PHYSICAL REVIEW D **108,** 064027 (2023)

Numerical relativity surrogate model with memory effects and post-Newtonian hybridization

Jooheon Yoo,[1] Keefe Mitman,[2] Vijay Varma,[3] Michael Boyle,[1] Scott E. Field,[4] Nils Deppe,[2] François Hébert,[2]
Lawrence E. Kidder,[1] Jordan Moxon,[2] Harald P. Pfeiffer,[3] Mark A. Scheel,[2] Leo C. Stein,[5]
Saul A. Teukolsky,[1,2] William Throwe,[1] and Nils L. Vu[2,3]

[1]*Cornell Center for Astrophysics and Planetary Science, Cornell University, Ithaca, New York 14853, USA*
[2]*Theoretical Astrophysics 350-17, California Institute of Technology, Pasadena, California 91125, USA*
[3]*Max Planck Institute for Gravitational Physics (Albert Einstein Institute),*
Am Mühlenberg 1, D-14476 Potsdam, Germany
[4]*Department of Mathematics, Center for Scientific Computing and Data Science Research,*
University of Massachusetts, Dartmouth, Massachusetts 02747, USA
[5]*Department of Physics and Astronomy, University of Mississippi, University, Mississippi 38677, USA*

(Received 7 June 2023; accepted 15 August 2023; published 14 September 2023)

Numerical relativity simulations provide the most precise templates for the gravitational waves produced by binary black hole mergers. However, many of these simulations use an incomplete waveform extraction technique—extrapolation—that fails to capture important physics, such as gravitational memory effects. Cauchy-characteristic evolution (CCE), by contrast, is a much more physically accurate extraction procedure that fully evolves Einstein's equations to future null infinity and accurately captures the expected physics. In this work, we present a new surrogate model, `NRHybSur3dq8_CCE`, built from CCE waveforms that have been mapped to the post-Newtonian (PN) BMS frame and then hybridized with PN and effective one-body (EOB) waveforms. This model is trained on 102 waveforms with mass ratios $q \leq 8$ and aligned spins $\chi_{1z}, \chi_{2z} \in [-0.8, 0.8]$. The model spans the entire LIGO-Virgo-KAGRA (LVK) frequency band (with $f_{\mathrm{low}} = 20$ Hz) for total masses $M \gtrsim 2.25 M_{\odot}$ and includes the $\ell \leq 4$ and $(\ell, m) = (5, 5)$ spin-weight -2 spherical harmonic modes, but not the (3, 1), (4, 2) or (4, 1) modes. We find that `NRHybSur3dq8_CCE` can accurately reproduce the training waveforms with mismatches $\lesssim 2 \times 10^{-4}$ for total masses $2.25 M_{\odot} \leq M \leq 300 M_{\odot}$ and can, for a modest degree of extrapolation, capably model outside of its training region. Most importantly, unlike previous waveform models, the new surrogate model successfully captures memory effects.

DOI: 10.1103/PhysRevD.108.064027

I. INTRODUCTION

To date, there have been a total of 90 joint detections of gravitational wave (GW) signals by the LIGO[1] [1] and Virgo [2] collaborations. But, with increased sensitivity in future observation runs and the inclusion of KAGRA[2] [3] as well as other proposed future detectors, such as the Einstein Telescope [4], the Cosmic Explorer [5], and the space-based LISA[3] [6], the number of gravitational wave observations is expected to increase dramatically [7,8]. To fully take advantage of the ever-expanding catalog of gravitational wave signals from compact binaries, it is crucial that we have high-fidelity waveform templates to compare the observed signals to. This is because accurate waveform templates are necessary for reliably extracting astrophysical source properties that provide important information about the binaries' formation channels and also for performing unique tests of general relativity.

Numerical relativity (NR) is the only *ab initio* method for solving Einstein's equations for the coalescence of two compact objects and has played a fundamental role in both GW theory and GW astronomy [9–12]. Even so, despite continued efforts by the NR community to make simulations more computationally efficient, they are still prohibitively expensive for key multi-query applications, such as parameter estimation. Because of this bottleneck, numerous waveform models have been developed [13–18] that can be evaluated much faster than evolving an entire NR simulation. By construction, these semi-analytical models rely on physically motivated or phenomenological assumptions to reduce the complexity of parameter space. They then calibrate the remaining free parameters by comparing to the waveforms produced by NR simulations. While these waveform models tend to be

---

[1] The Laser Interferometer Gravitational-Wave Observatory.
[2] The Kamioka Gravitational Wave Detector.
[3] The Laser Interferometer Space Antenna.

fast enough for performing GW data analyses, they are not nearly as accurate or reliable as NR waveforms.

NR surrogate models are a more recent addition to the collection of compact binary waveform models [19–25]. Unlike semianalytic models, NR surrogates instead take a data-driven approach: training the waveform model directly on the waveforms output by simulations without the need to make any assumptions about the physics. Consequently, these surrogates recover NR waveforms much more faithfully than other semianalytical models. However, because of this unique data-driven approach to waveform modeling, NR surrogate models can only be constructed for the regions of parameter space in which NR simulations exist.[4]

Both NR surrogates and other semianalytical models have played a crucial role in studying previous detections of gravitational waves [28–35]. Up until now, however, every waveform model has always either been trained or calibrated by working with ‘extrapolated’ waveforms. Because NR simulations of binaries are necessarily run in a finite volume, one needs a method of ‘extracting’ the asymptotic waveform from the finite-volume data that is produced by the simulation. This is because the gravitational radiation that we observe on Earth can be well-approximated by the radiative solutions to Einstein’s equations at future null infinity. In NR, an ‘extrapolated’ waveform refers to the asymptotic waveform from an extraction procedure called extrapolation.[5] However, a major limitation of these extrapolated waveforms, which are the waveforms currently included in the SXS [12,37] and other catalogs [39,40], is that they do not accurately capture phenomena known as memory effects [41–44].

Gravitational memory effects correspond to persistent phenomena that two observers can measure after the passage of gravitational radiation [41–44]. In particular, to measure memory effects, observers must measure the change in spacetime induced by the passage of radiation between two times: one before and one after the radiation. While there are several types of memory effects [45,46], the two most prominent, and therefore detectable, effects are the displacement [41–44] and the spin [47] memories. The displacement memory is what two initially comoving observers will measure, while the spin memory is what two observers with initial relative velocities will measure, in conjunction with the usual displacement that they experience due to their nonzero relative velocities [46].

Apart from this classification of memory by the changes experienced by observers, there is also a classification in the way in which memory is sourced: ordinary and null.[6] Ordinary memory refers to the memory that is sourced by changes in the $\ell \geq 2$ mass multipole moment of “ordinary” unbound masses, while null memory refers to the memory that is sourced by a change in the energy radiated per unit solid angle due to the radiation of “null” gravitons. Consequently, null memory can be thought of as a form of the ordinary memory in which the unbound masses are individual gravitons. Generally, ordinary memory will be most prominent in unbound scattering processes, such as hyperbolic black hole encounters [49–54]. In contrast, null memory will be most prominent in bound scattering processes, such as binary black hole mergers. Because of this, in this work we will primarily be interested in the null memory.

While memory effects are an undetected phenomenon, various works have investigated their detectability using a forecast of future binary merger observations [55–57] and their observational consequences [58–60]. Furthermore, apart from their allure as a means to test Einstein’s theory of relativity, memory effects have also attracted significant attention in the theory community because of their inherent connection to asymptotic symmetries and soft theorems [45,47,61–65]. Accordingly, it is crucial that templates for gravitational waves contain memory.

While extrapolated NR waveforms ultimately fail to correctly capture memory effects,[7,8] there is a much more robust type of asymptotic waveform extraction, called Cauchy-characteristic evolution (CCE), which fully evolves Einstein’s equations to future null infinity and correctly resolves the various memory effects [68–71].

In this work, we build a hybridized NR surrogate model, `NRHybSur3dq8_CCE`, which is trained on CCE waveforms that have been created using the CCE module of the code SPECTRE [69,70,72]. Furthermore, to build our hybrid waveforms, using the technique outlined in [73], we first map our NR waveforms to the post-Newtonian (PN) BMS frame [73,74] before we hybridize them with effective one-body (EOB) phase-corrected PN waveforms. By doing so, we find that `NRHybSur3dq8_CCE` performs on par with a related extrapolated version of this surrogate,

---

[4] Surrogate models have also been constructed for semianalytical EOB models [26,27].

[5] Extrapolation uses Regge-Wheeler-Zerilli extraction to compute the strain waveform on a series of concentric spheres of constant coordinate radius and then extrapolates these values to future null infinity by fitting a power series in $1/r$ [36–38].

[6] Originally this classification was linear and nonlinear [41–43], but this terminology changed in recent years to more accurately reflect the physics sourcing this phenomenon [48].

[7] The reason why memory effects are not correctly resolved in extrapolated waveforms is because memory effects tend to have a much longer spatial dependence than the oscillatory components of the waveform. Consequently, when expressing the waveform as a series of $1/r$ terms, the convergence of the memory’s contribution to the waveform is very slow and hard to capture [36–38].

[8] While some of the missing memory in extrapolated waveforms can be computed through postprocessing [66,67], there are certain types of memory effects that cannot be corrected, e.g., spin memory, which makes extrapolated waveforms impractical for formal analyses of memory.

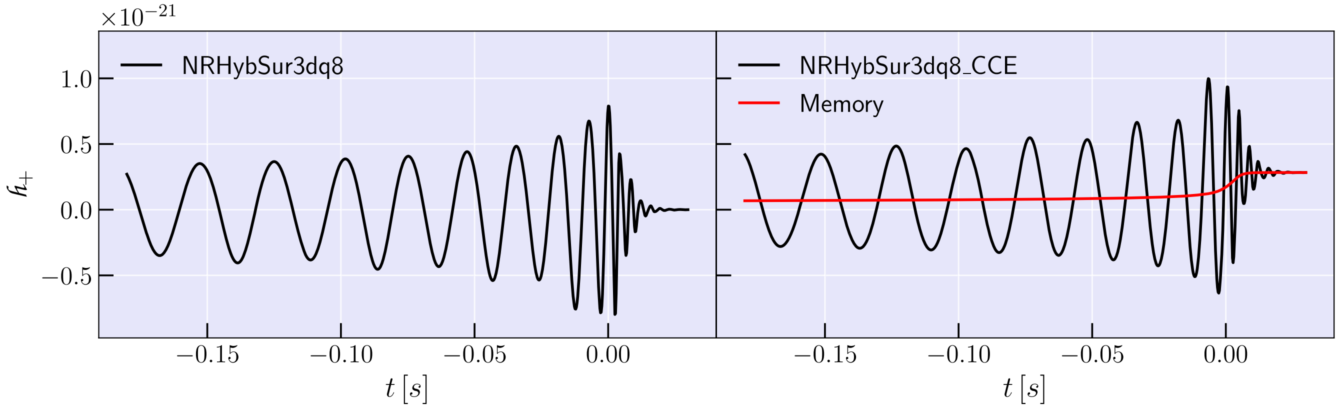

FIG. 1. Plus polarization of the strain for a GW150914-like event computed using the previous surrogate `NRHybSur3dq8` (left) and the new surrogate `NRHybSur3dq8_CCE` (right). The exact parameters that are used to calculate these two waveforms are $m_1 = 36M_\odot$, $m_2 = 29M_\odot, \chi_{1z} = 0.32, \chi_{2z} = -0.44$, $D_L = 410$ Mpc, $\iota = \pi/2$, and $\varphi = 0$. Because we wish to highlight the main difference between these surrogates, we used an inclination angle of $\iota = \pi/2$, i.e., the "edge on" orientation, for which the memory (red) (computed using Eq. (17b) of Ref. [66]) is maximized. Note that the memory roughly scales as $\sin^2(\iota)$.

`NRHybSur3dq8`, and in certain scenarios even outperforms `NRHybSur3dq8`, all while containing previously unresolved physical effects. As an illustration of our waveform model, we provide Fig. 1, which shows the correct waveform for a GW150914-like binary black hole merger event.

Throughout this paper, we adopt the common notations used by previous works. The mass ratio is denoted as $q = m_1/m_2$, where m_1 (m_2) denotes the mass of heavier (lighter) black hole, while the aligned spin of the heavier (lighter) black hole (in the direction of the binary's orbital angular momentum) is denoted as χ_{1z} (χ_{2z}). We use D_L to denote the luminosity distance, ι to denote the inclination angle between the orbital angular momentum and the line-of-sight to the detector, and φ to denote the azimuthal angle. Furthermore, when outputting our waveform templates, we represent the two polarizations of the gravitational wave—the plus and cross polarizations—as a single complex waveform, $h = h_+ - ih_\times$, which we then further decompose into a sum of spin-weight −2 spherical harmonic modes denoted as $h_{\ell m}$:

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

Here ${}_{-2}Y_{\ell m}$ are the spin-weight −2 spherical harmonics. In Eq. (1), the quadrupole modes ($\ell = |m| = 2$) typically dominate the sum; however, the other modes are also important for estimating binary source properties [75–79]. Therefore, our new model `NRHybSur3dq8_CCE` includes $\ell \leq 4$ and (5, 5) spin-weighted spherical harmonic modes, but not the (3, 1), (4, 2) and (4, 1) modes. The reason for excluding these three modes is explained in Appendix. Like its predecessor, the new model `NRHybSur3dq8_CCE` is an aligned-spin model, restricted to binary black holes (BBHs) whose spins are aligned with the system's orbital angular momentum. Thus, due to orbital-plane symmetry, we do not have to model the $m < 0$ modes separately as they can be obtained from $m > 0$ through the well-known relation $h_{\ell(-m)} = (-1)^\ell h^*_{\ell m}$, where $h^*_{\ell m}$ represents the complex conjugate of $h_{\ell m}$.

The rest of the paper is organized as follows. In Sec. II, we describe the entire construction of `NRHybSur3dq8_CCE`. In Sec. III, we then evaluate the errors involved in building this model. In particular, we check the error due to hybridization, the error of the surrogate itself, the success of the surrogate in extrapolating to values outside its training range, and the difference between `NRHybSur3dq8` and `NRHybSur3dq8_CCE`. Finally, in Sec. IV, we conclude with a few closing remarks. `NRHybSur3dq8_CCE`, our new surrogate model, has been made publicly available through the python package GWSURROGATE [80].

II. METHODS

In this section, we outline the steps that are required for building the new surrogate model `NRHybSur3dq8_CCE`. More specifically, in the subsequent text, we discuss the parameter space that our training waveforms will cover, Bondi-van der Burg-Metzner-Sachs (BMS) frame fixing, hybridization, and, finally, the routine for constructing the surrogate model `NRHybSur3dq8_CCE`.

A. Training set generation

To build the new surrogate model `NRHybSur3dq8_CCE`, we need a set of training waveforms in addition to their corresponding binary parameters. One cannot know, *a priori*, the optimal distribution of binary parameters for training the surrogate model. Fortunately, a previous surrogate model, `NRHybSur3dq8`, already explored the parameter space that we are interested in: mass ratio

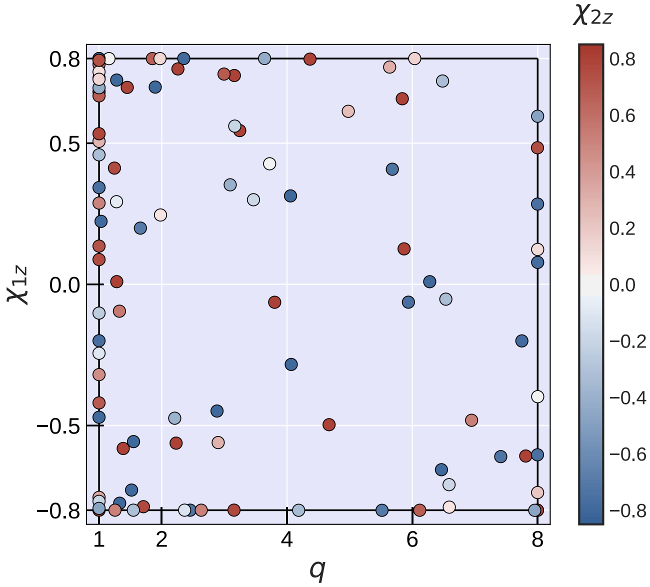

FIG. 2. Training set parameters used in the construction of the new surrogate model `NRHybSur3dq8_CCE`. There are a total of 102 training data points used for `NRHybSur3dq8_CCE`, which exactly match those of `NRHybSur3dq8` minus two points, for which the initial world tube data for CCE was not available. The boundary of the training region is represented with the black rectangle: $1 \leq q \leq 8$ and $-0.8 \leq \chi_{1z}, \chi_{2z} \leq 0.8$.

$q \in [1, 8]$ and $|\chi_{1z}|, |\chi_{2z}| \leq 0.8$, where χ_{1z} (χ_{2z}) is the spin of the heavier (lighter) black hole in the direction of the orbital angular momentum [22]. Hence, we use the same set of existing NR simulations (SXS:BBH:1419–1509, but not SXS:BBH:1468 or SXS:BBH:1488) that was used for training `NRHybSur3dq8`. For an equal mass simulation with unequal spins, we can exchange the two BHs to obtain an extra training data point. This is performed by applying a rotation (along the z-axis, defined as the axis of the orbital angular momentum of the BBH) by π to the waveform of $(q, \chi_{1z}, \chi_{2z}) = (1, \chi, \tilde{\chi})$ to obtain an extra waveform corresponding to $(q, \chi_{1z}, \chi_{2z}) = (1, \tilde{\chi}, \chi)$ for $\chi \neq \tilde{\chi}$. From the above 89 NR simulations, there are 13 of these cases, leading to 102 distinct training data.[9] In Fig. 2, we show the three-dimensional distribution $(q, \chi_{1z}, \chi_{2z})$ of our training parameters.

For each NR simulation used in our model, we extract the asymptotic waveform at future null infinity $\mathcal{I}^+$ using the SPECTRE code's implementation of CCE [69,70,72]. We run CCE on each of the four finite-radius worldtubes that the Cauchy evolution outputs. While in principle the CCEs that use these various worldtubes should yield valid and identical waveforms, we find that because of how the initial data for CCE is constructed there tends to be one worldtube that yields a more physically correct asymptotic waveform. We determine the best worldtube and waveform by examining which waveform's time derivative has the lowest L^2 norm after the ringdown phase. We find that this test is consistent with the previous method of checking which waveform and Weyl scalars minimally violate the five Bianchi identities [66,71,73,74], but tends to yield waveforms with less junk radiation in the inspiral phase. As for the resolution of the CCE, we simply use the highest setting possible which yields errors in the CCE that are well below the errors from the Cauchy evolution. The waveforms that we use are interpolated to a uniform time step of $0.1M$, which is a dense enough time array to capture the important features of the waveform, including those that emerge near and during the merger phase.

Like extrapolated waveforms, CCE waveforms contain "initial data transients" due to the imperfect initial data on the first null hypersurface of the characteristic evolution. These unphysical features, however, tend to persist much longer than those observed in the extrapolated waveforms. Fortunately, we find that by truncating the earlier parts of the CCE waveforms, we can avoid this issue when constructing our training set in every mode except the (3, 1) and (4, 2) modes. Therefore, we exclude these modes in our new model. For more details about these modes and why we choose to exclude them, see Appendix.

B. BMS frame fixing

When building surrogate models, it is important to ensure that the training waveforms are in the same frame. Otherwise, undesired gauge artifacts can complicate or even interfere with the various fitting and interpolation steps that are used when building the surrogate model. The earlier surrogate model, `NRHybSur3dq8`, for example, implemented a center-of-mass corrected version of the extrapolated waveforms, in which the waveforms were mapped to a "Newtonian" center-of-mass frame of the binary using the coordinate trajectories of the black hole apparent horizons from the simulation [12].

However, the true gauge degree of freedom possessed by the gravitational waves at future null infinity is not the usual Poincaré group, but the BMS group, which includes an infinite-dimensional group of transformations that are called supertranslations [81,82] in addition to the usual Poincaré transformations. As a result, before using NR waveforms for any analysis, one should first fix these BMS freedoms.[10] In Refs. [73,74] this task of fixing the BMS frame was performed by computing the various BMS

---

[9] Note that two NR simulations (SXS:BBH:1468 and 1488) are missing the world tube data that is necessary to produce the CCE waveforms of interest. This is why in this work we have 102 training data rather than the 104 training data used in [22].

[10] Note though that while fixing the BMS frame is important for modeling purposes, e.g., constructing or comparing waveform models, when examining waveforms at a point on the two-sphere, the only frame freedom that is relevant is the Poincaré freedom. This is because when looking at a point on the sky, supertranslations become degenerate with time translations.

charges that correspond to the symmetries of the BMS group and then finding the transformations that change those charges in a desired way. For example, to fix the translation and boost symmetries Refs. [73,74] found a transformation that mapped the center-of-mass charge to have a mean of zero. More specifically, they found the transformation which minimized the time integral of the L^2 norm of this charge over a three-orbit window. Similarly, for fixing the system's supertranslation freedom, the same works also found what supertranslation to apply to the waveforms by examining a charge known as the Moreschi supermomentum—an extension of the usual Bondi four-momentum. By finding the supertranslation that mapped the Moreschi supermomentum to the value expected from PN theory, it was found that NR waveforms could be made to much better agree with PN waveforms once they were mapped to the same BMS frame of PN. This frame is called the PN BMS frame.

Because the surrogate model that we are building is for hybrid waveforms in which PN waveforms are stitched to the NR waveforms, the natural BMS frame to work with is, similarly, the PN BMS frame. Therefore, to fix the BMS frame of our waveforms we use the frame fixing procedure described in Ref. [73] and the python module SCRI [83–86]. That is, we fix the translation and boost freedoms by mapping the center-of-mass charge to have a mean of zero, we fix the rotation freedom by mapping the angular velocity vector to be aligned with that of a PN waveform, and we fix the supertranslation freedom by mapping the NR Moreschi supermomentum to agree with the PN Moreschi supermomentum [73]. We perform this frame fixing using a three-orbit window that starts ~2500–3500M before the peak of the L^2 norm of the NR strain. This choice of BMS frame implies that the $\ell \geq 1$ components of the Moreschi supermomentum of our BBH vanish at $t \to -\infty$. This is equivalent to matching the PN memory terms with the NR system's memory over the hybridization window. This choice also implies that `NRHybSur3dq8_CCE` 's strains vanish at $t \to -\infty$.

C. Hybridization

Because of computational limits, NR simulations of BBHs typically only cover the last 20 orbits of inspiral. Thus, they are not nearly long enough to span the full LVK detection band for stellar mass binaries. More precisely, the initial frequency of (2, 2) mode of these waveforms falls within the LVK band, taken to begin at $f_{\mathrm{low}} = 20$ Hz, for total masses $M = m_1 + m_2 \gtrsim 60 M_{\odot}$. To address this limitation and extend the validity of our model to lower values of total mass, we hybridize the CCE waveforms that are obtained from NR with the early inspiral parts of EOB phase-corrected PN waveforms. We create pure PN waveforms using the python package GWFRAMES [87]. For the PN orbital phase we include nonspinning terms up to 4 PN order [88–92] and spinning terms up to 2.5 PN order [93–95]. For the PN amplitude we include non-memory terms to 3.5 PN order [96–98], nonspinning memory terms to 3 PN order, and spinning memory terms up to 2 PN order [73,99]. We use the TaylorT4 [100] approximant to compute the PN phase, but we replace this with an EOB-derived phase for the following reasons.

As noted in the previous work with `NRHybSur3dq8` [22], the accuracy of the inspiral parts from PN waveforms can be improved by replacing the PN phase with the phase that is derived from an EOB model, which undergoes an NR calibration. This improvement is typically larger for high mass-ratio systems, where the PN deviation from NR tends to be more significant, as is shown in Fig. 3. For the phase correction[11] to the PN waveforms used in this new surrogate model, we use the EOB model `SEOBNRv4_opt` [101,102].[12]

Currently, the waveforms produced by CCE contain "initial data transients" or "junk radiation" because of imperfect initial data that forces us to discard early parts of the waveforms [71]. We find that the transients present in the CCE waveforms typically last noticeably longer than the junk radiation of the extrapolated waveforms [70]. Because of this, we instead use a hybridization window that is closer to the merger: roughly $2500 - 3500M$ before the peak of the L^2 norm of the strain. The later window further necessitates our use of the EOB-corrected phase for the inspiral part fo the PN waveforms.

Once the NR waveforms are mapped to the same frame as the EOB-corrected PN waveforms via the procedures described in II B,[13] we then hybridize NR and PN together, for each spin-weighted spherical harmonic mode, $\hbar_{\ell m}$, as:

$$\hbar_{\ell m}^{\mathrm{hyb}} = \hbar_{\ell m}^{\mathrm{PN}} + f\left(\frac{t - t_1}{t_2 - t_1}\right)(\hbar_{\ell m}^{\mathrm{NR}} - \hbar_{\ell m}^{\mathrm{PN}}), \tag{2}$$

using the following transition function

$$f(x) = \begin{cases} 0 & x \leq 0, \\ \left(1 + \exp\left[\frac{1}{x-1} + \frac{1}{x}\right]\right)^{-1} & 0 < x < 1, \\ 1 & x \geq 1. \end{cases} \tag{3}$$

---

[11] The phase correction procedure is identical to that described in Sec. IV B of Ref. [22].

[12] While `SEOBNRv4_opt` is trained on extrapolated waveforms rather than CCE waveforms, we do not expect this feature to noticeably impact the surrogate because the phase evolution of these different waveforms should still be comparable. Still, it would be interesting to see how our surrogate model construction changes when using EOB waveforms that have been calibrated with CCE waveforms.

[13] Even though we are using EOB-corrected PN waveforms as the target waveform in the BMS frame fixing procedure, because the EOB correction tends to zero as $t \to -\infty$, the frame that we map our NR waveforms to is still consistent with the PN BMS frame. The EOB-correction simply helps ensure that our mapping to the PN BMS frame is as accurate as possible.

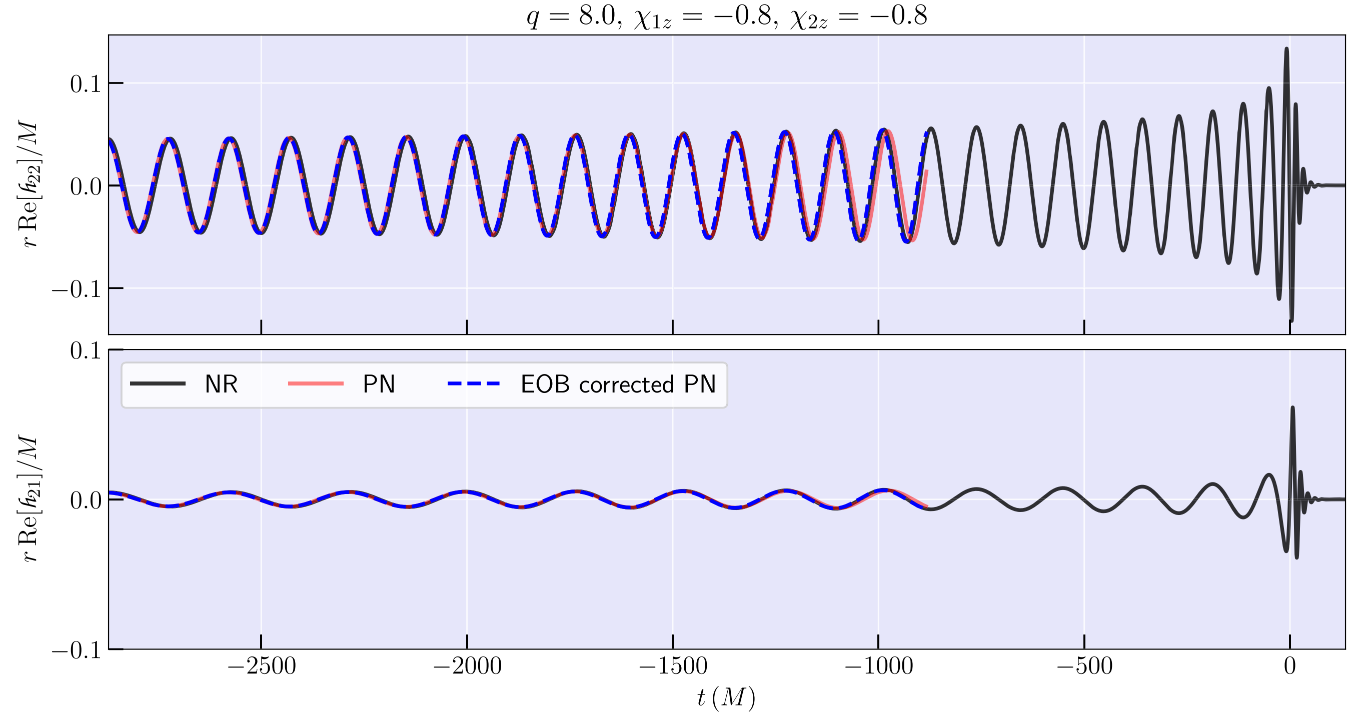

FIG. 3. NR (black), PN (red), and EOB-corrected PN (blue) waveforms for an example simulation with binary parameters $q = 8.0$, $\chi_{1z} = -0.8$, and $\chi_{2z} = -0.8$. The real part of both the (2, 2) and (2, 1) modes are shown in the top and bottom panels. Notice that the EOB-corrected PN waveform is more faithful to the NR waveform than the PN waveform.

Our choice of transition function matches that used in Ref. [74]. In the previous surrogate, `NRHybSur3dq8`, instead of hybridizing the complex waveforms directly, the waveforms were decomposed into amplitude and frequency before these components were hybridized independently. This was done to avoid undesirable artifacts that their transition function introduced, which are shown in Fig. 4 of Ref. [22]. Because CCE waveforms contain memory, which acts as a time-dependent offset in the waveforms, the decomposition into amplitude and frequency is not as useful as it is for the extrapolated waveforms which do not contain memory. Therefore, we choose to not use the previous hybridization method and instead directly stitch the NR and PN waveforms together. We do observe some minor glitches in the orbital frequency of a few of our hybrid waveforms in the hybridization window. However, these effects are largely negligible when compared to the other modeling errors for `NRHybSur3dq8_CCE`.

D. Postprocessing the training data

In this section, we now describe how we construct the surrogate model from the hybrid waveforms of Sec. II C.

1. Down-sampling and common time array

First, we apply a time shift to each training waveform such that the peak of the L^2 norm

$$A_{\mathrm{tot}}(t) = \sqrt{\sum_{\ell,m} |\mathcal{h}_{\ell m}(t)|^2} \tag{4}$$

is aligned at $t = 0$. The peak time of this curve is determined from a quadratic fit using 5 time samples that are adjacent to the discrete maximum of A_{tot}. When we compute the sum in Eq. (4), we use every mode of the hybrid waveform, including the $m < 0$ modes.

To begin with, the length of each hybrid waveform is determined by ensuring that the initial orbital frequency is $\omega_{\mathrm{orb}} = 2 \times 10^{-4}$ rad$/M$, where ω_{orb} is approximated from ϕ_{22}, the phase of the (2, 2) mode, using

$$\omega_{\mathrm{orb}} = \frac{1}{2}\frac{d\phi_{22}}{dt} \tag{5}$$

This frequency choice, however, results in waveforms with different lengths for different mass ratios and spins. The surrogate-building procedure, however, requires that every training waveform share a common time array. Therefore, to remedy this issue we truncate the training waveforms such that they begin with time $t = -5.8 \times 10^8 M$, which is the first time of the shortest waveform in the surrogate's training set. After truncation, the training set's largest starting orbital frequency is $\omega_{\mathrm{orb}} = 2.8 \times 10^{-4}$ rad$/M$. Consequently, this frequency is the low-frequency limit of validity for the surrogate model.

For the LVK observatories, if we assume that 20 Hz is the lowest GW frequency that can be measured, then the

(2, 2) mode of the surrogate model can be considered valid for total masses $M \geq 0.9 M_\odot$. The highest spin-weighted spherical harmonic mode included in `NRHybSur3dq8_CCE` is the (5, 5) mode for which the corresponding frequency is a factor of 5/2 more than that of the (2, 2) mode. Thus, the entire surrogate is valid for total masses $M \geq 2.25 M_\odot$.

Because the hybrid waveforms are millions of M long, it is not practical to sample the entire waveform with a small uniform time step like $0.1M$, as is typically used for NR-only surrogates [23]. Fortunately, the early inspiral, low-frequency portion of the waveform does not require as dense a sampling as the later high-frequency portion. Therefore, we instead down-sample the time arrays of the truncated waveforms such that there are only 5 points per orbit for the shortest hybrid waveform of the training set. However, for $t \geq -1000M$, we switch to uniformly spaced time samples with a time step of $0.1M$ to ensure that we have sufficiently dense sampling for the late inspiral and merger-ringdown phases where the frequency reaches its peak. We retain times up to $135M$ after the peak to ensure that we fully capture the numerically resolvable parts of the ringdown phase.

Given the common down-sampled time array, we then use cubic splines to interpolate all of the waveforms in the training set onto the common time array. However, we first transform the waveforms to the coorbital frame, which we construct via

$$\text{co-orbital frame:} \begin{cases} \mathpzc{h}^C_{\ell m} = \mathpzc{h}_{\ell m} e^{i m \phi_{\mathrm{orb}}} \\ \mathpzc{h}_{22} = A_{22} e^{i \phi_{22}} \\ \phi_{\mathrm{orb}} = \phi_{22}/2 \end{cases} \tag{6}$$

where $\mathpzc{h}_{\ell m}$ is the inertial frame waveform, ϕ_{orb} is the orbital phase, and A_{22} and ϕ_{22} are the amplitude and phase of the (2, 2) mode. The co-orbital frame is roughly corotating with the binary and is obtained by applying a time-dependent rotation about the $\hat{z}$-axis by an amount measured by the instantaneous orbital phase. As the waveforms are slowly varying functions of time in the co-orbital frame, by transforming to this frame we can increase the interpolation accuracy. For the (2, 2) mode, we sample the amplitude A_{22} and phase ϕ_{22}, while for all other modes we use the real and imaginary parts of $\mathpzc{h}^C_{\ell m}$.

2. Phase alignment

After interpolating to a common time array, we align the phases of the waveforms by rotating the waveforms about the $\hat{z}$-axis such that the orbital phase ϕ_{orb} is zero at time $t = -1000M$. This ensures that each waveform corresponds to a binary with its heavier black hole on the $\hat{x}$-axis at that time. Note that this frame is constructed using information from the waveform at future null infinity, and as a result these BH positions need not correspond to the gauge-dependent BH positions in the NR simulations.

TABLE I. Number of basis functions for each data piece.

Data piece	Number of basis functions
A_{22}, ϕ_{22}	15
$\mathrm{Re}\mathpzc{h}^C_{21}$, $\mathrm{Im}\mathpzc{h}^C_{21}$	12
$\mathrm{Re}\mathpzc{h}^C_{20}$	12
$\mathrm{Im}\mathpzc{h}^C_{30}$	12
$\mathrm{Re}\mathpzc{h}^C_{43}$, $\mathrm{Im}\mathpzc{h}^C_{43}$	12
$\mathrm{Re}\mathpzc{h}^C_{55}$, $\mathrm{Im}\mathpzc{h}^C_{55}$	8
Others	10

3. Data decomposition

As mentioned earlier, it is easier to build a model for slowly varying functions of time. Because of this, we decompose the inertial frame strain $\mathpzc{h}_{\ell m}$, which is oscillatory, into simpler waveform data pieces and build a separate surrogate model for each of these data pieces. For the (2, 2) mode, we decompose this mode into the amplitude A_{22} and the phase ϕ_{22}, while for the other $m \neq 0$ waveform modes, we model the real and imaginary parts of the coorbital frame strain, $\mathpzc{h}^C_{lm}$, using Eq. (6). For the $m = 0$ modes of nonprecessing systems, $\mathpzc{h}^C_{\ell m}$ is purely real (imaginary) for even (odd) ℓ. Because of this, we only model the nontrivial part for the $m = 0$ modes.

Because our hybrid waveforms are rather long—extending over roughly 3×10^4 orbits—ϕ_{22} roughly spans 10^5 radians. Accurately modeling the phase evolution of such long hybrid waveforms poses a challenge. We find that we can resolve this issue, however, by subtracting the leading-order Taylor T3 PN phase [103], ϕ^{T3}_{22}, and simply modeling the phase residual, $\phi^{\mathrm{res}}_{22} = \phi_{22} - \phi^{\mathrm{T3}}_{22}$, as was performed in Ref. [22]. The leading-order prediction from the TaylorT3 PN approximant [103] is

$$\phi^{\mathrm{T3}}_{22} = \phi^{\mathrm{T3}}_{\mathrm{ref}} - \frac{2}{\eta \theta^5} \tag{7}$$

with

$$\theta = [\eta (t_{\mathrm{ref}} - t)/(5M)]^{-1/8}, \tag{8}$$

where $\phi^{\mathrm{T3}}_{\mathrm{ref}}$ is an arbitrary integration constant, t_{ref} is an arbitrary time offset, and $\eta = q/(1+q)^2$ is the symmetric mass ratio. Because, by definition, θ diverges at $t = t_{\mathrm{ref}}$, to avoid such divergences we set $t_{\mathrm{ref}} = 1000M$, which is sufficiently long after the end of the waveform. We also choose $\phi^{\mathrm{T3}}_{\mathrm{ref}}$ such that $\phi^{\mathrm{T3}}_{22} = 0$ at $t = -1000M$, i.e., the time at which we align the phase, as outlined in Sec. II D 2.

E. Surrogate building

Given the decomposed waveform data pieces, we build a surrogate model for each individual data piece using the same procedure as that of Sec. V. C of Ref. [22], with a few minor modifications, which we summarize below.

For each waveform data piece, we begin by constructing a linear basis in parameter space, so that we can reduce the training dataset to a smaller representative dataset. The basis functions that we use are chosen in the following iterative manner [104–107], called the "greedy algorithm":

1. Pick out the training data with the largest L^2 norm and add it to the basis set as the first basis function;
2. Compute the projection error between each of the training data and the basis set;
3. Determine which of the training data has the highest projection error and add this to the basis set;
4. Repeat steps 2–3 until a predetermined number of basis functions for each data piece is obtained.

For step 4, we determine the number of basis functions used for each of the data pieces through trial and error. That is, we increase the number of basis functions until the inclusion of new basis functions introduces noise into the model or gives diminishing returns in terms of minimizing the projection errors. The number of basis functions that is used for each data piece is shown in Table I.

Next, we build empirical time interpolants [105,108–111] with the same number of empirical nodes as the number of basis functions that are used to model the data piece. Following the methodology of Ref. [22] we also require that the start of the waveform always be included as one of the empirical nodes. This provides an "anchor point" that ensures that the waveform data pieces start with the correct value. In Ref. [22], no empirical nodes were picked at times past $t > 50M$ to ensure that little to no numerical noise was being modeled, particularly for the phase data piece. We follow the same convention for the phase data pieces; however, for the other data pieces we allow time nodes past $t > 50M$ to ensure that the surrogate correctly models the memory throughout the entirety of the ringdown phase.

Finally, for each empirical time node, we construct a parametric fit for the waveform data piece, following the Gaussian process regression (GPR) fitting method described in the supplementary material of Ref. [112], using the python package SCIKIT-LEARN [113]. For the fit itself, we use the parameterization used in Ref. [22]. That is, we use the parameters $\log(q), \hat{\chi}$, and χ_a, where $\hat{\chi}$ is the leading order spin parameter [114–117] for the GW phase in the PN expression

$$\hat{\chi} = \frac{\chi_{\mathrm{eff}} - 38\eta(\chi_{1z} + \chi_{2z})/113}{1 - 76\eta/113}, \tag{9}$$

with

$$\chi_{\mathrm{eff}} = \frac{q\chi_{1z} + \chi_{2z}}{1 + q}, \tag{10}$$

and χ_a is the antisymmetric spin defined as

$$\chi_a = \frac{1}{2}(\chi_{1z} - \chi_{2z}). \tag{11}$$

III. ERROR QUANTIFICATION

With the methodology behind the construction of our surrogate model outlined in Sec. II, we now examine the quality of `NRHybSur3dq8_CCE` by conducting a variety of model consistency checks and waveform comparisons. This process involves examining both the error in the hybridization and the error in the surrogate model itself as well as the ability of the surrogate to model waveforms outside of its training parameter space.

A. Hybridization errors

Before building the surrogate model, we first have to map the NR waveforms to be in the same frame as the EOB-corrected PN waveforms. As described in Sec. II B, this fixes the BMS freedom of the NR waveforms by mapping the waveforms to the PN BMS frame using the procedure outlined in Ref. [73]. Following this, we then stitch the two waveforms to create a hybrid waveform that we use as our training waveform. One of the important checks for the quality of our final training waveform is to understand the errors that result from performing this hybridization procedure between NR and PN waveforms. There are two natural ways to conduct this check.

First, we can simply compare the hybrid waveform to the NR waveform in the hybridization window. This check will determine the combined discrepancy that results from the discrepancy between the PN and NR waveforms and from using the smooth transition function in Eq. (3). Next, to examine the error introduced from using PN for the early inspiral parts of our hybrid waveform, we can compare the hybrid to a simulation that has identical binary parameters, but a larger initial BH separation. Fortunately, of the simulations that we include in our surrogate's training dataset, 37 of them have these longer extensions that we can use for a comparison with the hybrid waveforms.

To quantify the error for these comparisons, we use the following function:

$$\mathcal{E}[\hbar^{(1)}, \hbar^{(2)}] \equiv \frac{1}{2}\frac{\sum_{\ell,m}\int_{t_1}^{t_2}|\hbar^{(1)}_{\ell m}(t) - \hbar^{(2)}_{\ell m}(t)|^2 dt}{\sum_{\ell,m}\int_{t_1}^{t_2}|\hbar^{(1)}_{\ell m}(t)|^2 dt}. \tag{12}$$

For the first comparison, we compute the error within the hybridization window, where t_1 is roughly $3000M$ before merger and t_2 is the time at which the system has undergone 3 orbits since t_1. For the latter comparison, we first map the long NR waveforms to the BMS frame of the training NR waveforms and then compute the error between these waveforms from $t_1 = -6000M$ to the end of the hybridization widow. The results of these comparisons are shown in Fig. 4. We observe that our two estimates for the hybridization error are low, but tend to be higher than the estimate of the NR resolution error. Note, though, that the NR resolution errors, which were computed for a smaller subset of NR simulations for which the higher

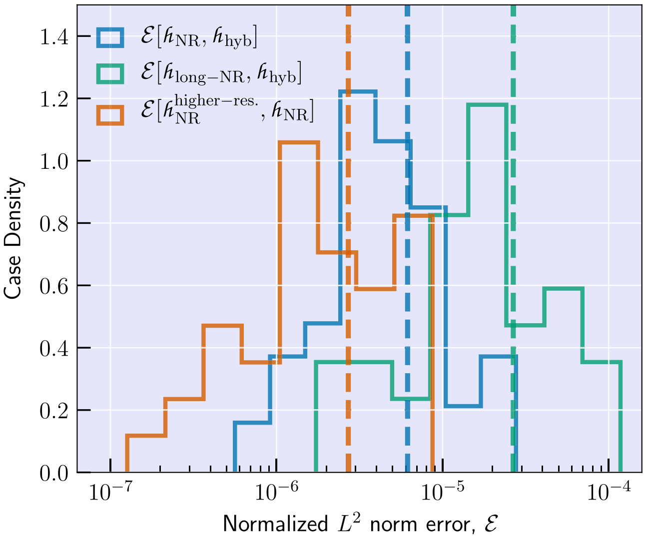

FIG. 4. Two unique estimates of the hybridization error. $\mathcal{E}[h_{\rm NR}, h_{\rm hyb}]$, using Eq. (12), computes the error between the hybrid and the NR waveforms in the hybridization window. Meanwhile, $\mathcal{E}[h_{\rm long-NR}, h_{\rm hyb}]$ computes the error between the hybrid and the long NR waveforms from an initial time of $t_1 = -6000M$ to the end of hybridization window. We also include $\mathcal{E}[h_{\rm NR}^{\rm higher-res}, h_{\rm NR}]$ as a resolution error between the two highest resolution NR waveforms, computed within the hybridization window. $\mathcal{E}[h_{\rm NR}, h_{\rm hyb}]$ was computed for every one of the 102 training waveforms, while $\mathcal{E}[h_{\rm long-NR}, h_{\rm hyb}]$ and $\mathcal{E}[h_{\rm NR}^{\rm higher-res}, h_{\rm NR}]$ were computed for the 37 waveforms for which longer and higher resolution simulations were available. The dashed lines represent the median values.

resolution data was available, come from simulations that are for more comparable mass binaries. Because the difficulty of NR simulations increases with mass ratio, we therefore expect that the NR resolution error that we computed is a minor underestimate of the true error. Regardless, Fig. 4 suggests that our training waveforms are primarily limited by the accuracy of the PN waveforms, rather than the NR resolution error.

B. Model errors

We now evaluate the accuracy of the surrogate model NRHybSur3dq8_CCE by comparing the waveforms that it produces to the hybridized PN/NR waveforms that were used to train it. We quantify this model accuracy by computing the frequency-domain mismatch $\mathcal{M}$ between two waveforms h_1 and h_2 via

$$\mathcal{M} = 1 - \frac{\langle h_1, h_2 \rangle}{\sqrt{\langle h_1, h_1 \rangle \langle h_2, h_2 \rangle}} \tag{13}$$

with

$$\langle h_1, h_2 \rangle = 4\mathrm{Re}\left[\int_{f_{\rm min}}^{f_{\rm max}} \frac{\tilde{h}_1(f)\tilde{h}_2^*(f)}{S_n(f)} df\right], \tag{14}$$

where $\tilde{h}(f)$ denotes the Fourier transform of the strain $h(t)$, * the complex conjugate, and $S_n(f)$ the one-sided power spectral density of, say, a GW detector. The mismatches are optimized over shifts in time, polarization angle, and the initial orbital phase following the procedure described in Appendix D of Ref. [20]. For each pair of waveforms, we compute mismatches at a total of 37 sky points that are uniformly distributed over the two-sphere.

Before performing the Fourier transform, we taper both ends of the time domain waveform.[14] The tapering at the start of the waveform is done over 1.5 cycles of the (2, 2) mode. Because the waveforms that we are examining include memory effects, tapering them in the ringdown region can produce a significant level of windowing effects[15] in the Fourier spectrum. Therefore, before computing mismatches we first pad the end of the waveforms with their final values for $1000M$ and then taper them over this padded region. We find that a padding length of $1000M$ is enough to significantly reduce windowing effects from tapering and that the mismatch result is not very sensitive to this choice of padding length.

Because we use all available hybrid waveforms for the training of the model NRHybSur3dq8_CCE, if we compute the mismatch of our model against hybrid waveforms we would obtain a training error, rather than an estimate of the true modeling error. Thus, we instead estimate the out-of-sample error by performing leave-five-out analyses for the NRHybSur3dq8 and NRHybSur3dq8_CCE models. We construct exactly 20 trial surrogates, leaving out 5 or 6 waveforms from the training set for each surrogate. By calculating the mismatch between each surrogate and the left-out waveforms, we are then able to assess the performance of each surrogate against waveforms that were not used in the training process.

The left panel of Fig. 5 shows the mismatches for NRHybSur3dq8_CCE that are computed using a flat (white) noise curve ($S_n = 1$) over the late inspiral part (NR part) of the hybrid waveforms, truncating the waveforms to start at $t = -3500M$ for NRHybSur3dq8 and at $t = -2500M$ for NRHybSur3dq8_CCE.[16] We define $f_{\rm min}$ to be the frequency of the (2, 2) mode at the end of the initial tapering window, and $f_{\rm max} = 5 f_{22}^{\rm peak}$, where $f_{22}^{\rm peak}$ is the frequency of the (2, 2) mode at its peak. The choice of

---

[14] A value mismatch at both ends of the waveform tends to result in the presence of Gibbs phenomenon in the Fourier spectrum. To avoid this, we taper the waveform to zero at both ends using a Planck window [118].

[15] A new scheme for preprocessing that potentially reduces this windowing effect is proposed in Ref. [119].

[16] The discrepancy is due to the later hybridization window that was used for NRHybSur3dq8_CCE as explained in Sec. II C.

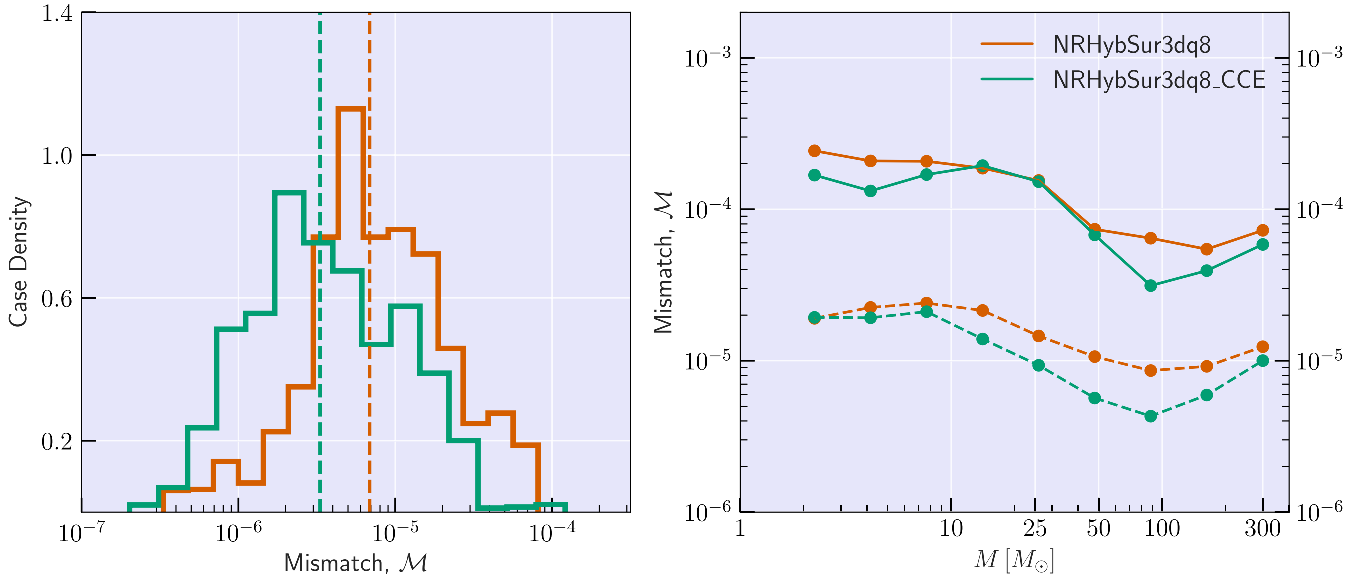

FIG. 5. Frequency-domain mismatches for the `NRHybSur3dq8` and `NRHybSur3dq8_CCE` surrogates when compared to their respective training hybridized waveforms (extrapolated and CCE, respectively). The results for `NRHybSur3dq8` are taken from Fig. 6 of Ref. [22] and are plotted for comparison with `NRHybSur3dq8_CCE`. For both of these surrogate models, we compute leave-five-out errors at several points in the sky of the source frame using all available modes in each model: $\ell \leq 4$ and (5, 5), but excluding (3, 1), (4, 2), and (4, 1) for `NRHybSur3dq8_CCE` and (4, 1) and (4, 0) for `NRHybSur3dq8`. Left: mismatches computed using a flat noise curve for only the late inspiral part (NR part) of the hybrid waveforms. The dashed vertical line represents the median mismatch value. Right: mismatches computed using the Advanced-LIGO noise curve as a function of the system's total mass. The solid (dashed) line represents 95th percentile (median) mismatch values.

$f_{\max}$ ensures that we capture the peak frequency of every mode.

The right panel of Fig. 5 show results that are more relevant to GW observations: namely, the mismatches computed with the Advanced-LIGO design sensitivity zero-detuned-high-power noise curve [120] with $f_{\min} = 20$ Hz and $f_{\max} = 2000$ Hz for various total masses. We also include the mismatch result for `NRHybSur3dq8` from Ref. [22] for comparison. On the horizontal axis, different total masses correspond to different portions of the waveform falling within the [20 Hz, 2000 Hz] window. This roughly implies that the low end of the total mass axis is a proxy for the fidelity of the early inspiral part of the waveforms, while the high end is a proxy for the late inspiral parts. We show the mismatch for various total masses: from the lower limit of the range of validity of the surrogate, i.e., $M \gtrsim 2.25 M_\odot$, up to $M = 300 M_\odot$. For each total mass point that we plot, we show both the median and the 95th percentile mismatches.

For the surrogate modeling errors, we obtain values that are comparable to—and often better than—those of `NRHybSur3dq8`, despite the additional modeling challenges resulting from the new contributions due to the presence of memory. The 95th percentile mismatches fall below $\sim 3 \times 10^{-4}$ for the entire mass range.

Last, we test how well the surrogate models both the displacement and spin memory contributions. To do this, we compute the displacement and spin null memories, i.e., Eqs. (17b) and (17d) from Ref. [66] which only depend on the strain, using both the training waveforms and the surrogate evaluation, and then calculate the normalized L^2 norm error between the two using Eq. (12) over the $t \in [-1000M, 135M]$ window. The choice of the window's starting time is arbitrary but early enough to capture most of the memory. For this check, we again perform leave-five-out analyses to estimate `NRHybSur3dq8_CCE` 's modeling error. As a reference for these errors, we also show the numerical resolution error from the same set of 37 simulations used for Fig. 4. As shown in Fig. 6, the modeling error for the two memory effects is at a reasonable level, albeit higher than our estimate of the resolution error. As shown in Fig. 7, even for the case that corresponds to the largest error value in Fig. 6, the memory effects computed from the surrogate evaluation closely agree with those computed from NR waveforms.

We suspect that the spin memory accuracy is typically worse than that of the displacement memory because the spin memory is smaller than the displacement memory and is thus harder to resolve. The discrepancy between the modeling error and the resolution error suggests that there is room for future improvements in modeling the memory contribution. Building future surrogate models using CCE waveforms from higher resolution and longer NR simulations might help improve this modeling error; however, a new data decomposition scheme or even a new modeling

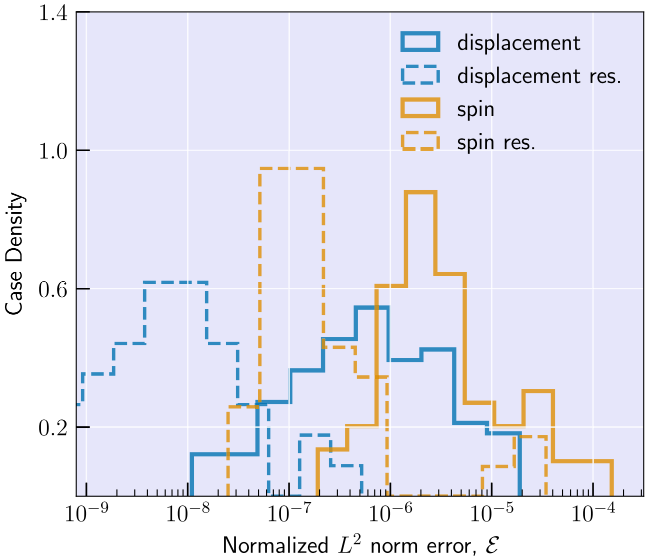

FIG. 6. Errors for the displacement (blue) and spin (orange) memory effects computed by comparing the training waveforms to `NRHybSur3dq8_CCE` 's output using Eq. (12). For this result, we show the leave-five-out cross-validation errors (solid line). As a reference, we also include the errors computed between the two highest resolution waveforms (dashed line).

strategy could be necessary to obtain an improved modeling of memory effects.

C. Extrapolating outside training region

The errors that we have examined thus far have been restricted to the training region of the parameter space: $q \leq 8, \chi_{1z}, \chi_{2z} \in [-0.8, 0.8]$. It is possible, however, to evaluate the surrogate outside the training region, e.g., for larger mass ratio, q, or even higher primary or secondary spin magnitudes: $|\chi_{1z}|$, $|\chi_{2z}|$. Consequently, to understand the extrapolation capability of the model, we compute errors of the model against a few existing simulations (SXS:BBH:0185, 0189, 0199, 1124 2085, 2105, 2132, and 2515) [12,19,37,121] that have relatively high mass ratios or spin magnitudes.

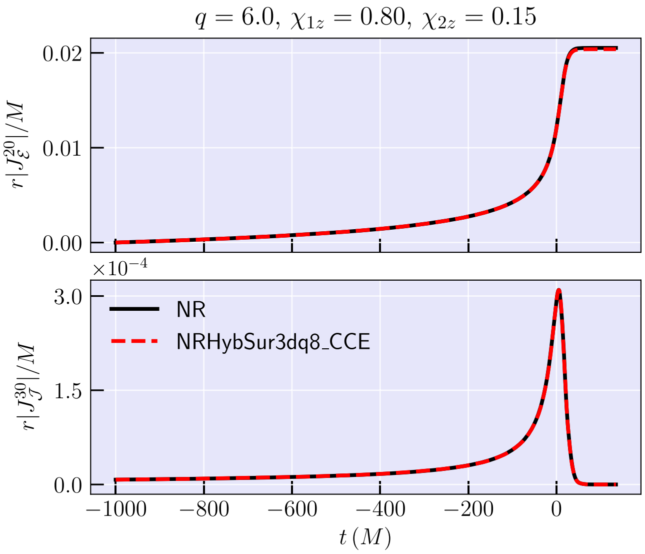

FIG. 7. Amplitudes of the most dominant modes of the displacement and spin memory effects when computed from the NR waveform and the surrogate evaluation for the case with the largest errors in Fig. 6. The top and bottom panels show the (2, 0) mode of the displacement null memory and the (3, 0) mode of the spin null memory.

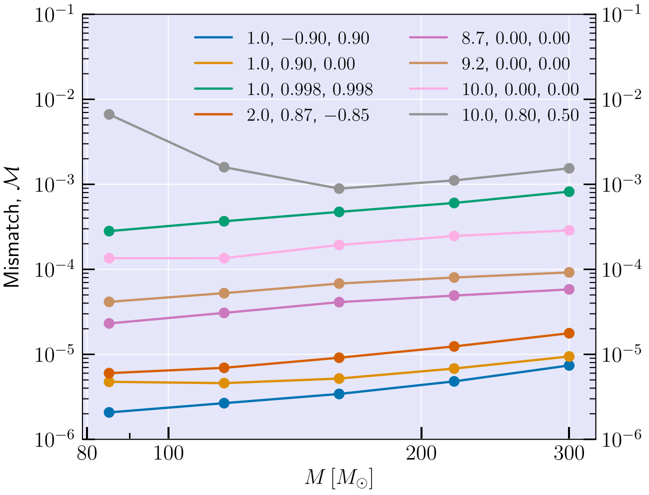

FIG. 8. Noise-weighted frequency-domain mismatches between `NRHybSur3dq8_CCE` and NR simulations that are outside the model's training region. The numbers in the legend correspond to the NR simulation's mass ratio q, primary spin χ_{1z}, and secondary spin χ_{2z}. The mismatches that are shown are computed at several points in the sky of the source frame using the Advanced-LIGO noise curve. Each of the solid lines represents the 95th percentile mismatch values.

As shown in Fig. 8, the mismatch results, while worse than those shown in Fig. 5, are nonetheless reasonable. The highest three mismatch results correspond to the three most extreme parameter simulations: mass ratio $q = 10$ or spin magnitude $|\chi_{1z}|, |\chi_{2z}| = 0.998$, for which we do not expect the `NRHybSur3dq8_CCE` to perform well. Apart from these, we find that the surrogate performs well for a modest degree of extrapolation, with many of the mismatches falling below values near $\sim 10^{-4}$.

D. Systematic bias in waveforms that omit memory or add it through postprocessing

Finally, in Fig. 9, we show the mismatch between the training waveforms used for `NRHybSur3dq8_CCE` and those that were used for `NRHybSur3dq8`. The primary purpose of this analysis is to obtain a rough estimate of the level of systematic bias that one could expect from analyzing a GW signal that contains memory using a waveform that does not. Apart from this, we also show the mismatch between an extrapolated waveform, once we have added the expected memory contribution to it using Eq. (17b) of

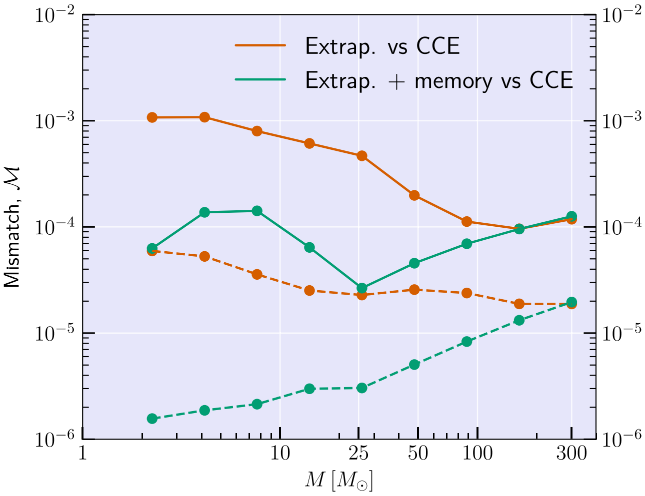

FIG. 9. Noise-weighted frequency-domain mismatches between extrapolated and CCE hybrid waveforms in red. These are the hybrid waveforms that were used to train `NRHybSur3dq8` and `NRHybSur3dq8_CCE`. Apart from this, we also include the mismatches between the memory-corrected extrapolated and CCE waveforms in green. The mismatches that are shown are computed at several points in the sky of the source frame using the Advanced-LIGO noise curve. The solid (dashed) line represents the 95th percentile (median) mismatch values.

Ref. [66], and a CCE waveform from the same simulation. This highlights that, while there is a noticeable difference between CCE waveforms and extrapolated waveforms, these discrepancies are largely reduced by adding memory to the extrapolated waveform. Because of this result, we suspect that memory-detection studies that have used this memory-correction technique, like Refs. [55–57,122,123], would likely obtain similar estimates had they used our new surrogate model `NRHybSur3dq8_CCE`. Note though that `NRHybSur3dq8_CCE` will be better for performing analyses of the spin memory, seeing as the contribution of spin memory to extrapolated waveforms cannot be as easily corrected [66]. Regardless, it would still be interesting to see if the conclusions made by these studies on memory, or even parameter estimation results, change when using `NRHybSur3dq8_CCE` instead of `NRHybSur3dq8`.

IV. CONCLUSION

In this work, we present a new surrogate model, `NRHybSur3dq8_CCE`, the first GW model to contain both the oscillatory and memory components of the strain. Consequently, `NRHybSur3dq8_CCE` is the first model to fully capture the expected GW physics of binary black hole mergers. The model is trained on 102 NR/PN hybrid waveforms from aligned-spin binary BH systems with mass ratios $q \leq 8$ and aligned spins $|\chi_{1z}|, |\chi_{2z}| \leq 0.8$. These hybrid waveforms are constructed by first mapping CCE waveforms to the PN BMS frame before hybridizing them with PN waveforms whose phase has been corrected using EOB waveforms. Performing this frame fixing helps eliminate unwanted gauge artifacts that could potentially interfere with modeling. The model includes $\ell \leq 4$ and (5, 5) spin-weighted spherical harmonic modes, but not the (3, 1), (4, 2), or (4, 1) modes, and spans the entire LVK band (with $f_{\mathrm{low}} = 20$ Hz) for total masses $M \geq 2.25 M_\odot$. By conducting a series of leave-five-out cross-validation analyses, we find that `NRHybSur3dq8_CCE` can accurately reproduce the hybrid waveforms that were used to train it with mismatches below $\sim 3 \times 10^{-4}$ for total masses in the range $2.25 M_\odot \leq M \leq 300 M_\odot$. These accuracies are on par with—and often better than—the previous aligned-spin NR surrogate model, `NRHybSur3dq8`, despite the modeling challenges that result from the inclusion of new modes and memory effects. Apart from this, we also importantly find that `NRHybSur3dq8_CCE` can successfully capture the null memory contributions with mismatches below $\sim 2 \times 10^{-4}$. Last, `NRHybSur3dq8_CCE` is also found to reproduce waveforms outside of its trained region of parameter space for a moderate degree of extrapolation; however, we advise caution when extrapolating the model. This new model is made publicly available through the python package GWSURROGATE [80].

With the expected advances in detector sensitivity for both current and future gravitational wave observatories, waveform templates with memory effects will prove to be crucial for analyzing future compact binary detections. The new surrogate model `NRHybSur3dq8_CCE` serves as the first step in an important endeavor to produce a complete set of waveform templates that contain these undetected effects and thus correctly capture the expected gravitational wave physics of binary black hole mergers.

ACKNOWLEDGMENTS

This work was supported in part by the Sherman Fairchild Foundation and by NSF Grants No. PHY-2011961, No. PHY-2011968, and No. OAC-2209655 at Caltech, and NSF Grants No. PHY-2207342 and No. OAC-2209655 at Cornell. SPECTRE uses CHARM++/CONVERSE [124], which was developed by the Parallel Programming Laboratory in the Department of Computer Science at the University of Illinois at Urbana-Champaign. V. V. acknowledges support from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant agreement No. 896869. L. C. S. was partially supported by NSF CAREER Award PHY-2047382. S. E. F acknowledges partial support from NSF Grant PHY-2110496 and by UMass Dartmouth's Marine and Undersea Technology (MUST) Research Program funded by the Office of Naval Research (ONR) under Grant No. N00014-23-1–2141. This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the NSF. This project made use of PYTHON libraries including SciPy [125] and NumPy

[126], and the figures were produced using Matplotlib [127] and SEABORN [128].

APPENDIX: CHALLENGES IN MODELING CERTAIN MODES

1. Impact of long-lived transient junk

The waveforms produced by CCE contain initial data transients that typically persist much longer than those of the Cauchy evolution. Because of this effect, as was outlined in Sec. II C, we choose to remove the early parts of our NR waveforms when constructing our surrogate's training waveforms. Nevertheless, even after we perform this truncation, some of the waveforms that we use for the surrogate still exhibit unphysical effects that we do not see when using extrapolated waveforms, or CCE waveforms from higher resolution or longer Cauchy evolutions.

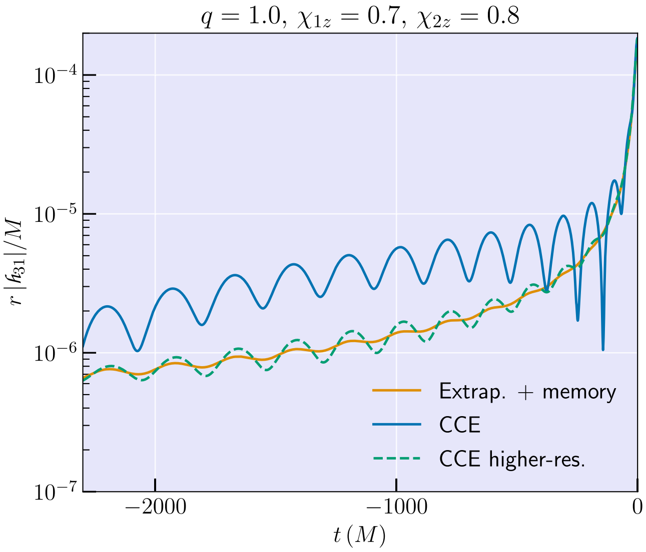

FIG. 10. An example of the unphysical oscillations that are seen in the amplitude of the (3, 1) mode for one of the CCE training waveforms (blue). As a reference, we also show the extrapolated waveform (orange) for the same simulation and the higher-resolution CCE waveform (green).

We find that initial data transients are most pronounced in the (3, 1) and (4, 2) modes, and are often identified by unphysical amplitude oscillations, as shown in Fig. 10. As is the case with all models, the surrogate model `NRHybSur3dq8_CCE` is only as good as the data that it is trained on. In fact, we find that having even only a few training waveforms with these issues can result in noisy and unphysical features being modeled by the surrogate.

We find that these unphysical oscillations in certain modes are significantly reduced for waveforms extracted from both higher resolution and longer Cauchy evolutions. Unfortunately, such simulations only exist for one third of our training dataset. Therefore, we instead choose to not model the (3, 1) and (4, 2) modes in our new surrogate `NRHybSur3dq8_CCE`. We also omit the (4, 1) mode because it is subdominant to the (4, 2) mode and, as a result, does not significantly impact the overall waveform accuracy provided that the (4, 2) is already excluded. For future surrogate models that are built using CCE waveforms, it is important that we have higher resolution and longer NR simulations to avoid these issues until the problem of constructing initial data for CCE is resolved.

2. Impact of a simulation-dependent BMS frame

As stressed in Sec. II B, it is crucial to ensure that the training waveforms are in the same BMS frame to avoid undesired frame artifacts that tend to complicate the

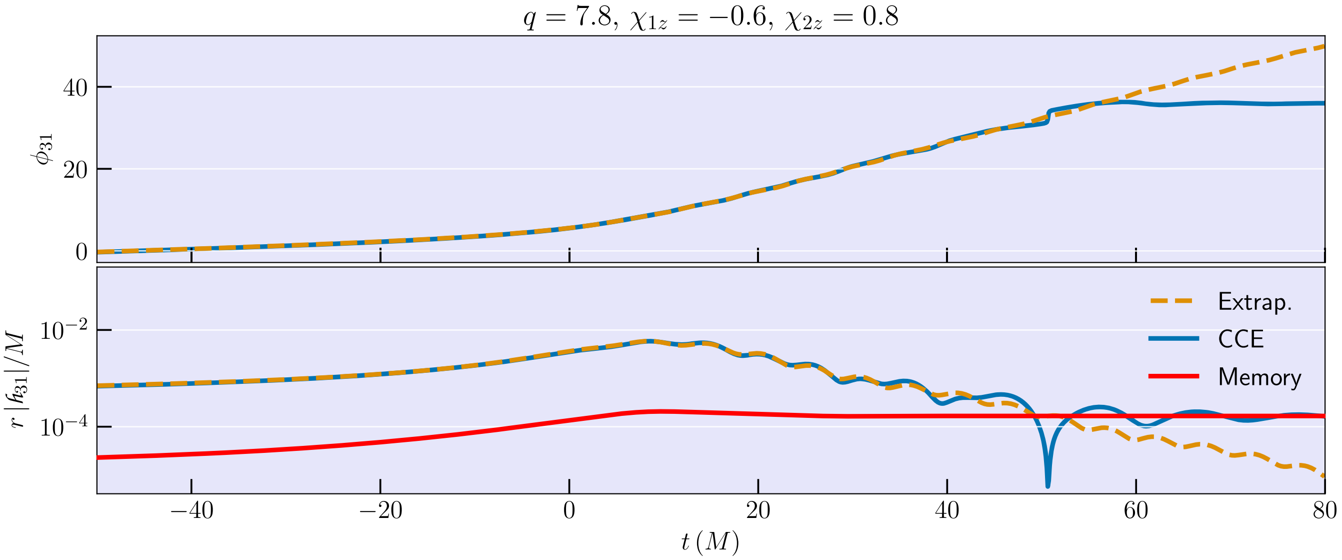

FIG. 11. Phase and amplitude of the (3, 1) mode for a CCE (blue) and an extrapolated (orange) waveform. We also include the memory contribution (red) to highlight the fact that the CCE waveform does not decay to zero as $t \to \infty$.

waveform modeling. Because the surrogate model that we are building is for hybrid waveforms, we work with waveforms that are in the PN BMS frame. However, while this ensures that the waveforms are in the same frame during the inspiral phase,[17] this is not necessarily true for the ringdown phase [129]. This is because, for example, the remnant black holes can have different kick velocities or supertranslation fields because of complicated effects that arise during the merger phase. Consequently, there are unresolved frame artifacts during the ringdown phase that can impact the modeling of the strain waveforms.

These additional challenges of modeling the strain in the ringdown using the current PN BMS framework are most pronounced for the (3, 1) and (4, 2) modes where the strain is oscillatory and does not decay to zero due to the impact of memory. Because these modes do not decay to zero as $t \to \infty$, we find that the decomposition of the strain into coorbital frame data does not work well as the phase of these two modes is ill-defined. We highlight this phase issue in Fig. 11, which shows the amplitude and the phase of the (3, 1) mode for both a CCE waveform and an extrapolated waveform. One potential remedy to this problem is to instead work in a co-BMS frame, in which there is little-to-no time evolution of the BMS charges, i.e., a noninertial frame similar to the corotating frame that simplifies the waveform data. However, such a project is nontrivial and we therefore postpone it for future work.

---

[17] Formally, this only ensures that the waveforms are in the same BMS frame at $t \to -\infty$.

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

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

[3] T. Akutsu *et al.* (KAGRA Collaboration), Overview of KAGRA: Detector design and construction history, Prog. Theor. Exp. Phys. **2021**, 05A101 (2021).

[4] M. Punturo *et al.*, The Einstein telescope: A third-generation gravitational wave observatory, Classical Quantum Gravity **27**, 194002 (2010).

[5] David Reitze *et al.*, Cosmic Explorer: The U.S. contribution to gravitational-wave astronomy beyond LIGO, Bull. Am. Astron. Soc. **51**, 035 (2019).

[6] Pau Amaro-Seoane *et al.* (LISA Collaboration), Laser interferometer space antenna, arXiv:1702.00786.

[7] 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. Relativity **21**, 3 (2018).

[8] Benjamin P. Abbott *et al.* (LIGO Scientific Colllaboration), Exploring the sensitivity of next generation gravitational wave detectors, Classical Quantum Gravity **34**, 044001 (2017).

[9] Frans Pretorius, Evolution of Binary Black Hole Spacetimes, Phys. Rev. Lett. **95**, 121101 (2005).

[10] Manuela Campanelli, C. O. Lousto, P. Marronetti, and Y. Zlochower, Accurate Evolutions of Orbiting Black-Hole Binaries without Excision, Phys. Rev. Lett. **96**, 111101 (2006).

[11] John G. Baker, Joan Centrella, Dae-Il Choi, Michael Koppitz, and James van Meter, Gravitational Wave Extraction from an Inspiraling Configuration of Merging Black Holes, Phys. Rev. Lett. **96**, 111102 (2006).

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

[13] Serguei Ossokine *et al.*, Multipolar effective-one-body waveforms for precessing binary black holes: Construction and validation, Phys. Rev. D **102**, 044055 (2020).

[14] Sebastian Khan, Frank Ohme, Katerina Chatziioannou, and Mark Hannam, Including higher order multipoles in gravitational-wave models for precessing binary black holes, Phys. Rev. D **101**, 024056 (2020).

[15] Geraint Pratten *et al.*, Computationally efficient models for the dominant and subdominant harmonic modes of precessing binary black holes, Phys. Rev. D **103**, 104056 (2021).

[16] Héctor Estellés, Marta Colleoni, Cecilio García-Quirós, Sascha Husa, David Keitel, Maite Mateu-Lucena, Maria de Lluc Planas, and Antoni Ramos-Buades, New twists in compact binary waveform modelling: A fast time domain model for precession, Phys. Rev. D **105**, 084040 (2022).

[17] Alessandro Nagar, Gunnar Riemenschneider, Geraint Pratten, Piero Rettegno, and Francesco Messina, Multipolar effective one body waveform model for spin-aligned black hole binaries, Phys. Rev. D **102**, 024077 (2020).

[18] Sarp Akcay, Rossella Gamba, and Sebastiano Bernuzzi, Hybrid post-Newtonian effective-one-body scheme for spin-precessing compact-binary waveforms up to merger, Phys. Rev. D **103**, 024014 (2021).

[19] 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).

[20] 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).

[21] 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).

[22] 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).

[23] 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).

[24] 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).

[25] 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).

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

[27] Bhooshan Gadre, Michael Pürrer, Scott E. Field, Serguei Ossokine, and Vijay Varma, A fully precessing higher-mode surrogate model of effective-one-body waveforms, arXiv:2203.00381.

[28] 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).

[29] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GWTC-2.1: Deep extended catalog of compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run, arXiv:2108.01045.

[30] R. Abbott *et al.* (LIGO Scientific, VIRGO, and KAGRA Collaborations), GWTC-3: Compact binary coalescences observed by ligo and virgo during the second part of the third observing run, arXiv:2111.03606.

[31] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW190521: A Binary Black Hole Merger with a Total Mass of $150 M_\odot$, Phys. Rev. Lett. **125**, 101102 (2020).

[32] Vijay Varma, Maximiliano Isi, Sylvia Biscoveanu, Will M. Farr, and Salvatore Vitale, Measuring binary black hole orbital-plane spin orientations, Phys. Rev. D **105**, 024045 (2022).

[33] Vijay Varma, Sylvia Biscoveanu, Maximiliano Isi, Will M. Farr, and Salvatore Vitale, Hints of Spin-Orbit Resonances in the Binary Black Hole Population, Phys. Rev. Lett. **128**, 031101 (2022).

[34] Vijay Varma, Sylvia Biscoveanu, Tousif Islam, Feroz H. Shaik, Carl-Johan Haster, Maximiliano Isi, Will M. Farr, Scott E. Field, and Salvatore Vitale, Evidence of Large Recoil Velocity from a Black Hole Merger Signal, Phys. Rev. Lett. **128**, 191102 (2022).

[35] Vijay Varma, Maximiliano Isi, and Sylvia Biscoveanu, Extracting the Gravitational Recoil from Black Hole Merger Signals, Phys. Rev. Lett. **124**, 101104 (2020).

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

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

[38] Nigel T. Bishop and Luciano Rezzolla, Extraction of gravitational waves in numerical relativity, Living Rev. Relativity **19**, 2 (2016).

[39] Karan Jani, James Healy, James A. Clark, Lionel London, Pablo Laguna, and Deirdre Shoemaker, Georgia tech catalog of gravitational waveforms, Classical Quantum Gravity **33**, 204001 (2016).

[40] James Healy, Carlos O. Lousto, Yosef Zlochower, and Manuela Campanelli, The RIT binary black hole simulations catalog, Classical Quantum Gravity **34**, 224001 (2017).

[41] Y. B. Zel'dovich and A. G. Polnarev, Radiation of gravitational waves by a cluster of superdense stars, Sov. Astron. **18**, 17 (1974).

[42] V. B. Braginskii and Kip S. Thorne, Gravitational-wave bursts with memory and experimental prospects, Nature (London) **327**, 123 (1987).

[43] Kip S. Thorne, Gravitational-wave bursts with memory: The Christodoulou effect, Phys. Rev. D **45**, 520 (1992).

[44] Demetrios Christodoulou, Nonlinear Nature of Gravitation and Gravitational-Wave Experiments, Phys. Rev. Lett. **67**, 1486 (1991).

[45] Geoffrey Compère, Infinite Towers of Supertranslation and Superrotation Memories, Phys. Rev. Lett. **123**, 021101 (2019).

[46] Alexander M. Grant and David A. Nichols, Persistent gravitational wave observables: Curve deviation in asymptotically flat spacetimes, Phys. Rev. D **105**, 024056 (2022).

[47] Sabrina Pasterski, Andrew Strominger, and Alexander Zhiboedov, New gravitational memories, J. High Energy Phys. 12 (2016) 053.

[48] Lydia Bieri and David Garfinkle, Perturbative and gauge invariant treatment of gravitational wave memory, Phys. Rev. D **89**, 084039 (2014).

[49] M. Turner, Gravitational radiation from point-masses in unbound orbits: Newtonian results, Astrophys. J. **216**, 610 (1977).

[50] Thibault Damour, Federico Guercilena, Ian Hinder, Seth Hopper, Alessandro Nagar, and Luciano Rezzolla, Strong-field scattering of two black holes: Numerics versus analytics, Phys. Rev. D **89**, 081503 (2014).

[51] Gihyuk Cho, Achamveedu Gopakumar, Maria Haney, and Hyung Mok Lee, Gravitational waves from compact binaries in post-Newtonian accurate hyperbolic orbits, Phys. Rev. D **98**, 024039 (2018).

[52] Alessandro Nagar, Piero Rettegno, Rossella Gamba, and Sebastiano Bernuzzi, Effective-one-body waveforms from dynamical captures in black hole binaries, Phys. Rev. D **103**, 064013 (2021).

[53] Thibault Damour and Piero Rettegno, Strong-field scattering of two black holes: Numerical relativity meets post-Minkowskian gravity, Phys. Rev. D **107**, 064051 (2023).

[54] Piero Rettegno, Geraint Pratten, Lucy Thomas, Patricia Schmidt, and Thibault Damour, Strong-field scattering of two spinning black holes: Numerical relativity versus post-Minkowskian gravity, arXiv:2307.06999.

[55] Moritz Hübner, Colm Talbot, Paul D. Lasky, and Eric Thrane, Measuring gravitational-wave memory in the first LIGO/Virgo gravitational-wave transient catalog, Phys. Rev. D **101**, 023011 (2020).

[56] Oliver M. Boersma, David A. Nichols, and Patricia Schmidt, Forecasts for detecting the gravitational-wave memory effect with Advanced LIGO and Virgo, Phys. Rev. D **101**, 083026 (2020).

[57] Alexander M. Grant and David A. Nichols, Outlook for detecting the gravitational-wave displacement and spin memory effects with current and future gravitational-wave detectors, Phys. Rev. D **107**, 064056 (2023).

[58] Michael Ebersold and Shubhanshu Tiwari, Search for nonlinear memory from subsolar mass compact binary mergers, Phys. Rev. D **101**, 104041 (2020).

[59] Shubhanshu Tiwari, Michael Ebersold, and Eleanor Z. Hamilton, Leveraging gravitational-wave memory to distinguish neutron star-black hole binaries from black hole binaries, Phys. Rev. D **104**, 123024 (2021).

[60] Dixeena Lopez, Shubhanshu Tiwari, and Michael Ebersold, Memory matters: Gravitational wave memory of compact binary coalescence in the presence of matter effects, arXiv:2305.04761.

[61] Andrew Strominger and Alexander Zhiboedov, Gravitational memory, BMS supertranslations and soft theorems, J. High Energy Phys. 01 (2016) 086.

[62] Temple He, Vyacheslav Lysov, Prahar Mitra, and Andrew Strominger, BMS supertranslations and Weinberg's soft graviton theorem, J. High Energy Phys. 05 (2015) 151.

[63] Steven Weinberg, Infrared photons and gravitons, Phys. Rev. **140**, B516 (1965).

[64] Sasha Haco, Stephen W. Hawking, Malcolm J. Perry, and Andrew Strominger, Black hole entropy and soft hair, J. High Energy Phys. 12 (2018) 098.

[65] Stephen W. Hawking, Malcolm J. Perry, and Andrew Strominger, Soft Hair on Black Holes, Phys. Rev. Lett. **116**, 231301 (2016).

[66] Keefe Mitman *et al.*, Adding gravitational memory to waveform catalogs using BMS balance laws, Phys. Rev. D **103**, 024031 (2021).

[67] Colm Talbot, Eric Thrane, Paul D. Lasky, and Fuhui Lin, Gravitational-wave memory: Waveforms and phenomenology, Phys. Rev. D **98**, 064031 (2018).

[68] Nigel T. Bishop, Roberto Gomez, Luis Lehner, and Jeffrey Winicour, Cauchy-characteristic extraction in numerical relativity, Phys. Rev. D **54**, 6153 (1996).

[69] Jordan Moxon, Mark A. Scheel, and Saul A. Teukolsky, Improved Cauchy-characteristic evolution system for high-precision numerical relativity waveforms, Phys. Rev. D **102**, 044052 (2020).

[70] Jordan Moxon, Mark A. Scheel, Saul A. Teukolsky, Nils Deppe, Nils Fischer, Francois Hébert, Lawrence E. Kidder, and William Throwe, The SPECTRE Cauchy-characteristic evolution system for rapid, precise waveform extraction, Phys. Rev. D **107**, 064013 (2023).

[71] Keefe Mitman, Jordan Moxon, Mark A. Scheel, Saul A. Teukolsky, Michael Boyle, Nils Deppe, Lawrence E. Kidder, and William Throwe, Computation of displacement and spin gravitational memory in numerical relativity, Phys. Rev. D **102**, 104007 (2020).

[72] Nils Deppe *et al.*, SPECTREv2023.02.09, 10.5281/zenodo.7626579(2023).

[73] Keefe Mitman *et al.*, Fixing the BMS frame of numerical relativity waveforms with BMS charges, Phys. Rev. D **106**, 084029 (2022).

[74] Keefe Mitman *et al.*, Fixing the BMS frame of numerical relativity waveforms, Phys. Rev. D **104**, 024051 (2021).

[75] 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).

[76] 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).

[77] Collin Capano, Yi Pan, and Alessandra Buonanno, Impact of higher harmonics in searching for gravitational waves from nonspinning binary black holes, Phys. Rev. D **89**, 102003 (2014).

[78] 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).

[79] 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).

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

[81] H. Bondi, M. G. J. van der Burg, and A. W. K. Metzner, Gravitational waves in general relativity. 7. Waves from axisymmetric isolated systems, Proc. R. Soc. A **269**, 21 (1962).

[82] R. K. Sachs, Gravitational waves in general relativity. 8. Waves in asymptotically flat space-times, Proc. R. Soc. A **270**, 103 (1962).

[83] Mike Boyle, Dante Iozzo, and Leo C. Stein, moble/scri: v1.2, https://github.com/moble/scri (2020).

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

[85] Michael Boyle, Lawrence E. Kidder, Serguei Ossokine, and Harald P. Pfeiffer, Gravitational-wave modes from precessing black-hole binaries, arXiv:1409.4431.

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

[87] Mike Boyle, Dan Hemberger, and Dante Iozzo, moble/gwframes, https://github.com/moble/GWFrames (2020).

[88] Luc Blanchet, Thibault Damour, Gilles Esposito-Farese, and Bala R. Iyer, Gravitational Radiation from Inspiralling Compact Binaries Completed at the Third Post-Newtonian Order, Phys. Rev. Lett. **93**, 091101 (2004).

[89] Luc Blanchet, Gravitational radiation from post-Newtonian sources and inspiralling compact binaries, Living Rev. Relativity **17**, 2 (2014).

[90] Piotr Jaranowski and Gerhard Schäfer, Dimensional regularization of local singularities in the 4th post-Newtonian two-point-mass Hamiltonian, Phys. Rev. D **87**, 081503 (2013).

[91] Donato Bini and Thibault Damour, Analytical determination of the two-body gravitational interaction potential at the fourth post-Newtonian approximation, Phys. Rev. D **87**, 121501 (2013).

[92] Donato Bini and Thibault Damour, High-order post-Newtonian contributions to the two-body gravitational interaction potential from analytical gravitational self-force calculations, Phys. Rev. D **89**, 064063 (2014).

[93] Lawrence E. Kidder, Coalescing binary systems of compact objects to postNewtonian 5/2 order. 5. Spin effects, Phys. Rev. D **52**, 821 (1995).

[94] Clifford M. Will and Alan G. Wiseman, Gravitational radiation from compact binary systems: Gravitational wave forms and energy loss to second post-Newtonian order, Phys. Rev. D **54**, 4813 (1996).

[95] Alejandro Bohe, Sylvain Marsat, Guillaume Faye, and Luc Blanchet, Next-to-next-to-leading order spin-orbit effects in the near-zone metric and precession equations of compact binaries, Classical Quantum Gravity **30**, 075017 (2013).

[96] Luc Blanchet, Guillaume Faye, Bala R. Iyer, and Siddhartha Sinha, The third post-Newtonian gravitational wave polarisations and associated spherical harmonic modes for inspiralling compact binaries in quasi-circular orbits, Classical Quantum Gravity **25**, 165003 (2008); **29**, 239501(E) (2012).

[97] Guillaume Faye, Sylvain Marsat, Luc Blanchet, and Bala R. Iyer, The third and a half post-Newtonian gravitational wave quadrupole mode for quasi-circular inspiralling compact binaries, Classical Quantum Gravity **29**, 175004 (2012).

[98] Guillaume Faye, Luc Blanchet, and Bala R. Iyer, Non-linear multipole interactions and gravitational-wave octupole modes for inspiralling compact binaries to third-and-a-half post-Newtonian order, Classical Quantum Gravity **32**, 045016 (2015).

[99] Marc Favata, Post-Newtonian corrections to the gravitational-wave memory for quasi-circular, inspiralling compact binaries, Phys. Rev. D **80**, 024002 (2009).

[100] 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).

[101] 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).

[102] Caleb Devine, Zachariah B. Etienne, and Sean T. McWilliams, Optimizing spinning time-domain gravitational waveforms for Advanced LIGO data analysis, Classical Quantum Gravity **33**, 125025 (2016).

[103] Thibault Damour, Bala R. Iyer, and B. S. Sathyaprakash, A comparison of search templates for gravitational waves from binary inspiral, Phys. Rev. D **63**, 044023 (2001); **72**, 029902(E) (2005).

[104] Scott E. Field, Chad R. Galley, Frank Herrmann, Jan S. Hesthaven, Evan Ochsner, and Manuel Tiglio, Reduced Basis Catalogs for Gravitational Wave Templates, Phys. Rev. Lett. **106**, 221102 (2011).

[105] 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).

[106] Peter Binev, Albert Cohen, Wolfgang Dahmen, Ronald DeVore, Guergana Petrova, and Przemyslaw Wojtaszczyk, Convergence rates for greedy algorithms in reduced basis methods, SIAM J. Math. Anal. **43**, 1457 (2011).

[107] Ronald DeVore, Guergana Petrova, and Przemyslaw Wojtaszczyk, Greedy algorithms for reduced bases in banach spaces, Constr. Approx. **37**, 455 (2013).

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

[109] 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).

[110] Jan S. Hesthaven, Benjamin Stamm, and Shun Zhang, Efficient greedy algorithms for high-dimensional parameter spaces with applications to empirical interpolation and reduced basis methods, ESAIM: M2AN **48**, 259 (2014).

[111] Harbir Antil, Scott E. Field, Frank Herrmann, Ricardo H. Nochetto, and Manuel Tiglio, Two-step greedy algorithm for reduced order quadratures, J. Sci. Comput. **57**, 604 (2013).

[112] 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).

[113] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake VanderPlas, Alexandre Passos, David Cournapeau,

Matthieu Brucher, Matthieu Perrot, and Edouard Duchesnay, SCIKIT-LEARNMachine learning in python, J. Mach. Learn. Res. **12**, 2825 (2011).

[114] 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).

[115] 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).

[116] Curt Cutler and Eanna 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).

[117] Eric Poisson and Clifford M. Will, Gravitational waves from inspiraling compact binaries: Parameter estimation using second postNewtonian wave forms, Phys. Rev. D **52**, 848 (1995).

[118] 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).

[119] Yitian Chen, Michael Boyle, and Saul A. Teukolsky, Improved frequency spectra of gravitational waves with memory in a binary-black-hole simulation (to be published).

[120] LIGO Scientific Collaboration, Updated Advanced LIGO sensitivity design curve, Technical Report No. T1800044-v5, 2018, https://dcc.ligo.org/LIGO-T1800044/public.

[121] Tony Chu, Heather Fong, Prayush Kumar, Harald P. Pfeiffer, Michael Boyle, Daniel A. Hemberger, Lawrence E. Kidder, Mark A. Scheel, and Bela Szilagyi, On the accuracy and precision of numerical waveforms: Effect of waveform extraction methodology, Classical Quantum Gravity **33**, 165001 (2016).

[122] Tousif Islam, Scott E Field, Gaurav Khanna, and Niels Warburton, Survey of gravitational wave memory in intermediate mass ratio binaries, Phys. Rev. D **108**, 024046 (2023).

[123] Silvia Gasparotto, Rodrigo Vicente, Diego Blas, Alexander C. Jenkins, and Enrico Barausse, Can gravitational-wave memory help constrain binary black-hole parameters? A LISA case study, Phys. Rev. D **107**, 124033 (2023).

[124] Laxmikant Kale *et al.*, Uiuc-ppl/charm: Charm++ version 6.10.2, 10.5281/zenodo.3972617 (2020).

[125] Pauli Virtanen *et al.*, SciPy 1.0 Contributors, SciPy 1.0: Fundamental algorithms for scientific computing in Python, Nat. Methods **17**, 261 (2020).

[126] Charles R. Harris *et al.*, Array programming with NumPy, Nature (London) **585**, 357 (2020).

[127] J. D. Hunter, Matplotlib: A 2d graphics environment, Comput. Sci. Eng. **9**, 90 (2007).

[128] Michael L. Waskom, seaborn: Statistical data visualization, J. Open Source Softwaare **6**, 3021 (2021).

[129] Lorena Magaña Zertuche *et al.*, High precision ringdown modeling: Multimode fits and BMS frames, Phys. Rev. D **105**, 104015 (2022).