defined by Ref. [100]) at the end of the initial tapering window, and $f_{\max}$ is chosen to be four times the waveform angular velocity at $t = 0$; the extra factor of 4 is chosen to resolve up to $m = 4$ spherical-harmonic modes, with an extra margin of a factor of 2. We compute mismatches with a flat noise curve ($S_n = 1$) as well as with the advanced-LIGO design sensitivity noise curve [101]. Mismatches are computed following the procedure described in Appendix D of Ref. [74]. In particular, we optimize over shifts in time, polarization angle, and initial orbital phase. Both plus and cross polarizations are treated on an equal footing by using a two-detector setup where one detector sees only the plus and the other only the cross polarization. We compute the mismatches at 37 points uniformly distributed on the sky in the source frame, and we use all available modes of a given waveform model.

Figure 4 summarizes the out-of-sample mismatches for NRSur7dq4 against the NR waveforms. In Fig. 4(a), we show mismatches computed using a flat noise curve. We compare this with the truncation error in the NR waveforms themselves, estimated by computing the mismatch between the two highest available resolutions of each NR simulation. The errors in the surrogate model are well within the estimated truncation errors of the NR simulations. In addition, we also show the errors for the waveform model SEOBNRv3 [24,31],
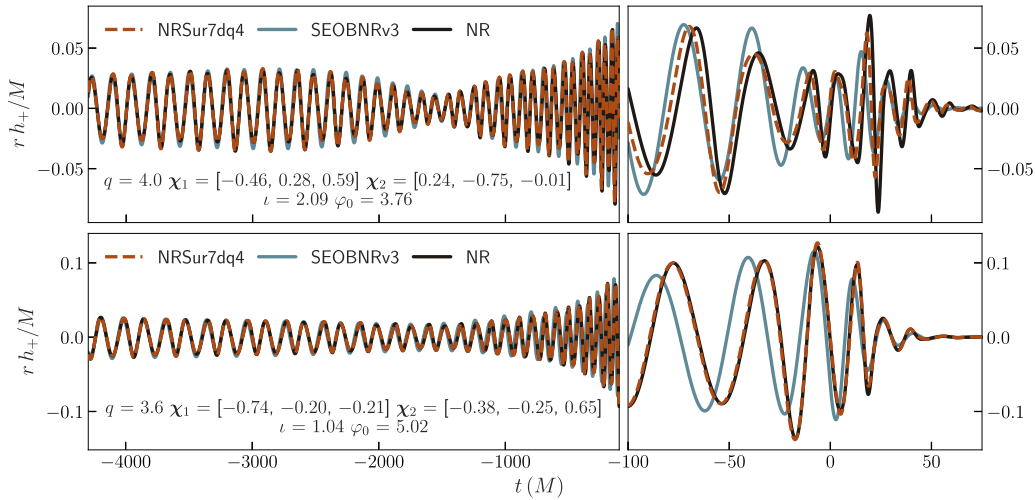
FIG. 5. The plus polarization of the waveforms for the cases that result in the largest mismatch for NRSur7dq4 (top) and SEOBNRv3 (bottom) in Fig. 4(a). We also show the corresponding NR waveforms. Each waveform is projected using all available modes for that model, along the direction that results in the largest mismatch for NRSur7dq4 (SEOBNRv3) in the top (bottom) panel. Note that NRSur7dq4 is evaluated using trial surrogates that are not trained using these cases. The binary parameters and the direction in the source frame are indicated in the figure text. All waveforms are time shifted such that the peak of the total amplitude occurs at $t = 0$ [using all available modes, according to Eq. (5)]. The waveform modes are then rotated to have their orbital angular momentum aligned with the z axis and such that the orbital phase is equal to zero at $t = -4300M$.

which also includes spin precession effects.[3] The surrogate errors are at least an order of magnitude lower than those of SEOBNRv3.

Apart from SEOBNRv3, another model commonly used in data analysis applications is IMRPhemomPv2 [27]. IMRPhemomPv2 was shown to be comparable in accuracy to SEOBNRv3 in Ref. [56], at least in order of magnitude. Therefore, for simplicity, we do not show comparisons of IMRPhemomPv2 to NR here. Note that updated versions of both SEOBNRv3 (based on Ref. [22]) and IMRPhemomPv2 (see Ref. [21]) are under development but are not currently available publicly. We note that these models are calibrated only against aligned-spin NR simulations, using a much smaller set of simulations than our model. Both these factors contribute to the accuracy of these models. On the other hand, these models are expected to be valid for larger mass ratios and spin magnitudes than our model, although their accuracy in that region is unknown due to lack of sufficient number of simulations.

We note that the NR truncation mismatch distribution in Fig. 4(a) has a tail extending to $\mathcal{MM} \sim 0.1$. We find that these cases occur when the spins of the two highest resolutions of the simulation are inconsistent with each other because of unresolved effects during junk-radiation emission, meaning that the two resolutions represent different physical systems. This means that comparing the resolutions for these cases gives us an error estimate that is too conservative and does not reflect the actual truncation error of the simulations. We

expect the actual truncation error to be closer to the errors reproduced by the surrogate model (which is trained on the high-resolution data set) in Fig. 4(a). Evidence for these claims is provided in Appendix B.

Figure 4(b) shows mismatches computed using the Advanced LIGO design sensitivity noise curve [101]. In this case, results depend on the total mass M of the system. Consequently, we show the median and 95th percentile values at different M, rather than full histograms. Once again, the surrogate errors are comparable to those of the NR simulations and are at least an order of magnitude lower than that of SEOBNRv3. Over the mass range 50–200$M_\odot$, mismatches for NRSur7dq4 are always $\lesssim 8 \times 10^{-3}$ at the 95 percentile level.

Figure 5 shows a comparison of waveforms computed via NRSur7dq4, SEOBNRv3, and NR for the cases that lead to the largest error for NRSur7dq4 and SEOBNRv3 in Fig. 4(a). The surrogate shows reasonable agreement with NR, even for its worst case, while SEOBNRv3 shows a noticeably larger deviation in both cases.

In Figs. 4 and 5, we use all available modes for NRSur7dq4 and SEOBNRv3. NRSur7dq4 models all modes $\ell \leqslant 4$, while SEOBNRv3 models only the $\ell = 2$ modes. For the NR waveforms in Figs. 4 and 5, we include all modes $\ell \leqslant 5$ to account for the error due to neglecting $\ell > 4$ modes in NRSur7dq4. To better understand what fraction of the SEOBNRv3 error comes from neglecting modes with $\ell > 2$, we repeat the calculations leading to the SEOBNRv3 histogram in Fig. 4(a) in Fig. 6, while restricting all waveforms to $\ell = 2$. While there is a noticeable move toward lower mismatches when restricted to $\ell = 2$, the median and 95th percentile values change only marginally, suggesting that the main error source for SEOBNRv3 are the $\ell = 2$ modes themselves.

---

[3]Note that SEOBNRv3 spins are specified at a reference frequency, rather than a time before merger. We choose the reference frequency such that the waveform begins at $t = -4300M$ before the waveform amplitude peak [as defined in Eq. (5)].
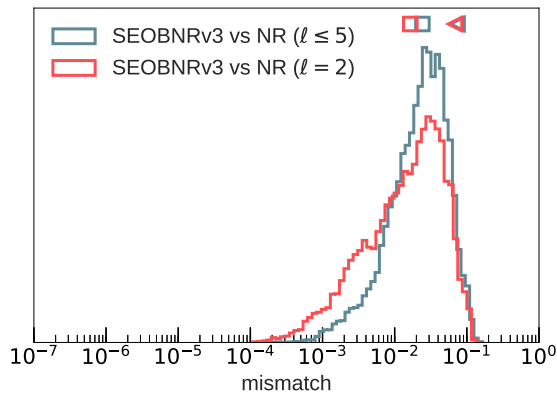
FIG. 6. Same as Fig. 4(a) but using only $\ell = 2$ modes for NR when compared to SEOBNRv3. The blue histogram from Fig. 4(a), where SEOBNRv3 is compared to NR with all $\ell \leqslant 5$ modes, is reproduced here for comparison. The square (triangle) markers at the top indicate the median (95th percentile) values.

B. Remnant surrogate errors

We evaluate the accuracy of the remnant surrogate NR-Sur7dq4Remnant by comparing against the NR simulations through a cross-validation study as in Sec. VI A. Out-of-sample errors for the remnant properties predicted by NR-Sur7dq4Remnant are shown in Fig. 7. The 95th percentile errors are $\approx 5 \times 10^{-4} M$ for mass, $\approx 2 \times 10^{-3}$ for spin magnitude, $\approx 4^{-3}$ radians for spin direction, $\approx 4 \times 10^{-4} c$ for kick magnitude, and ≈ 0.2 rad for kick direction. Our errors are at the same level as the NR resolution error, estimated by comparing the two highest NR resolutions. The largest errors in the kick direction can be of order ≈ 1 rad. The bottom-right panel of Fig. 7 shows the joint distribution of kick magnitude and kick direction error for NRSur7dq4Remnant, showing that direction errors are larger at low kick magnitudes. Our error in kick direction is below ≈ 0.2 rad whenever $v_f \gtrsim 2 \times 10^{-3} c$.

We also compare the performance of our fits against several existing fitting formulas for remnant mass, spin, and kick which we denote as follows: HBMR ([32,33] with $n_M = n_J = 3$), UIB [34], HL [35], HLZ [36], and CLZM ([37–41] as summarized in Ref. [42]). To partially account for spin precession, these fits are corrected as described in Ref. [102] and used in current LIGO-Virgo analyses [6,103]: Spins are evolved using PN from relaxation to the Schwarzschild innermost stable circular orbit, and final UIB and HL spins are postprocessed by adding the sum of the in-plane spins in quadrature. Figure 7 shows that our procedure to predict remnant mass, spin magnitude, and kick magnitude for precessing systems is more accurate than these existing fits by at least an order of magnitude.

Our fits appear to outperform the NR simulations when estimating the spin direction. Once again, this is due to the post-junk-radiation initial spins of the two highest resolutions being inconsistent with each other for some of our simulations, so that different resolutions represent different physical systems (cf. Appendix B). Therefore, the errors estimated by comparing the two highest resolutions is a poor estimate of the actual truncation error for these cases. The actual truncation

error is likely to be close to the errors reproduced by the surrogate.

The NRSur7dq4Remnant fits in Fig. 7 are evaluated using the NR spins at $t = -100M$ as inputs. In typical applications, one may have access to the spins only at the start of the waveform, rather than at $t = -100M$. For this case, as described in Sec. V, we use a combination of PN and NRSur7dq4 to evolve the spins from any given starting frequency to $t = -100M$. These spins are then used to evaluate the NRSur7dq4Remnant fits. Thus, spins can be specified at any given orbital frequency and are evolved consistently before estimating the final BH properties. This is a crucial improvement (introduced by Ref. [57]) over previous results, which, being calibrated solely to nonprecessing systems, suffer from ambiguities regarding the time and frequency at which spins are defined.

Figure 8 shows the errors in NRSur7dq4Remnant when the spins are specified at an orbital frequency $f_0 = 10$ Hz. These errors are computed by comparing against 23 long NR ($3 \times 10^4 M$ to $10^5 M$ in length) simulations [84] with mass ratios $q \leqslant 4$ and generically oriented spins with magnitudes $\chi_1, \chi_2 \sim 0.5$. None of these simulations were used to train the fits. Longer PN evolutions are needed at lower total masses, and the errors are therefore larger. These errors will decrease with an improved spin evolution procedure. Note, however, that our predictions are still more accurate than those of existing fitting formulas (cf. Fig. 7).

VII. CONCLUSION

We present NR surrogate models for precessing BBH systems with generic spins and unequal masses. In particular, we model the two most-used outputs of NR simulations: the gravitational waveform and the properties (mass, spin, and recoil kick) of the final BH formed after the merger. Trained against 1528 NR simulations with mass ratios $q \leqslant 4$, spin magnitudes $\chi_{1,2} \leqslant 0.8$, and generic spin directions, both these models are shown to reproduce the NR simulations with accuracies comparable to those of the simulations themselves.

The waveform model, NRSur7dq4, includes all spin-weighted spherical harmonic modes up to $\ell = 4$. The precession frame dynamics and spin evolution of the BHs are also modeled as by-products. Through a cross-validation study, we show that the mismatches for NRSur7dq4 against NR computed with the Advanced LIGO design sensitivity noise curve are always $\lesssim 8 \times 10^{-3}$ at the 95 percentile level over the mass range 50–200$M_\odot$. This is at least an order of magnitude improvement over existing waveform models. NRSur7dq4 is made publicly available through the gwsurrogate [58] PYTHON package, with example evaluation code at Ref. [60].

For the final BH model, NRSur7dq4Remnant, the 95th percentile errors are $\approx 5 \times 10^{-4} M$ for mass, $\approx 2 \times 10^{-3}$ for spin magnitude, and $\approx 4 \times 10^{-4} c$ for kick magnitude. Once again, these are lower than that of existing models by at least an order of magnitude. In addition, we also model the spin and kick directions. Moreover, the GPR methods employed here naturally provide error estimates along with the fitted values. These uncertainty estimates can be incorporated into data analysis applications to marginalize over systematic uncertainties. NRSur7dq4Remnant is made publicly available
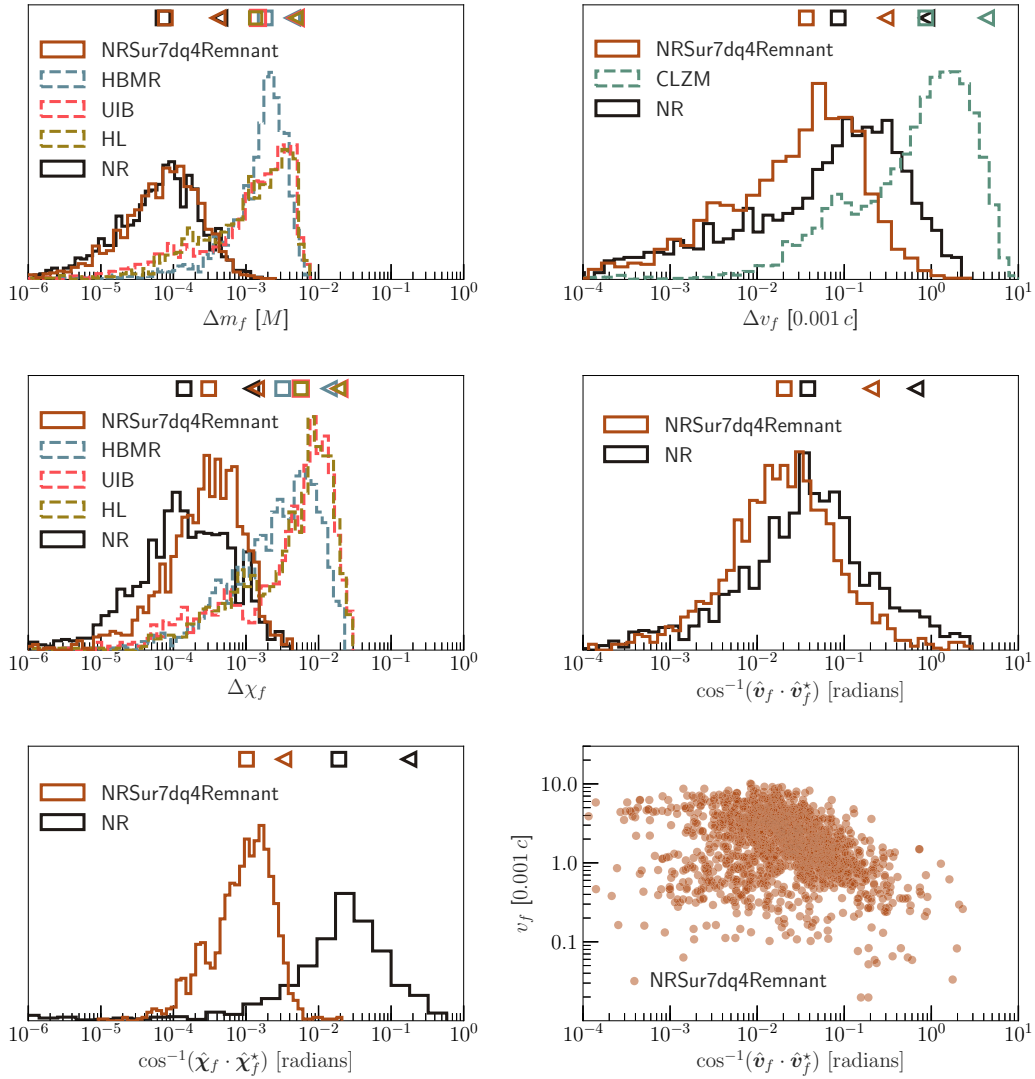
FIG. 7. Error histograms for NRSur7dq4Remnant for the remnant mass, spin magnitude, spin direction, kick magnitude, and kick direction for precessing BBH with mass ratios $q \leqslant 4$ and spin magnitudes $\chi_1, \chi_2 \leqslant 0.8$. The direction error is the angle between the predicted vector and a fiducial vector, taken to be the high-resolution NR case and indicated by *. Square (triangle) markers indicate median (95th percentile) values. Also shown are the NR resolution errors and errors for different existing fitting formulas. In the bottom-right panel, we show the joint distribution of kick magnitude and kick-direction error.

through the surfinBH [59] PYTHON package, which includes an example evaluation code.

A. Future work

In Appendix A, we test the performance of these surrogate models when extrapolated outside their training range to $q = 6$. We find that our models become worse at these mass ratios but are still comparable or better than existing models. Unfortunately, suitable precessing simulations are currently not available for testing at intermediate mass ratios $4 < q < 6$. In general, we advise caution with extrapolation. A natural improvement of both NRSur7dq4 and NRSur7dq4Remnant is to extend their range of validity with new training simulations at higher mass ratios and spin magnitudes. We note, however, that both these regimes are increasingly expensive to model in NR.

Another important limitation of these models is that they are restricted to the same length as the NR simulations (starting time of $\approx 4300M$ before the peak or about 20 orbits). For LIGO, assuming a starting GW frequency of 20 Hz, the (2, 2) mode of the surrogate is valid for total masses $M \gtrsim 66 M_\odot$. This number, however, depends on the mass ratio. Figure 9 shows the mass range of validity of NRSur7dq4 as a function of mass ratio. We compare this with the parameters of the 10 BBH detections seen by LIGO and Virgo in the first two observing runs [9]. NRSur7dq4 sufficiently covers the posterior spread of most but not all of these detections, the main limitation being the number of orbits covered by the model. However, see Ref. [104] for an example of NR surrogates used in data analysis with GW signals.

A promising avenue to extend the length of the waveforms is to "hybridize" the simulations using PN waveforms in the early inspiral. This approach already was found to be
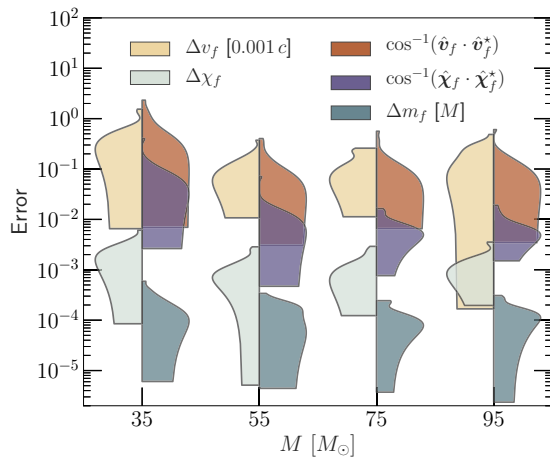
FIG. 8. Errors for NRSur7dq4Remnant in predicting remnant properties when spins are specified at an orbital frequency of $f_0 = 10$ Hz. For four different total masses, we compute the differences between the surrogate prediction of various remnant properties with the value obtained in the NR simulation. For each mass, these differences are shown as a vertical histogram. Note that the distributions in these plots are normalized to have a fixed height, not fixed area.

successful for the case of aligned-spin BBH [82] but still needs to be generalized to precessing spins. Furthermore, it is not clear if the current length of the NR simulations is sufficient to guarantee good attachment of the PN and NR waveforms for precessing BBH.

Despite these limitations, in their regime of validity, the models presented in the paper are the most accurate models currently available for precessing BBHs. As shown in this paper, our models rival the accuracy of the NR simulations, while being very cheap to evaluate. As more BBHs are detected at higher signal-to-noise ratios, fast yet accurate models
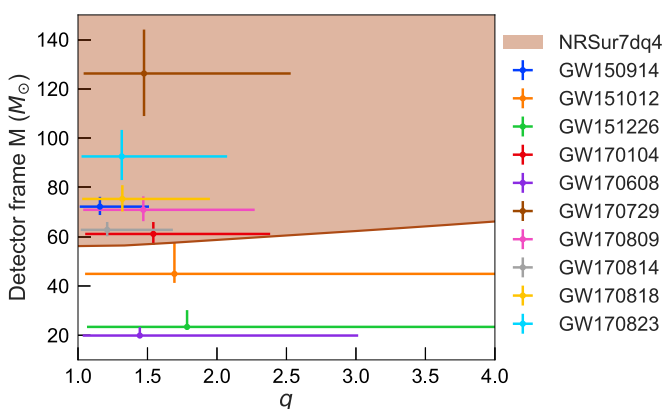
FIG. 9. The shaded region shows the regime of validity of the (2,2) mode of NRSur7dq4 with a starting frequency of 20 Hz. Also shown are the parameter ranges for the 10 BBH signals seen by LIGO and Virgo during the first two observing runs [9]. The markers indicate the median values of the marginalized posteriors for the detector frame total mass M and mass ratio q. The error bars indicate the range between the 5th percentile and 95th percentile values of the posteriors.

such as these will contribute to turning GW astronomy into high precision science.

APPENDIX A: EVALUATING SURROGATES AT LARGER MASS RATIOS

In this Appendix, we assess the performance of the NR-Sur7dq4 and NRSur7dq4Remnant models when evaluated at mass ratio $q = 6$. Doing so is effectively an extrapolation because $q = 6$ is outside the training range of the surrogates ($q \leqslant 4$). The surrogate models are compared against 100 NR simulations with $q = 6$ and generically precessing spins with magnitudes $\chi_1, \chi_2 \leqslant 0.8$. These simulations have been assigned the identifiers SXS:BBH:2164–SXS:BBH:2263 and are made publicly available through the SXS public catalog [83].

Figure 10 shows the $q = 6$ extrapolation mismatches for NRSur7dq4. Also shown are the mismatches for SEOB-NRv3 when compared against the same simulations. The mismatches are computed in the same manner as in Fig. 4(a), which we reproduce here for comparison. The surrogate errors become noticeably worse when extrapolating to $q = 6$, but are still much smaller than the corresponding errors for SEOBNRv3.

Figure 11 shows the performance of NRSur7dq4Remnant when extrapolating to $q = 6$. We show the errors when the fits are evaluated using the NR spins at $t = -100M$ as well as when the spins are specified at the start of the NR simulations. In the latter case, we use the extrapolated dynamics surrogate of NRSur7dq4 to evolve the spins to $t = -100M$
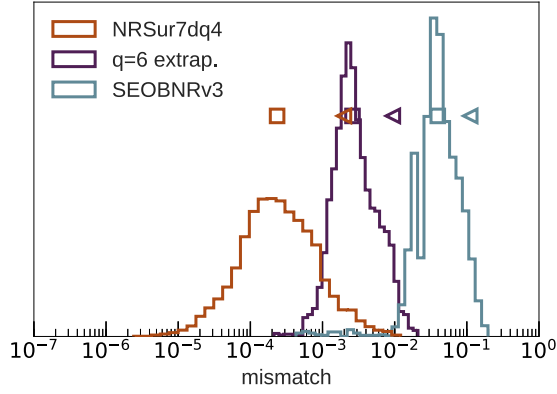
FIG. 10. Mismatch histogram when extrapolating the NR-Sur7dq4 waveform model to mass ratio $q = 6$. Also shown are mismatches for SEOBNRv3. The mismatches are computed using a flat noise curve. The training range errors from Fig. 4(a) are reproduced here for comparison. The square (triangle) markers indicate median (95th percentile) values.

and then evaluate the fits. We reproduce the training range errors from Fig. 7 for comparison. Also shown are the errors for the existing fitting formulas described in Sec. VI B when compared against the same simulations. We find that NRSur7dq4Remnant performs noticeably worse when extrapolated to $q = 6$ but is still slightly better than the existing fitting formulas, except for the final spin where the existing fitting formulas perform slightly better.

In general, we find that the NRSur7dq4 and NR-Sur7dq4Remnant models become worse with extrapolation to $q = 6$ but are still better or comparable to existing models. Unfortunately, we do not have enough suitable precessing simulations with $4 < q < 6$ with which to test at what mass ratio the degradation of these surrogate models becomes significant. We leave these tests, as well as extending the models to larger mass ratios by adding NR simulations, to future work.

APPENDIX B: ON THE HIGH MISMATCH TAIL IN NR ERRORS

The histogram of NR errors in Fig. 4(a) shows a significant tail to the right, i.e., at large mismatches. In Sec. VI A, this tail was attributed to different resolutions of the same NR simulation having different physical parameters, namely the "initial" spins, which are measured at the relaxation time [84] after the poorly resolved junk-radiation transients have settled. In this Appendix, we provide some evidence for this claim. Figure 12 shows the maximum mismatch (with a flat noise curve) over points in the sky versus the difference in the relaxation-time dimensionless spins between the two highest resolutions. We refer to the two highest resolutions as HiRes and MedRes, and their corresponding relaxation-time dimensionless spins are denoted by $(\chi_1^{\text{HiRes}}, \chi_2^{\text{HiRes}})$ and $(\chi_1^{\text{MedRes}}, \chi_2^{\text{MedRes}})$, respectively. We note that the largest mismatch occurs when the spin difference is largest between the two resolutions. For a significant fraction of the simulations, the spins can be different by about 0.1; for these cases the two resolutions essentially represent two different physical
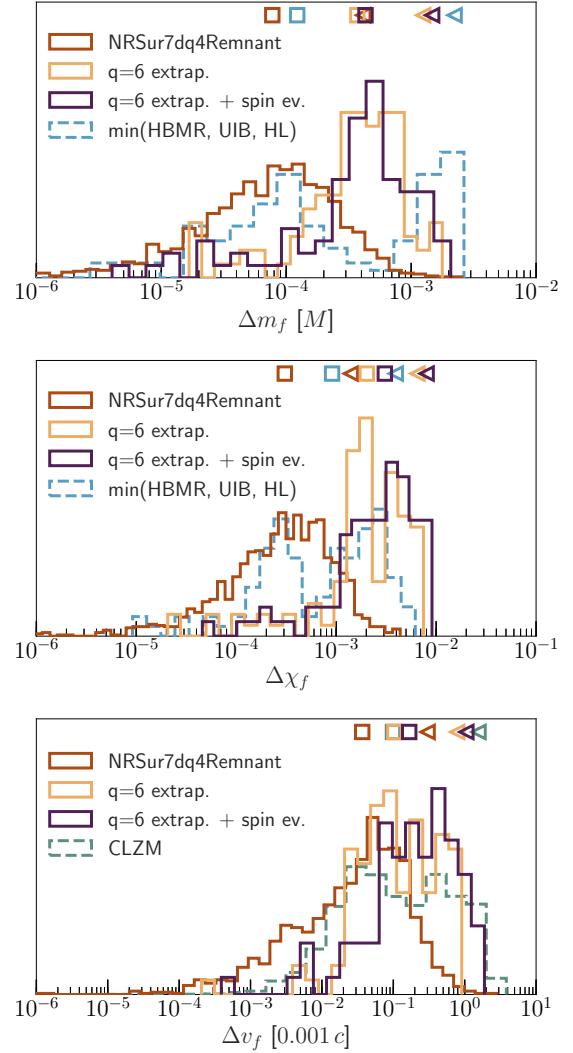

FIG. 11. Error histograms of the remnant mass, spin magnitude, and kick magnitude when extrapolating NRSur7dq4Remnant to mass ratio $q = 6$. The training range errors from Fig. 7 are reproduced here for comparison. We show errors using the NR spins at $t = -100M$ (yellow) as well as the initial NR spins (blue) as inputs for the model. Also shown are the errors for existing fitting formulas described in Sec. VI B; for the final mass and spin, we only show the minimum error among the HBMR, UIB, and HL fits. The square (triangle) markers indicate median (95th percentile) values.

systems, so the difference in waveforms between the two resolutions fails to be a good estimate of the truncation error in the simulations.

Figure 12 suggests that the high NR mismatch tail of Fig. 4(a) is artificially large, and if the two resolutions were to correspond to the same physical system, the tail would be shorter. We test this in Fig. 13, where we compare the surrogate against the MedRes simulations, but use the spins of the MedRes simulation $(\chi_1^{\text{MedRes}}, \chi_2^{\text{MedRes}})$ to evaluate the surrogate. The surrogate mismatches against the HiRes simulations as well as the NR resolution mismatches (HiRes vs MedRes) are reproduced from Fig. 4(a) for comparison. We note that the surrogate mismatches when compared against the MedRes simulations always lie below $\approx 10^{-2}$ and do not have
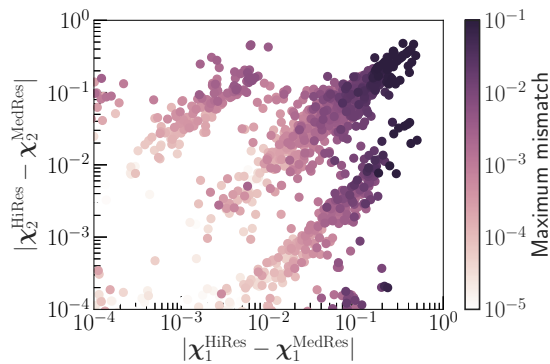
FIG. 12. Dependence of the NR resolution error on the difference in the relaxation-time spins of the two highest resolutions (labeled HiRes and MedRes). The horizontal (vertical) axis shows the difference between the spin of the heavier (lighter) BH. The colors show the largest (flat noise) mismatch between the waveforms of the two resolutions over different points in the sky. Large mismatches occur when the difference between the relaxation-time spins of the two resolutions is large.

the high mismatch tail seen for the NR resolution mismatches. In this test, we are treating the surrogate, which is trained on the HiRes simulations, as a proxy for the HiRes dataset. Evaluating the surrogate with the parameters of a MedRes simulation is treated as a proxy for performing the HiRes simulation with the same parameters. Therefore, the green histogram in Fig. 13 can be treated as the "true" resolution error when the parameters of the resolutions are the same. As expected for this case, this estimate of the resolution error agrees with the errors for the surrogate model (red histogram).

Together, Figs. 12 and 13 show that the high NR mismatch tail in Fig. 4(a) is due to the difference in the parameters of the different NR resolutions. We believe this difference arises from spurious initial transients known as "junk radiation."
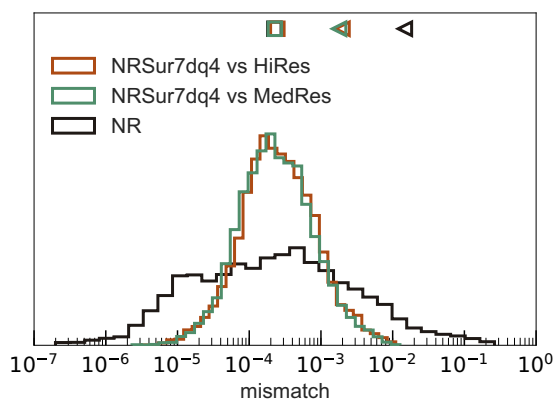

FIG. 13. Mismatch histograms for NRSur7dq4 when compared against the two highest available NR simulations (referred to as HiRes and MedRes). Also shown are mismatches between the two resolutions (labeled NR). The "NRSur7dq4 vs HiRes" and NR errors are the same as the red and black histograms, respectively, in Fig. 4(a). These are flat noise mismatches, computed at several points in the sky. The square (triangle) markers indicate median (95th percentile) values.
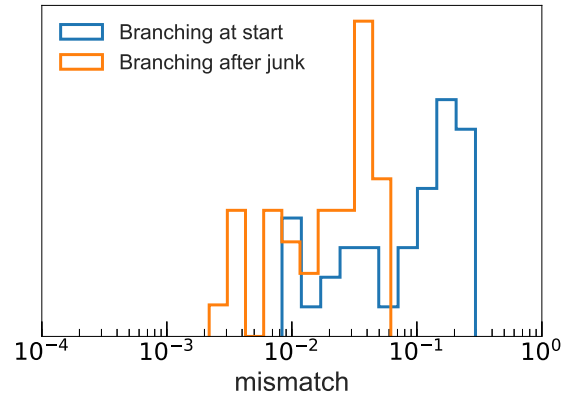

FIG. 14. NR resolution mismatches for the simulation leading to the largest NR mismatch in Fig. 4(a). The different samples in the histogram correspond to comparisons at different angles on the sky. The blue histogram shows the current resolution errors when the two resolutions start with the same initial data at the start of the simulation. All points in the blue histogram are the same as those included in Fig. 4(a). The green histogram shows the resolution errors for the same case when the two resolutions start with the same initial data at $\approx 1000M$ after start, at which point the junk radiation has left the simulation domain.

These transients result from initial data that do not precisely represent a snapshot of a binary that has evolved from $t = -\infty$. The transients quickly leave the simulation domain after about one or two binary orbits. It is computationally expensive to resolve the high spatial and temporal frequencies of the transients, so we typically choose not to resolve these transients at all, and instead we simply discard the initial part of the waveform. Because some of the transients carry energy and angular momentum down the BHs, the masses and spins are modified, so we measure "initial" masses and spins at a relaxation time [84] deemed sufficiently late that the transients have decayed away. Because we do not fully resolve the transients, their effect on the masses and spins are not always convergent with resolution.

This issue should ideally be resolved with improved, junk-free initial data (see Ref. [105] for steps in this direction). In the meantime, we propose a change in how SpEC performs different resolutions for the same simulation. Currently, initial data are constructed by solving the Einstein constraint equations [77,106]. The same constraint-satisfying initial data are then interpolated onto several grids of different resolution, and Einstein's equations are evolved on each grid independently. Our proposal is to first evolve the initial data using the high resolution grid until the transients leave the simulation domain, then interpolate the data at that time onto grids of lower resolution, and evolve Einstein's equations on these lower resolution grids independently. This way all resolutions start with the same initial data at a time after transients have decayed away instead of at the start of the simulation, and the masses and spins of the black holes should be convergent.

This proposal is tested in Fig. 14 for the case leading to the largest NR mismatch in Fig. 4(a). We perform the resolution branching at $t \sim 1000M$ after the start of the high-resolution simulation. The outer boundary is at $\approx 600M$ and this is

sufficient time for junk radiation to leave the simulation domain. We find that the mismatches decrease significantly

when the resolution branching is done postjunk, as the resolutions now correspond to the same physical system.

[1] J. Aasi *et al.* (LIGO Scientific Collaboration), Advanced LIGO, Class. Quantum Grav. **32**, 074001 (2015).

[2] F. Acernese *et al.* (Virgo Collaboration), Advanced Virgo: A second-generation interferometric gravitational wave detector, Class. Quantum Grav. **32**, 024001 (2015).

[3] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Observation of Gravitational Waves from a Binary Black Hole Merger, Phys. Rev. Lett. **116**, 061102 (2016).

[4] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral, Phys. Rev. Lett. **119**, 161101 (2017).

[5] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence, Phys. Rev. Lett. **116**, 241103 (2016).

[6] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2, Phys. Rev. Lett. **118**, 221101 (2017); **121**, 129901(E) (2018).

[7] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW170608: Observation of a 19-solar-mass binary black hole coalescence, Astrophys. J. **851**, L35 (2017).

[8] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence, Phys. Rev. Lett. **119**, 141101 (2017).

[9] 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).

[10] B. P. Abbott *et al.* (KAGRA, LIGO Scientific, and VIRGO Collaborations), Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo, and KAGRA, Living Rev. Rel. **21**, 3 (2018).

[11] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Binary black hole population properties inferred from the first and second observing runs of Advanced LIGO and Advanced Virgo, arXiv:1811.12940 [astro-ph.HE].

[12] C. Cutler and E. E. Flanagan, Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral wave form? Phys. Rev. D **49**, 2658 (1994).

[13] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Properties of the Binary Black Hole Merger GW150914, Phys. Rev. Lett. **116**, 241102 (2016).

[14] J. Veitch, V. Raymond, B. Farr, W. Farr, P. Graff, S. Vitale, B. Aylott, K. Blackburn, N. Christensen, M. Coughlin *et al.*, Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library, Phys. Rev. D **91**, 042003 (2015).

[15] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Tests of General Relativity with GW150914, Phys. Rev. Lett. **116**, 221101 (2016); **121** 129902(E) (2018).

[16] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaboration), Tests of general relativity with the binary black hole signals from the LIGO-Virgo Catalog GWTC-1, arXiv:1903.04467 [gr-qc].

[17] A. Ghosh, N. K. Johnson-Mcdaniel, A. Ghosh, C. K. Mishra, P. Ajith, W. Del Pozzo, C. P. L. Berry, A. B. Nielsen, and L. London, Testing general relativity using gravitational wave signals from the inspiral, merger, and ringdown of binary black holes, Class. Quantum Grav. **35**, 014002 (2018).

[18] M. Campanelli, C. O. Lousto, Y. Zlochower, and D. Merritt, Maximum Gravitational Recoil, Phys. Rev. Lett. **98**, 231102 (2007).

[19] J. A. Gonzalez, M. Hannam, U. Sperhake, B. Bruegmann, and S. Husa, Supermassive Recoil Velocities for Binary Black-Hole Mergers with Antialigned Spins, Phys. Rev. Lett. **98**, 231101 (2007).

[20] D. Gerosa and A. Sesana, Missing black holes in brightest cluster galaxies as evidence for the occurrence of superkicks in nature, Mon. Not. Roy. Astron. Soc. **446**, 38 (2015).

[21] S. Khan, K. Chatziioannou, M. Hannam, and F. Ohme, Phenomenological model for the gravitational-wave signal from precessing binary black holes with two-spin effects, Phys. Rev. D **100**, 024059 (2019).

[22] R. Cotesta, A. Buonanno, A. Bohé, A. Taracchini, I. Hinder, and S. Ossokine, Enriching the symphony of gravitational waves from binary black holes by tuning higher harmonics, Phys. Rev. D **98**, 084028 (2018).

[23] L. London, S. Khan, E. Fauchon-Jones, C. García, M. Hannam, S. Husa, X. Jiménez-Forteza, C. Kalaghatgi, F. Ohme, and F. Pannarale, First Higher-Multipole Model of Gravitational Waves from Spinning and Coalescing Black-Hole Binaries, Phys. Rev. Lett. **120**, 161102 (2018).

[24] Y. Pan, A. Buonanno, A. Taracchini, L. E. Kidder, A. H. Mroué, H. P. Pfeiffer, M. A. Scheel, and B. Szilágyi, Inspiral-merger-ringdown waveforms of spinning, precessing black-hole binaries in the effective-one-body formalism, Phys. Rev. D **89**, 084006 (2014).

[25] A. 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).

[26] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jiménez Forteza, and A. 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).

[27] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, Simple Model of

Complete Precessing Black-Hole-Binary Gravitational Wave-forms, Phys. Rev. Lett. **113**, 151101 (2014).

[28] A. Taracchini, A. Buonanno, Y. Pan, T. Hinderer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, A. H. Mroue, H. P. Pfeiffer *et al.*, Effective-one-body model for black-hole binaries with generic mass ratios and spins, Phys. Rev. D **89**, 061502(R) (2014).

[29] Y. Pan, A. Buonanno, M. Boyle, L. T. Buchman, L. E. Kidder, H. P. Pfeiffer, and M. A. Scheel, Inspiral-merger-ringdown multipolar waveforms of nonspinning black-hole binaries using the effective-one-body formalism, Phys. Rev. D **84**, 124052 (2011).

[30] A. K. Mehta, C. K. Mishra, V. Varma, and P. Ajith, Accurate inspiral-merger-ringdown gravitational waveforms for nonspinning black-hole binaries including the effect of subdominant modes, Phys. Rev. D **96**, 124010 (2017).

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

[32] F. Hofmann, E. Barausse, and L. Rezzolla, The final spin from binary black holes in quasi-circular orbits, Astrophys. J. **825**, L19 (2016).

[33] E. Barausse, V. Morozova, and L. Rezzolla, On the mass radiated by coalescing black-hole binaries, Astrophys. J. **758**, 63 (2012); **786**, 76(E) (2014).

[34] X. Jiménez-Forteza, D. Keitel, S. Husa, M. Hannam, S. Khan, and M. Pürrer, Hierarchical data-driven approach to fitting numerical relativity data for nonprecessing binary black holes with an application to final spin and radiated energy, Phys. Rev. D **95**, 064024 (2017).

[35] J. Healy and C. O. Lousto, Remnant of binary black-hole mergers: New simulations and peak luminosity studies, Phys. Rev. D **95**, 024037 (2017).

[36] J. Healy, C. O. Lousto, and Y. Zlochower, Remnant mass, spin, and recoil from spin aligned black-hole binaries, Phys. Rev. D **90**, 104004 (2014).

[37] J. A. Gonzalez, U. Sperhake, B. Brügmann, M. Hannam, and S. Husa, Total Recoil: The Maximum Kick from Nonspinning Black-Hole Binary Inspiral, Phys. Rev. Lett. **98**, 091101 (2007).

[38] M. Campanelli, C. O. Lousto, Y. Zlochower, and D. Merritt, Large merger recoils and spin flips from generic black-hole binaries, Astrophys. J. **659**, L5 (2007).

[39] C. O. Lousto and Y. Zlochower, Further insight into gravitational recoil, Phys. Rev. D **77**, 044028 (2008).

[40] C. O. Lousto, Y. Zlochower, M. Dotti, and M. Volonteri, Gravitational recoil from accretion-aligned black-hole binaries, Phys. Rev. D **85**, 084015 (2012).

[41] C. O. Lousto and Y. Zlochower, Nonlinear gravitational recoil from the mergers of precessing black-hole binaries, Phys. Rev. D **87**, 084027 (2013).

[42] D. Gerosa and M. Kesden, PRECESSION: Dynamics of spinning black-hole binaries with PYTHON, Phys. Rev. D **93**, 124066 (2016).

[43] J. Healy and C. O. Lousto, Hangup effect in unequal mass binary black hole mergers and further studies of their gravitational radiation and remnant properties, Phys. Rev. D **97**, 084002 (2018).

[44] F. Herrmann, I. Hinder, D. M. Shoemaker, P. Laguna, and R. A. Matzner, Binary black holes: Spin dynamics and gravitational recoil, Phys. Rev. D **76**, 084032 (2007).

[45] L. Rezzolla, E. Barausse, E. N. Dorband, D. Pollney, C. Reisswig, J. Seiler, and S. Husa, On the final spin from the coalescence of two black holes, Phys. Rev. D **78**, 044002 (2008).

[46] L. Rezzolla, P. Diener, E. N. Dorband, D. Pollney, C. Reisswig, E. Schnetter, and J. Seiler, The final spin from the coalescence of aligned-spin black-hole binaries, Astrophys. J. **674**, L29 (2008).

[47] M. Kesden, Can binary mergers produce maximally spinning black holes? Phys. Rev. D **78**, 084030 (2008).

[48] W. Tichy and P. Marronetti, The final mass and spin of black hole mergers, Phys. Rev. D **78**, 081501(R) (2008).

[49] E. Barausse and L. Rezzolla, Predicting the direction of the final spin from the coalescence of two black holes, Astrophys. J. **704**, L40 (2009).

[50] Y. Zlochower and C. O. Lousto, Modeling the remnant mass, spin, and recoil from unequal-mass, precessing black-hole binaries: The intermediate mass ratio regime, Phys. Rev. D **92**, 024022 (2015); **94**, 029901(E) (2016).

[51] T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, Spin-induced orbital precession and its modulation of the gravitational waveforms from merging binaries, Phys. Rev. D **49**, 6274 (1994).

[52] D. Gerosa, M. Kesden, E. Berti, R. O'Shaughnessy, and U. Sperhake, Resonant-plane locking and spin alignment in stellar-mass black-hole binaries: A diagnostic of compact-binary formation, Phys. Rev. D **87**, 104028 (2013).

[53] S. Vitale, R. Lynch, R. Sturani, and P. Graff, Use of gravitational waves to probe the formation channels of compact binaries, Class. Quantum Grav. **34**, 03LT01 (2017).

[54] B. Farr, D. E. Holz, and W. M. Farr, Using spin to understand the formation of LIGO and Virgo's black holes, Astrophys. J. **854**, L9 (2018).

[55] D. Gerosa, E. Berti, R. O'Shaughnessy, K. Belczynski, M. Kesden, D. Wysocki, and W. Gladysz, Spin orientations of merging black holes formed from the evolution of stellar binaries, Phys. Rev. D **98**, 084036 (2018).

[56] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, C. D. Ott, M. Boyle, L. E. Kidder, H. P. Pfeiffer, and B. Szilágyi, Numerical relativity waveform surrogate model for generically precessing binary black hole mergers, Phys. Rev. D **96**, 024058 (2017).

[57] V. Varma, D. Gerosa, L. C. Stein, F. Hébert, and H. Zhang, High-Accuracy Mass, Spin, and Recoil Predictions of Generic Black-Hole Merger Remnants, Phys. Rev. Lett. **122**, 011101 (2019).

[58] J. Blackman, S. Field, C. Galley, and V. Varma, gwsurrogate, https://pypi.python.org/pypi/gwsurrogate/.

[59] V. Varma *et al.*, surfinBH, pypi.org/project/surfinBH, doi: 10.5281/zenodo.1418525.

[60] Binary black-hole surrogate waveform catalog, http://www.black-holes.org/surrogates/.

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

[62] C. Capano, Y. Pan, and A. Buonanno, Impact of higher harmonics in searching for gravitational waves from nonspinning binary black holes, Phys. Rev. D **89**, 102003 (2014).

[63] T. B. Littenberg, J. G. Baker, A. Buonanno, and B. J. Kelly, Systematic biases in parameter estimation of binary black-hole mergers, Phys. Rev. D **87**, 104003 (2013).

[64] J. Calderón Bustillo, P. Laguna, and D. Shoemaker, Detectability of gravitational waves from binary black holes: Impact of precession and higher modes, Phys. Rev. D **95**, 104038 (2017).

[65] D. A. Brown, P. Kumar, and A. H. Nitz, Template banks to search for low-mass binary black holes in advanced gravitational-wave detectors, Phys. Rev. D **87**, 082004 (2013).

[66] V. Varma, P. Ajith, S. Husa, J. C. Bustillo, M. Hannam, and M. Pürrer, Gravitational-wave observations of binary black holes: Effect of nonquadrupole modes, Phys. Rev. D **90**, 124004 (2014).

[67] P. B. Graff, A. Buonanno, and B. S. Sathyaprakash, Missing link: Bayesian detection and measurement of intermediate-mass black-hole binaries, Phys. Rev. D **92**, 022002 (2015).

[68] I. Harry, J. Calderón Bustillo, and A. Nitz, Searching for the full symphony of black hole binary mergers, Phys. Rev. D **97**, 023004 (2018).

[69] J. Calderón Bustillo, S. Husa, A. M. Sintes, and M. Pürrer, Impact of gravitational radiation higher order modes on single aligned-spin gravitational wave searches for binary black holes, Phys. Rev. D **93**, 084019 (2016).

[70] L. Pekowsky, J. Healy, D. Shoemaker, and P. Laguna, Impact of higher-order modes on the detection of binary black hole coalescences, Phys. Rev. D **87**, 084008 (2013).

[71] P. Schmidt, M. Hannam, S. Husa, and P. Ajith, Tracking the precession of compact binaries from their gravitational-wave signal, Phys. Rev. D **84**, 024046 (2011).

[72] R. O'Shaughnessy, B. Vaishnav, J. Healy, Z. Meeks, and D. Shoemaker, Efficient asymptotic frame selection for binary black hole spacetimes using asymptotic radiation, Phys. Rev. D **84**, 124002 (2011).

[73] M. Boyle, R. Owen, and H. P. Pfeiffer, A geometric approach to the precession of compact binaries, Phys. Rev. D **84**, 124011 (2011).

[74] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, D. A. Hemberger, P. Schmidt, and R. Smith, A surrogate model of gravitational waveforms from numerical relativity simulations of precessing binary black hole mergers, Phys. Rev. D **95**, 104023 (2017).

[75] The spectral Einstein code, http://www.black-holes.org/SpEC.html.

[76] H. P. Pfeiffer, L. E. Kidder, M. A. Scheel, and S. A. Teukolsky, A multidomain spectral method for solving elliptic equations, Comput. Phys. Commun. **152**, 253 (2003).

[77] G. Lovelace, R. Owen, H. P. Pfeiffer, and T. Chu, Binary-black-hole initial data with nearly-extremal spins, Phys. Rev. D **78**, 084017 (2008).

[78] L. Lindblom, M. A. Scheel, L. E. Kidder, R. Owen, and O. Rinne, A new generalized harmonic evolution system, Class. Quantum Grav. **23**, S447 (2006).

[79] B. Szilagyi, L. Lindblom, and M. A. Scheel, Simulations of binary black hole mergers using spectral methods, Phys. Rev. D **80**, 124010 (2009).

[80] M. A. Scheel, M. Boyle, T. Chu, L. E. Kidder, K. D. Matthews, and H. P. Pfeiffer, High-accuracy waveforms for binary black hole inspiral, merger, and ringdown, Phys. Rev. D **79**, 024003 (2009).

[81] Simulating eXtreme spacetimes, http://www.black-holes.org/.

[82] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, L. E. Kidder, and H. P. Pfeiffer, Surrogate model of hybridized numerical relativity binary black hole waveforms, Phys. Rev. D **99**, 064045 (2019).

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

[84] M. Boyle *et al.*, The SXS Collaboration catalog of binary black hole simulations, Class. Quantum Grav. **36**, 195006 (2019).

[85] A. Buonanno, L. E. Kidder, A. H. Mroue, H. P. Pfeiffer, and A. Taracchini, Reducing orbital eccentricity of precessing black-hole binaries, Phys. Rev. D **83**, 104034 (2011).

[86] M. Boyle and A. H. Mroue, Extrapolating gravitational-wave data from numerical simulations, Phys. Rev. D **80**, 124045 (2009).

[87] M. Boyle, Transformations of asymptotic gravitational-wave data, Phys. Rev. D **93**, 084031 (2016).

[88] M. Boyle, Scri, https://github.com/moble/scri.

[89] S. Ossokine, M. Boyle, L. E. Kidder, H. P. Pfeiffer, M. A. Scheel, and B. Szilágyi, Comparing post-Newtonian and numerical-relativity precession dynamics, Phys. Rev. D **92**, 104028 (2015).

[90] D. Gerosa, F. Hébert, and L. C. Stein, Black-hole kicks from numerical-relativity surrogate models, Phys. Rev. D **97**, 104049 (2018).

[91] M. Ruiz, R. Takahashi, M. Alcubierre, and D. Nunez, Multipole expansions for energy and momenta carried by gravitational waves, Gen. Rel. Grav. **40**, 2467 (2008).

[92] S. E. Field, C. R. Galley, J. S. Hesthaven, J. Kaye, and M. Tiglio, Fast Prediction and Evaluation of Gravitational Waveforms Using Surrogate Models, Phys. Rev. X **4**, 031006 (2014).

[93] 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).

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

[95] P. Ajith, Addressing the spin question in gravitational-wave searches: Waveform templates for inspiralling compact binaries with nonprecessing spins, Phys. Rev. D **84**, 084037 (2011).

[96] E. Poisson and C. M. Will, Gravitational waves from inspiraling compact binaries: Parameter estimation using second post-Newtonian wave forms, Phys. Rev. D **52**, 848 (1995).

[97] A. Buonanno, Y.-b. Chen, and M. Vallisneri, Detecting gravitational waves from precessing binaries of spinning compact objects: Adiabatic limit, Phys. Rev. D **67**, 104025 (2003); **74**, 029904(E) (2006).

[98] M. Boyle, D. A. Brown, L. E. Kidder, A. H. Mroue, H. P. Pfeiffer, M. A. Scheel, G. B. Cook, and S. A. Teukolsky, High-accuracy comparison of numerical relativity simulations with post-Newtonian expansions, Phys. Rev. D **76**, 124038 (2007).

[99] 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, Class. Quantum Grav. **27**, 084020 (2010).

[100] M. Boyle, Angular velocity of gravitational radiation from precessing binaries and the corotating frame, Phys. Rev. D **87**, 104006 (2013).

[101] LIGO Scientific Collaboration, Updated Advanced LIGO sensitivity design curve, Tech. Rep., 2018, https://dcc.ligo.org/LIGO-T1800044/public.

[102] N. K. Johnson-McDaniel, A. Gupta, P Ajith, D. Keitel, O. Birnholtz, F. Ohme, and S. Husa, Determining the final spin of a binary black hole system including in-plane spins: Method and checks of accuracy, Tech. Rep. LIGO-T1600168, 2016, https://dcc.ligo.org/LIGO-T1600168/public.

[103] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaboration), Binary Black Hole Mergers in the First Advanced LIGO Observing Run, Phys. Rev. X **6**, 041015 (2016); **8**, 039903(E) (2018).

[104] P. Kumar, J. Blackman, S. E. Field, M. Scheel, C. R. Galley, M. Boyle, L. E. Kidder, H. P. Pfeiffer, B. Szilagyi, and S. A. Teukolsky, Constraining the parameters of GW150914 and GW170104 with numerical relativity surrogates, Phys. Rev. D **99**, 124005 (2019).

[105] V. Varma, M. A. Scheel, and H. P. Pfeiffer, Comparison of binary black hole initial data sets, Phys. Rev. D **98**, 104011 (2018).

[106] S. Ossokine, F. Foucart, H. P. Pfeiffer, M. Boyle, and B. Szilágyi, Improvements to the construction of binary black hole initial data, Class. Quantum Grav. **32**, 245010 (2015).