

# Effective-one-body waveforms for binary neutron stars using surrogate models

Benjamin D. Lackey<sup>1</sup>, Sebastiano Bernuzzi<sup>2,3</sup>, Chad R. Galley<sup>3</sup>, Jeroen Meidam<sup>4</sup>, Chris Van Den Broeck<sup>4</sup>

<sup>1</sup>*Department of Physics, Syracuse University, Syracuse, NY 13244, USA*

<sup>2</sup>*DiFeST, University of Parma, and INFN, I-43124, Parma, Italy*

<sup>3</sup>*Theoretical Astrophysics, Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, California 91125, USA*

<sup>4</sup>*Nikhef – National Institute for Subatomic Physics,*

*Science Park 105, 1098 XG Amsterdam, The Netherlands*

(Dated: October 18, 2016)

Gravitational-wave observations of binary neutron star systems can provide information about the masses, spins, and structure of neutron stars. However, this requires accurate and computationally efficient waveform models that take  $\lesssim 1$  s to evaluate for use in Bayesian parameter estimation codes that perform  $10^7 - 10^8$  waveform evaluations. We present a surrogate model of a nonspinning effective-one-body waveform model with  $\ell = 2, 3$ , and 4 tidal multipole moments that reproduces waveforms of binary neutron star numerical simulations up to merger. The surrogate is built from compact sets of effective-one-body waveform amplitude and phase data that each form a reduced basis. We find that 12 amplitude and 7 phase basis elements are sufficient to reconstruct any binary neutron star waveform with a starting frequency of 10 Hz. The surrogate has maximum errors of 3.8% in amplitude (0.04% excluding the last 100M before merger) and 0.043 radians in phase. The version implemented in the LIGO Algorithm Library takes  $\sim 0.07$  s to evaluate for a starting frequency of 30 Hz and  $\sim 0.8$  s for a starting frequency of 10 Hz, resulting in a speed-up factor of  $\sim 10^3 - 10^4$  relative to the original Matlab code. This allows parameter estimation codes to run in days to weeks rather than years, and we demonstrate this with a Nested Sampling run that recovers the masses and tidal parameters of a simulated binary neutron star system.

## I. INTRODUCTION

One of the primary targets for gravitational-wave detectors such as Advanced LIGO (aLIGO) [1], Advanced Virgo [2], KAGRA [3], and LIGO-India [4] is the inspiral of binary neutron star (BNS) systems. The evolution of the waveform provides detailed information about the masses and spins of the two neutron stars (NSs) as well as information about the NS structure and equation of state (EOS) encoded in the tidal interactions of the two NSs.

Measuring the parameters of the binary, however, requires waveform models that are both fast, for use in Bayesian parameter estimation codes, and accurate, to minimize systematic errors in the recovered parameters. Almost all previous Bayesian parameter-estimation studies of BNS systems [5–9] have used post-Newtonian (PN) waveform models [10, 11]. These models typically take  $\sim 1$  s or less to evaluate, and are therefore suitable for Markov-Chain-Monte-Carlo or Nested-Sampling codes (see [12] for a review) that require  $10^6 - 10^7$  waveform evaluations. Unfortunately, because the PN expansion is only known completely to 3.5PN order, the uncertainty in the waveform phase can be greater than  $\sim 10$  radians by the time the binary merges [13]. This can lead to significant biases in the measured masses and tidal interactions [6, 14, 15].

The effective one body (EOB) formalism first introduced in Ref. [16] provides an alternative that includes several effects beyond the standard PN expansion (see Ref. [17] for a review) and can be calibrated with numerical relativity binary black hole (BBH) simulations near merger [18–22]. Most implementations of EOB

waveforms, however, are significantly slower than for PN waveforms, sometimes taking tens of minutes to generate a single waveform, and this is unusably slow for most parameter estimation algorithms. Recent work on optimizing EOB waveform generation for BBH has resulted in a significant speed-up [23], but this optimization must be done for each new waveform model and will not work for numerical relativity simulations.

Reduced-order modeling (ROM) techniques provide a framework for reducing large data sets that can be used to build lightweight models that are rapidly evaluated as a substitute, or surrogate, in place of the slow waveform generation code. The method, introduced in [24], begins with a training set of waveforms that covers the waveform parameter space. A greedy algorithm [25–28] exposes the most relevant waveforms needed to accurately represent the full training set [29]. These relatively few number of judiciously chosen waveforms, also called a reduced basis, captures the dependence of the training set waveforms on parameters.

Waveforms for arbitrary parameter values (in the training region) can then be generated from the reduced basis by estimating the parametric dependence of the projection coefficients. This can be done in two ways. In one case, each coefficient is interpolated as a function of waveform parameters using the training data [30, 31]. In the other case, one uses the empirical interpolation method [24, 32, 33] to build an interpolant that is customized to the waveform data such that at a relatively few specific times (i.e., the interpolation nodes) one fits for the parametric variation of the waveform data [24]. This second approach is compact, robust to round-off noise, and allows for the intrinsic waveform errors to be

incorporated in the error of the final surrogate model (e.g., see [34]).

Reduced-order modeling in gravitational wave physics started with an observation that the inspiral dynamics of precessing BBHs can be dimensionally reduced, meaning that many configurations share similar, almost redundant, qualities that vary smoothly across parameter space [35]. The result implied that the multi-dimensional parameter space for precession waveforms might effectively be considerably smaller thus providing a possible avenue towards beating the “curse of dimensionality” for template bank coverage. At the same time, compression factors  $\sim 10$  of a small template bank of non-spinning gravitational waveforms were achieved using a singular value decomposition [36]. Subsequently, reduced-order modeling techniques have been used to efficiently represent/compress large waveform banks [29, 37–40] and to build fast and accurate surrogate models [24] of merger waveforms [24, 30, 31, 34, 41], which can be used in multiple-query applications like parameter estimation studies [42–44] that use reduced-order quadratures [45]. In fact, reduced-order models are crucial in modern GW search pipelines [46] and in parameter estimation studies to accelerate waveform generation and likelihood computations.

In this work, we construct a reduced-order surrogate model (or “surrogate”) for the  $\ell = m = 2$  mode of BNS waveforms generated with the EOB formalism. This EOB model, described in Ref. [47], incorporates tidal interactions that are parameterized by the quadrupolar  $\ell = 2$  tidal deformability  $\Lambda_2$  of each star as well as the  $\ell = 3$  and 4 tidal deformabilities  $\Lambda_3$  and  $\Lambda_4$ , respectively. These tidal interactions enter at the 5th, 7th, and 9th PN orders, respectively in a resummed form, and lead to an accumulating phase shift of  $\sim 1$  radian up to a gravitational-wave frequency of 400Hz and  $\sim 10$  radians up to the BNS merger frequency as shown in Fig. 1 below. (An alternative model that includes tidally excited resonances has recently been developed [48].) We construct separate reduced bases for the amplitude and phase, and find them to be extremely compact; 12 amplitude bases and 7 phase bases are sufficient to accurately reproduce any waveform in the training set. We then interpolate the amplitude and phase as a function of waveform parameters at the times chosen by the empirical interpolation method using Chebyshev interpolation.

Because EOB models with both tidal interactions and spin are just starting to become available, our surrogate only applies to nonspinning BNS systems. In future work we intend to incorporate NS spins once they are available in the EOB models for tidally interacting systems. We note that inspiraling BNS systems are not likely to have significant spins. The fastest known NS in a confirmed BNS system has a spin frequency of 44 Hz [49], corresponding to a dimensionless spin of  $\sim 0.04$ . Another potential BNS system has a NS with a spin frequency of 239 Hz [50], corresponding to a dimensionless spin of  $\sim 0.2$ . However, even a spin of  $\sim 0.03$  can lead to a sys-

tematic bias in the estimated tidal parameters that are as large as the statistical errors if not incorporated into the waveform model. [14, 15].

We organize the paper as follows. In Section II, we summarize the EOB model for BNS systems from which we construct the reduced-order surrogate model. We describe the steps to build the surrogate in Section III, and present its accuracy and speed for predicting EOB waveforms at new parameter values in Section IV. Finally, we summarize our results and discuss future work in Section V. In the Appendix, we describe the accuracy of approximating the  $\ell = 3$  and 4 tidal interactions in terms of the  $\ell = 2$  tidal interaction.

*Conventions:* Unless explicitly stated, we use units where  $G = c = 1$ .

## II. TIDAL EOB WAVEFORM MODEL

### A. TEOB<sub>Resum</sub>

In this work, we use the tidal EOB (TEOB) model developed in [47] and called TEOB<sub>Resum</sub>. TEOB<sub>Resum</sub> incorporates an enhanced (resummed) attractive tidal potential derived from recent analytical advances in the PN and gravitational self-force description of relativistic tidal interactions [51, 52]. The resummed tidal potential of TEOB<sub>Resum</sub> significantly improves the description of tidal interactions near the merger over the previous next-next-to-leading-order (NNLO) TEOB model [53, 54] and over the conventional PN models. In particular, TEOB<sub>Resum</sub> predicts high-resolution, multi-orbit numerical relativity results within their uncertainties and without fitting parameters to BNS numerical waveforms [47].

The main features of TEOB<sub>Resum</sub> are summarised in what follows. The Hamiltonian is  $H_{\text{EOB}} = M\sqrt{1 + 2\nu(\hat{H}_{\text{eff}} - 1)}$  with

$$\hat{H}_{\text{eff}}(u, p_{r_*}, p_\varphi) \equiv \sqrt{A(u; \nu) (1 + p_\varphi^2 u^2 + 2\nu(4 - 3\nu)u^2 p_{r_*}^4) + p_{r_*}^2}, \quad (1)$$

where the binary mass is  $M = M_A + M_B$ , the symmetric mass-ratio is  $\nu = M_A M_B / M^2$ ,  $u \equiv 1/r$ ,  $r$  is the EOB radial coordinate, and  $p_\varphi$  and  $p_{r_*}$  are the conjugate momenta (see e.g. [53]). The EOB potential

$$A(u; \nu) \equiv A_0(u; \nu) + A_T(u; \nu), \quad (2)$$

is the sum of a point-mass term and a tidal term.  $A_0(u; \nu)$  is defined as the (1, 5) Padé approximant of the formal 5PN expression  $A_0^{5\text{PN}}(u; \nu) = 1 - 2u + a_3 u^3 + a_4 u^4 + (a_5^c(\nu) + a_5^{\text{ln}} \ln u)u^5 + (a_6^c(\nu) + a_6^{\text{ln}} \ln u)u^6$ . The coefficients up to 4PN, i.e.  $(a_3, a_4, a_5^c(\nu), a_5^{\text{ln}})$ , are analytically known [55]. Although the 5PN term  $a_6^{\text{ln}}$  and the linear-in- $\nu$  part of  $a_6^c(\nu)$  are analytically known [56, 57], we use the value  $a_6^c(\nu) = 3097.3\nu^2 - 1330.6\nu + 81.38$  fit to NR

data in [58]. The tidal term,

$$A_T^{(+)}(u; \nu) \equiv - \sum_{\ell=2}^4 \left[ \kappa_\ell^A u^{2\ell+2} \hat{A}_A^{(\ell+)} + (A \leftrightarrow B) \right], \quad (3)$$

models the gravito-electric sector of the interaction [59], where

$$\kappa_\ell^A = 2 \frac{M_B M_A^{2\ell}}{M^{2\ell+1}} \frac{k_\ell^A}{C_A^{2\ell+1}} = 2Q^{-1} \left( \frac{X_A}{C_A} \right)^{2\ell+1} k_\ell^A, \quad (4a)$$

$$\kappa_\ell^B = 2 \frac{M_A M_B^{2\ell}}{M^{2\ell+1}} \frac{k_\ell^B}{C_B^{2\ell+1}} = 2Q \left( \frac{X_B}{C_B} \right)^{2\ell+1} k_\ell^B, \quad (4b)$$

are the  $\ell = 2, 3, 4$  tidal polarizability parameters (or tidal coupling constants) [51]. Labels  $A, B$  refer to the stars,  $M_A$  is the gravitational mass of star  $A$ ,  $R_A$  the areal radius,  $C_A = M_A/R_A$ ,  $X_A = M_A/M$ , and  $k_\ell^A$  are the dimensionless Love numbers [53, 60–62]. The expressions above assume  $M_A \geq M_B$ , so that  $Q = M_A/M_B \geq 1$ . In the equal-mass case, the tidal interaction and EOS information are fully encoded at leading order (LO) in the total dimensionless *quadrupolar* tidal coupling constant

$$\kappa_2^T \equiv \kappa_A^{(2)} + \kappa_B^{(2)}. \quad (5)$$

The relativistic correction factors  $\hat{A}_A^{(\ell+)}$  formally include all the high PN corrections to the leading-order. The particular choice of  $\hat{A}_A^{(\ell+)}$  defines the TEOB model considered in this paper. The PN-expanded NNLO, fractionally 2PN accurate, expression is

$$\hat{A}_A^{(\ell+)}(u) = 1 + \alpha_1^{(\ell)} u + \alpha_2^{(\ell)} u^2 \quad [\text{NNLO}], \quad (6)$$

with  $\alpha_{1,2}^{(2),(3)} \neq 0$  computed analytically and  $\alpha_{1,2}^{(4)} = 0$  [63]. This TEOB<sub>NNLO</sub> model has been compared against NR simulations in [47, 54], significant deviations are observed at dimensionless GW frequencies  $M\omega_{22} \gtrsim 0.8$ , i.e. after contact and during the last 2-3 orbits to merger. The TEOB<sub>Resum</sub> model is defined from TEOB<sub>NNLO</sub> by substituting the  $\ell = 2$  term in (6) with the expression

$$\hat{A}_A^{(2+)}(u) = 1 + \frac{3u^2}{1 - r_{\text{LR}}u} + \frac{X_A \tilde{A}_1^{(2+)} \text{1SF}}{(1 - r_{\text{LR}}u)^{7/2}} + \frac{X_A^2 \tilde{A}_2^{(2+)} \text{2SF}}{(1 - r_{\text{LR}}u)^p}, \quad (7)$$

where the functions  $\tilde{A}_1^{(2+)} \text{1SF}(u)$  and  $\tilde{A}_2^{(2+)} \text{2SF}(u)$  are given in [52] and  $p = 4$ . The key idea of TEOB<sub>Resum</sub> is to use as pole location in Eq. (7) the light ring  $r_{\text{LR}}(\nu; \kappa_A^{(\ell)})$  of the TEOB<sub>NNLO</sub> model, i.e., the location of the maximum of  $A^{\text{NNLO}}(r; \nu; \kappa_A^{(\ell)})/r^2$ . TEOB<sub>Resum</sub> is completed with a resummed waveform [64] that includes the NLO tidal contributions computed in [51, 65, 66].

A black hole limit of TEOB<sub>Resum</sub> is given by setting  $\kappa_{A,B}^{(\ell)} \rightarrow 0$ . Waveforms obtained this way, however, do not accurately represent BBH ones because the model

does not include next-to-quasicircular corrections tuned to BBH NR data (so it actually differs from the model of [58]). For this reason we exclude the  $\kappa_{A,B}^{(\ell)} = 0$  configurations from the surrogate model. That is not a serious limitation because BBH waveform models are independently available, and because BBH sources are not expected in the mass range covered by our surrogate. Additionally, including the correct BBH limit would introduce a discontinuity in the waveform's parameters space that would affect the overall accuracy of the surrogate. In a similar way, configurations with  $\kappa_B^{(\ell)} \rightarrow 0$  approximate black hole-neutron star binaries, but with an astrophysically unexpected small mass ratio  $Q$ .

The TEOB<sub>Resum</sub> waveform model is determined by seven input parameters (7D parameter space): the binary mass-ratio  $Q$  and the  $\ell = 2, 3, 4$  tidal polarizability parameters (or tidal coupling constants)  $\kappa_\ell^{A,B}$ . The latter are linked to the usual multipolar dimensionless tidal parameters, e.g. [51, 67]. For each star we define

$$\Lambda_\ell^A = \frac{2k_\ell^A}{C_A^{2\ell+1} (2\ell - 1)!}, \quad (8)$$

which are proportional to the  $Q$ -independent part of  $\kappa_\ell^A$ , and correspond to the multipolar quantities called  $\lambda_\ell$  in [67].

## B. Approximation of higher order tidal effects

In constructing a surrogate, it is extremely important to reduce the dimensionality of the parameter space as much as possible in order to avoid high-dimensional interpolation which is often inaccurate and computationally demanding. Fortunately, Yagi has found tight correlations between the  $\ell = 2$  tidal parameter and the  $\ell = 3$  and  $\ell = 4$  tidal parameters that are nearly independent of the choice of EOS for plausible NS EOS models [67]. Yagi then constructed fits  $\Lambda_3^{\text{fit}}(\Lambda_2)$  and  $\Lambda_4^{\text{fit}}(\Lambda_2)$  for the  $\ell = 3$  and 4 tidal parameters in terms of the  $\ell = 2$  tidal parameter. This reduces the 7D parameter space

$$(q, \Lambda_2^A, \Lambda_2^B, \Lambda_3^A, \Lambda_3^B, \Lambda_4^A, \Lambda_4^B), \quad (9)$$

where  $q = M_B/M_A \leq 1$  to the 3D parameter space

$$(q, \Lambda_2^A, \Lambda_2^B). \quad (10)$$

We evaluate the systematic uncertainty from these fits for 14 different EOS and for NS masses in the range  $M \in [0.9M_\odot, M_{\text{max}}]$ , where  $M_{\text{max}}$  is the maximum mass. For these EOS and masses, we find the  $\ell = 3$  fit  $\Lambda_3^{\text{fit}}(\Lambda_2)$  results in fractional errors in the range  $-0.098 \leq \Delta\Lambda_3/\Lambda_3 \leq 0.17$  with an average absolute fractional error of  $\langle |\Delta\Lambda_3/\Lambda_3| \rangle = 0.04$ . The  $\ell = 4$  fit  $\Lambda_4^{\text{fit}}(\Lambda_2)$  results in fractional errors in the range  $-0.22 \leq \Delta\Lambda_4/\Lambda_4 \leq 0.28$  with an average absolute fractional error of  $\langle |\Delta\Lambda_4/\Lambda_4| \rangle = 0.08$ . Details are given in the Appendix.

In Fig. 1, we show the contribution of each tidal effect to the phase evolution of the waveform for both the soft EOS SLY [68] and the stiff EOS MS1b [69]. (See [70] for the naming convention.) The phase shift due to the tidal interactions is usually  $\sim 10$  radian for the  $\ell = 2$  interaction,  $\sim 1$  radian for the  $\ell = 3$  interaction, and  $\sim 0.1$  radian for the  $\ell = 4$  interaction for frequencies up to the maximum amplitude. Also shown is the typical error expected in the phase that results from using the fits  $\Lambda_3^{\text{fit}}(\Lambda_2)$  and  $\Lambda_4^{\text{fit}}(\Lambda_2)$  instead of the true values of  $\Lambda_3$  and  $\Lambda_4$  determined by the EOS. This error is smaller than the  $\ell = 4$  tidal effect with an overall error of  $\lesssim 0.01$  radian.

### III. BUILDING THE TIDAL EOB WAVEFORM SURROGATE

TEOB<sub>Resum</sub> is implemented as a publicly available Matlab code available to download at [71]. As mentioned above, the time it takes for this code to generate a typical waveform in a ground-based gravitational wave detector's frequency band is about twenty minutes. Unfortunately, this computational time is far too long for the waveform generation code to be used in practical gravitational wave data analysis applications. A solution to this problem is provided by reduced-order surrogate modeling, which produces a fast-to-evaluate and compact model that can be used as a substitute for the original TEOB<sub>Resum</sub> code implementation with negligible loss in accuracy.

The recipe for building a reduced-order surrogate model, or surrogate for short, was introduced in [24] to which the reader is referred to for further details. There are four main steps in the surrogate building process:

0. Precondition the set of precomputed waveforms to vary as smoothly as possible with parameters. This often results in a very compact surrogate model while also improving the accuracy of the surrogate's predictions, in the end;
1. Build a reduced basis from a set of precomputed waveforms. This results in a compression in parameter space;
2. Build an empirical interpolant from the reduced basis. This results in a dual compression of the data in the time (or frequency) dimension;
3. Estimate or fit for the parametric dependence of the waveform data at specific values of the time (or frequency) samples in the data.

We discuss the details of these steps for building a surrogate for TEOB<sub>Resum</sub> waveforms in the following subsections. However, the set of precomputed waveforms (called a training set) needed for building a reduced basis representation in Step 1 often requires preconditioning the data so that the resulting surrogate model will be as compact and accurate as possible. This preconditioning

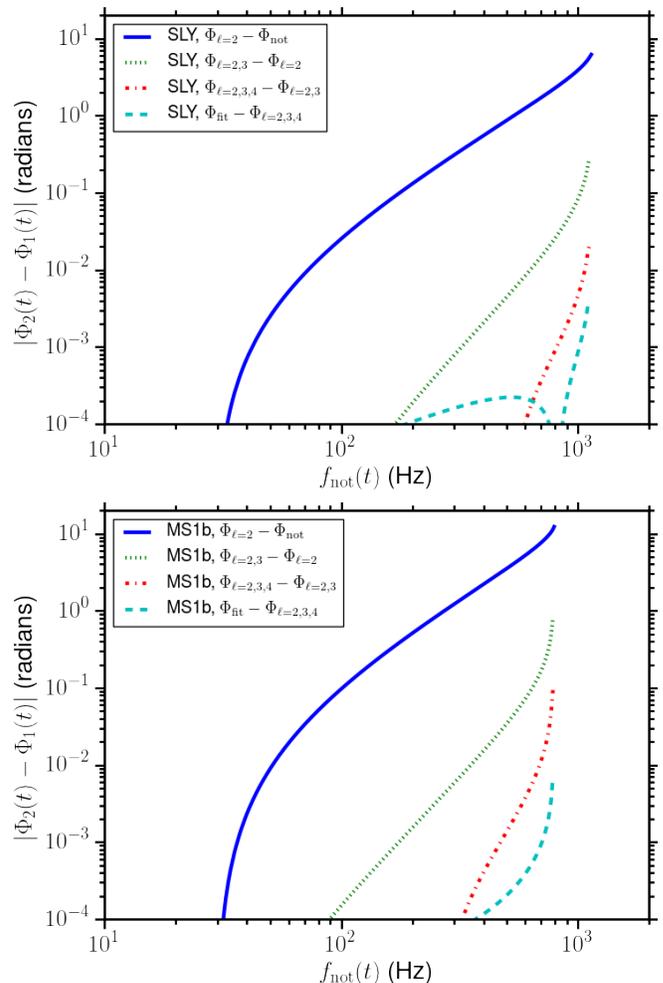


FIG. 1. Contribution of each tidal multipole moment to the phase evolution of an equal mass BNS system with component masses  $(M_A, M_B) = (1.4, 1.4) M_\odot$  beginning at 30 Hz for the soft EOS SLY (top) and the stiff EOS MS1b (bottom). The phase contribution is given by the difference in phase between waveforms with no tidal interactions  $\Phi_{\text{not}}(t)$ , only the  $\ell = 2$  interaction  $\Phi_{\ell=2}(t)$ , the  $\ell = 2, 3$  interactions  $\Phi_{\ell=2,3}(t)$ , and the  $\ell = 2, 3, 4$  interactions  $\Phi_{\ell=2,3,4}(t)$ . Also shown by the dashed curve is the error that results from using the fitting functions  $\Lambda_3^{\text{fit}}(\Lambda_2)$  and  $\Lambda_4^{\text{fit}}(\Lambda_2)$  instead of the values of  $\Lambda_3$  and  $\Lambda_4$  calculated from the EOS  $\Phi_{\text{fit}}(t)$ . Each curve is plotted as a parametric function of the phase difference between the two waveforms  $|\Phi_2(t) - \Phi_1(t)|$  versus the frequency of the waveform with no tidal interactions  $f_{\text{not}}(t)$ . In this way, the phase difference between waveforms is calculated at the same time instead of the same frequency. This can be more directly compared to phase errors in the surrogate model below which are calculated as a function of time.

step, which can be thought of as the zeroeth step in surrogate building, often involves several choices that must be made in advance, sometimes with input and foresight of steps further down the surrogate building process. For example, the size of the reduced basis generated in Step 1 depends crucially on the features and morphologies of the

training set waveforms, which can be minimized through the choices made in preconditioning the training data. As such, building a surrogate may involve a few iterations to converge to the particular strategy and set of choices that end up being suitable for achieving the desired evaluation speed and accuracy. In the next subsection, we discuss how the training data of  $\text{TEOB}_{\text{Resum}}$  waveforms was generated as well as the choices we made for preconditioning the data.

### A. Step 0: Training set and preconditioning

When constructing a training set, significant care is required in choosing the waveform parameters as the choice made can impact the accuracy (and sometimes the ability) to accurately estimate or fit the waveform data in Step 3 above. We work in units for the mass parameters where the waveform amplitude and time samples are rescaled by the total mass  $M = M_A + M_B$  so that a mass-ratio parameter is the only mass parameter needed for the surrogate. Three common choices are  $q = M_B/M_A \leq 1$ ,  $Q = M_A/M_B \geq 1$ , and the symmetric mass ratio  $\nu = M_A M_B / M^2 \leq 1/4$ . We have found that the symmetric mass ratio  $\nu$  is a poor choice for  $\text{TEOB}_{\text{Resum}}$  waveforms since the amplitude and phase at fixed times can sometimes change rapidly between grid points in the training set of parameters, which will make fitting for the parametric variation (see Step 3 above) difficult. The amplitude and phase vary less rapidly as functions of  $q$  and  $Q$ , and we find that  $q$  provides slightly better accuracy of the final surrogate model. We choose  $q \in [0.5, 1]$  as the mass ratio parameter.

For the tidal parameters, we use a rectangular grid of  $\Lambda_2^A \in [50, 5000]$  for the more massive NS and  $\Lambda_2^B \in [50, 5000]$  for the less massive NS. For any realistic EOS,  $\Lambda_2^B \geq \Lambda_2^A$ , so only a triangular half of this rectangular grid is physically plausible. However, because most implementations of accurate interpolation algorithms for multi-dimensional data require rectangular grids, we will sample the entire rectangular grid. Other choices for the two tidal parameters include  $\kappa_2^A$  and  $\kappa_2^B$  (defined in Eq. (4)), as well as  $\tilde{\Lambda}$  and  $\delta\tilde{\Lambda}$ , which are another linear combination of  $\Lambda_2^A$  and  $\Lambda_2^B$  used in parameter estimation [6, 14]. Both of these alternative choices suffer from the same problem. In particular, mapping a rectangular grid of  $\{\kappa_2^A, \kappa_2^B\}$  or  $\{\tilde{\Lambda}, \delta\tilde{\Lambda}\}$  to the corresponding values of  $\{\Lambda_2^A, \Lambda_2^B\}$  can take  $\Lambda_2$  outside the domain where the fits  $\Lambda_3^{\text{fit}}(\Lambda_2)$  and  $\Lambda_4^{\text{fit}}(\Lambda_2)$  (Eq. (A1)) are valid. This makes it impossible to evaluate the  $\ell = 3, 4$  tidal parameters.

With this choice of parameters,

$$\boldsymbol{\theta} := (q, \Lambda_2^A, \Lambda_2^B), \quad (11)$$

we next choose a discretization of the parameter space to define the training set that will be used in Step 1. As discussed in Sec. III D below, we will use Chebyshev interpolation to fit for the variation of the amplitude and phase

in terms of the waveform parameters  $\boldsymbol{\theta}$ . So, in constructing our training set, we choose waveform parameters at Chebyshev-Gauss-Lobatto nodes [72]. For each parameter, after linearly rescaling the range to  $x \in [-1, 1]$ , the location of the  $M$  nodes are given by

$$x_k = -\cos\left(\frac{k\pi}{M-1}\right), \quad (12)$$

where  $k = 0, \dots, M-1$ . We have found that a grid of  $16 \times 16 \times 16 = 4096$  parameters, shown in Fig. 2, is sufficient to reach the desired accuracy of the final model. These  $N = 4096$  points define our training set of parameters,  $\mathcal{T}_N := \{\boldsymbol{\theta}_i\}_{i=1}^N$ , that is used in building our surrogate model. This grid is more densely sampled at the edges of the parameter space, which is convenient because the algorithm we use for constructing the reduced basis (see Step 1) tends to choose parameters near the boundary of the training set. We then run the  $\text{TEOB}_{\text{Resum}}$  code to generate 4096 waveforms at these training set parameter values with a starting frequency that is less than 10 Hz for any combination of parameters with  $M_A \geq M_B \geq 1M_\odot$ , which corresponds to a length of  $\sim 2 \times 10^8 M$  in dimensionless units. This collection of waveforms constitutes our training data.

As discussed in Appendix E of Ref. [24], it is very helpful that the training set waveforms be accurately aligned at maximum amplitude. Otherwise, the amplitude and phase will not be smooth functions of the waveform parameters and the resulting surrogate model may not have a compact size, which automatically results in a loss of evaluation speed. We do this alignment by (i) densely sampling a waveform near the time of maximum amplitude, (ii) interpolating the amplitude with cubic splines, (iii) numerically finding the maximum of the interpolated amplitude, and (iv) shifting the waveform such that  $t = 0$  corresponds to the maximum amplitude. We then set the phase to zero at the common starting time for the shifted waveform. This alignment procedure is performed for all waveforms in the training data set.

Finally, we resample all of the training data waveforms to reduce the physical memory storage requirements. Before  $t = -10^3 M$ , we sample the amplitude and phase uniformly in phase with a spacing of  $\Delta\Phi = \pi$ . After this time we use a uniform in time sampling with a spacing of  $\Delta t = 0.1M$  to capture the more complicated behavior near merger. This leads to waveforms that only require  $\sim 7 \times 10^4$  samples compared to  $\gtrsim 10^8$  samples if we had sampled uniformly in time with sufficient accuracy to capture the behavior near merger. Our nonuniform downsampling allows us to store the entire waveform training data in  $\sim 7\text{GB}$  instead of many TB. More elaborate downsampling strategies exist including one that uses theoretically derived bounds on cubic spline errors to estimate the largest spacing between data points [30] and one that uses a greedy algorithm to select only those data points that are sufficient to recover the full data set up to a requested accuracy by a spline of a given,

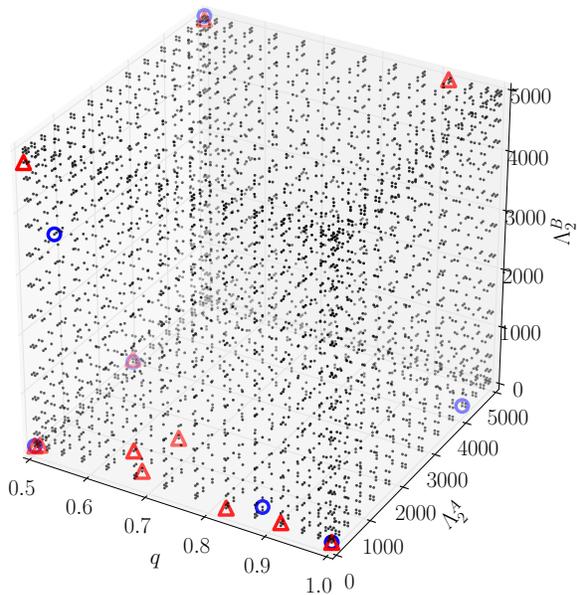


FIG. 2. The training set is constructed from the Chebyshev-Gauss-Lobatto nodes with 16 nodes in each dimension for a total of 4096 waveforms. The waveform parameters have the range  $q \in [0.5, 1]$ ,  $\Lambda_2^A \in [50, 5000]$ , and  $\Lambda_2^B \in [50, 5000]$ . The same grid is also used for the Chebyshev interpolation to evaluate the amplitude and phase at the empirical nodes  $\tau_j$ . Red  $\Delta$ 's represent the 12 waveforms chosen by the greedy algorithm to generate the amplitude reduced basis, while blue  $\circ$ 's represent the 7 waveforms chosen for the phase reduced basis.

arbitrary degree [73, 74].

### B. Step 1: Reduced basis

---

#### Algorithm 1 Greedy algorithm for reduced basis

---

- 1: **Input:**  $\{\theta_i, X(\cdot; \theta_i)\}_{i=1}^N, \epsilon$
  - 2: Set  $i = 0$  and define  $\sigma_0 = 1$
  - 3: **Seed choice** (arbitrary):  $\mu_1 \in \mathcal{T}, e_1 = X(\cdot; \mu_1)$
  - 4:  $\text{RB} = \{e_1\}$
  - 5: **while**  $\sigma_i \geq \epsilon$  **do**
  - 6:    $i = i + 1$
  - 7:    $\sigma_i = \max_{\theta \in \mathcal{T}} \|X(\cdot; \theta) - \mathcal{P}_i X(\cdot; \theta)\|^2$
  - 8:    $\mu_{i+1} = \arg\max_{\theta \in \mathcal{T}} \|X(\cdot; \theta) - \mathcal{P}_i X(\cdot; \theta)\|^2$
  - 9:    $e_{i+1} = X(\cdot; \mu_{i+1}) - \mathcal{P}_i X(\cdot; \mu_{i+1})$  (Gram-Schmidt)
  - 10:    $e_{i+1} = e_{i+1} / \|e_{i+1}\|$  (normalization)
  - 11:    $\text{RB} = \text{RB} \cup e_{i+1}$
  - 12: **end while**
  - 13: Set  $n = i$
  - 14: **Output:**  $\text{RB} = \{e_i\}_{i=1}^n$  and greedy points  $\{\mu_i\}_{i=1}^n$
- 

With the training set of waveforms in hand we now focus on reducing the data to its essential components in both parameters (this subsection) and time (in the next subsection). The reduction to a compact set of parameters can be achieved by building a reduced basis such that the projection  $\mathcal{P}$  of any training set waveform onto the basis will be indistinguishable from the original waveform up to some tolerance that is specified. We build a reduced basis using a greedy algorithm [25–28], which exposes the most relevant parameters in the training set that capture the salient features of the waveform training data.

The greedy algorithm we use is given in [24], to which we refer the reader for further details, and shown in Alg. 1. The algorithm terminates after  $n$  iterations when the projection error is smaller than a specified tolerance  $\epsilon$ . The output of the greedy algorithm includes a set of parameter tuples  $\{\mu_i\}_{i=1}^n$ , sometimes called *greedy parameters* or *greedy points*, and a reduced basis  $\{e_i(t)\}_{i=1}^n$ . We use the symbol  $X$  as a place holder for a waveform variable. In Alg. 1 and throughout,  $\|\cdot\|^2 = \langle \cdot, \cdot \rangle$  is the squared  $L_2$  norm where

$$\langle f, g \rangle = \int dt f^*(t)g(t) \quad (13)$$

is the integral of the product of two generally complex functions  $f(t)$  and  $g(t)$  and is sometimes called an *inner product*.

To make the reduced basis as compact as possible, we follow [24, 30, 31, 34] and represent the complex waveforms by their amplitudes and phases instead of real and imaginary parts. We then build a separate reduced basis for each of the amplitude and phase training data. This decomposition is made because the amplitude and phase of a waveform have less variation and features in both time and parameters than do the real and imaginary parts of the waveform, which have many oscillations. The greedy algorithm is sensitive to waveform morphologies and will tend to increase the size of the reduced basis in order to resolve these structures. As a result, using an amplitude and phase representation for the waveform allows, in our case, for an extremely compact pair of bases.

We executed the greedy algorithm separately on the amplitudes and the phases of the training set waveforms. In practice, we use an iterated, modified Gram-Schmidt process [75], which is known to be robust against the accumulation of numerical round-off, to generate the orthogonal basis vectors in line 9 of Alg. 1. We chose relative tolerances of  $\epsilon = 10^{-10}$  and  $10^{-15}$  for the amplitude and phase, respectively. We found that this resulted in extremely compact reduced basis sizes of  $n_A = 12$  for amplitudes and  $n_\Phi = 7$  for phases. These 12 and 7 greedy parameters are the minimal amount of information needed to represent the entire training set of  $\text{TEOB}_{\text{Resum}}$  waveform amplitudes and phases, respectively, to within our chosen tolerances, as measured with the squared  $L_2$  norm. Mathematically, if  $A$  and  $\Phi$  denote

the amplitude and phase of a waveform and if  $\{e_i^A(t)\}_{i=1}^{12}$  and  $\{e_i^\Phi(t)\}_{i=1}^7$  are their corresponding reduced bases then

$$A(t; \boldsymbol{\theta}) \approx \sum_{i=1}^{12} e_i^A(t) c_i^A(\boldsymbol{\theta}) \quad (14a)$$

$$\Phi(t; \boldsymbol{\theta}) \approx \sum_{i=1}^7 e_i^\Phi(t) c_i^\Phi(\boldsymbol{\theta}) \quad (14b)$$

where the coefficients are

$$c_i^A(\boldsymbol{\theta}) = \langle e_i^A(\cdot), A(\cdot; \boldsymbol{\theta}) \rangle \quad (15a)$$

$$c_i^\Phi(\boldsymbol{\theta}) = \langle e_i^\Phi(\cdot), \Phi(\cdot; \boldsymbol{\theta}) \rangle. \quad (15b)$$

The orthonormal reduced basis elements for the phase are shown in Fig. 3. For comparison, if we had constructed a reduced basis for the complex waveform itself then the basis size would be several hundred to reach a maximum projection error of  $10^{-10}$  across the entire training set of waveforms. The choice to decompose the waveforms into amplitude and phase is thus justified here.

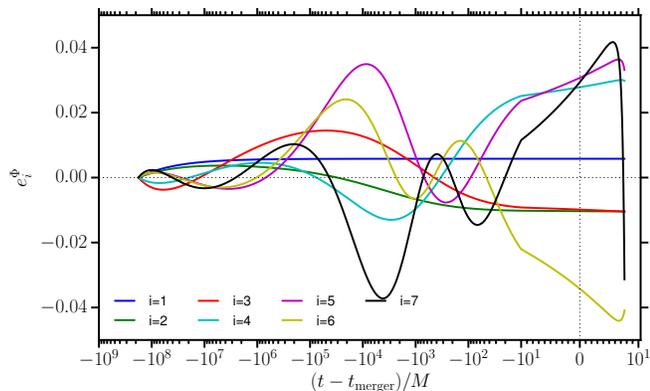


FIG. 3. The seven orthonormal reduced basis functions  $\{e_i^\Phi(t)\}_{i=1}^7$  for the waveform phase. These functions accurately capture the features in the waveform phases within the range of parameters considered. The  $x$ -axis is linear in the range  $[-10^1, 10^1]$  and logarithmic elsewhere.

The reduced basis greedy algorithm depends on a choice of seed, which is arbitrarily selected. The resulting sequence of parameter tuples selected by the greedy algorithm will depend on that choice of seed. A different choice still produces a reduced basis that represents all of the training data to within the specified tolerance, by construction. However, the sizes of the reduced bases built from different seeds tend to vary but will often lie within a few percent of each other so that the seed choice is immaterial [29, 73]. What matters is that we have a compact reduced basis to represent accurately the waveform amplitude and phase training data.

Greedy algorithms are extremely flexible for incorporating many decisions and choices for building a reduced-

order model. For example, in a greedy algorithm one can measure the projection errors with the  $L_2$  norm, as we did here, or with the  $L_\infty$  norm to provide a more stringent requirement of the reduced basis to represent the data in a point-wise sense. Other problem-specific error measures may be more appropriate (e.g., see [40]). One may also implement different greedy algorithms strategies for very large training spaces. For example, the training set can be randomly resampled at every iteration of the greedy algorithm, as described in [40]. Yet another strategy is to parallelize the computation of the inner product integrals used to compute the projections onto the reduced basis [76].

### C. Step 2: Empirical interpolation

The reduced basis representations of the amplitude and phase in (14) require knowing the coefficients in (15) or, equivalently, the data  $A(t; \boldsymbol{\theta})$  and  $\Phi(t; \boldsymbol{\theta})$  that one is projecting onto the basis. We wish to predict these coefficients in the linear representations. One way to do this is to simply fit for the parametric dependence (i.e.,  $\boldsymbol{\theta}$ ) of the coefficients themselves [30, 31]. However, the  $\boldsymbol{\theta}$  dependence of the coefficients can become increasingly noisy as the index  $i$  increases, which can make the fits unreliable and the ensuing surrogate model evaluations not meet stringent accuracy requirements at new parameter values [24, 77].

The second step of surrogate building aims to provide a sparse subset of times from which it is possible to reconstruct the waveform at any other time by using an *empirical interpolant* [32], which is informed by the structure and features of  $\text{TEOB}_{\text{Resum}}$  waveforms via the reduced basis found in Step 1. As we will discuss in the next subsection, fitting for the  $\boldsymbol{\theta}$  dependence of the coefficients is done at each of these time subsamples, which are called *empirical interpolation nodes*. The data being fitted turn out to be robust to the gradual appearance of round-off noise, unlike fitting directly for the projection coefficients mentioned in the previous paragraph [24].

---

#### Algorithm 2 Empirical Interpolation (EI) Method

---

- 1: **Input:**  $\{e_i\}_{i=1}^n$ ,  $t := \{t_i\}_{i=1}^L$
  - 2:  $i = \text{argmax}|e_1(t)|$  ( $\text{argmax}$  returns the largest entry of its argument).
  - 3: Set  $T_1 = t_i$
  - 4: **for**  $j = 2 \rightarrow n$  **do**
  - 5:     Build  $\mathcal{I}_{j-1}[e_j](t)$  from (22)–(24)
  - 6:      $\vec{r} = \mathcal{I}_{j-1}[e_j](t) - e_j(t)$
  - 7:      $i = \text{argmax}|\vec{r}|$
  - 8:      $T_j = t_i$
  - 9: **end for**
  - 10: **Output:** EI nodes  $\{T_i\}_{i=1}^n$ , interpolant operator  $\mathcal{I}_n$
-

The algorithm for building an empirical interpolant is given in [24], to which we refer the reader for further details, and shown in Alg. 2. The empirical interpolant is built from a second greedy algorithm and proceeds as follows [33]. (We focus the presentation on the waveform amplitude for clarity but the same steps are taken for the phase.) First, we choose a value of time  $\tau_1$  from the discrete set of available time samples  $t := \{t_i\}_{i=1}^L$  where  $L$  is the number of samples. In our case, we mentioned in Sec. III A that our data have  $L = 7 \times 10^4$  time samples that are nonuniformly distributed. This first time subsample  $\tau_1$  is a seed for this greedy algorithm. In practice, for reasons of conditioning, one chooses the seed to be the time sample at which the first reduced basis function  $e_1(t)$  is a maximum in absolute value so that  $\tau_1 = \arg \max_t |e_1(t)|$ . We follow this convention here.

For the next step, we build an empirical interpolant. We label the empirical interpolant  $\mathcal{I}$  of a function  $A(t; \boldsymbol{\theta})$  by the number  $m$  of time subsamples we currently have, namely,  $\mathcal{I}_m[A](t; \boldsymbol{\theta})$ . The notation here is to indicate that  $\mathcal{I}_m$  is an operator that acts on a function  $A(t; \boldsymbol{\theta})$ . If the function is independent of  $\boldsymbol{\theta}$  then so will the interpolant's operation on that function.

Currently,  $m = 1$  and we represent the interpolant as a linear combination of the first  $m$  reduced basis elements so that

$$\mathcal{I}_1[A](t; \boldsymbol{\theta}) = e_1(t)C_1(\boldsymbol{\theta}). \quad (16)$$

We assume that the empirical interpolant can always be written in affine form where the dependence on parameters  $\boldsymbol{\theta}$  and time  $t$  is factorized. To solve for the unknown coefficient  $C_1(\boldsymbol{\theta})$  in (16) we demand that the interpolant reproduce the data at  $t = \tau_1$  so that  $\mathcal{I}_1[A](t; \boldsymbol{\theta}) = A(t; \boldsymbol{\theta})$ . The solution is easily found and given by  $C_1(\boldsymbol{\theta}) = A(\tau_1; \boldsymbol{\theta})/e_1(\tau_1)$  and the empirical interpolant so far is given by

$$\mathcal{I}_1[A](t; \boldsymbol{\theta}) = B_1(t)A(\tau_1; \boldsymbol{\theta}) \quad (17)$$

where  $B_1(t) = e_1(t)/e_1(\tau_1)$ . In operator form, the  $m = 1$  empirical interpolant is  $\mathcal{I}_1[\cdot] = B_1(t)(\cdot|_{t=\tau_1})$ .

The second empirical interpolant node  $\tau_2$  is the time subsample at which the next reduced basis element  $e_2(t)$  and its interpolation with the current interpolant  $\mathcal{I}_1[e_2](t)$  is largest in absolute value,

$$\tau_2 := \arg \max_t |e_2(t) - \mathcal{I}_1[e_2](t)| \quad (18)$$

Notice that we are choosing the next time subsample in an effort to improve the empirical interpolant's point-wise representation of the reduced basis elements themselves. Recall that the reduced basis is all that is needed to accurately span the (training set of) waveforms. The set of nodes is now  $\{\tau_1, \tau_2\}$  and

$$\mathcal{I}_2[A](t; \boldsymbol{\theta}) = \sum_{i=1}^2 e_i(t)C_i(\boldsymbol{\theta}) \quad (19)$$

is the empirical interpolant at this step. The coefficients are found as in usual interpolation problems. At the interpolation nodes  $\{\tau_1, \tau_2\}$  we require that (19) equals to the data  $A(\tau_i; \boldsymbol{\theta})$  and then solve the linear equation

$$\sum_{i=1}^2 V_{ji}C_i(\boldsymbol{\theta}) = A(\tau_j; \boldsymbol{\theta}) \quad (20)$$

where  $V_{ji} := e_i(\tau_j)$ , which are the elements of a Vandermonde matrix. After finding the solution, the  $m = 2$  empirical interpolant is

$$\mathcal{I}_2[A](t; \boldsymbol{\theta}) = \sum_{j=1}^2 B_j(t)A(\tau_j; \boldsymbol{\theta}) \quad (21)$$

where  $B_j(t) = \sum_{i=1}^2 e_i(t)(V^{-1})_{ij}$  for  $i = 1, 2$ .

This process is repeated until we have used all  $n$  of the reduced basis elements to build the final empirical interpolant,

$$\mathcal{I}_n[A](t; \boldsymbol{\theta}) = \sum_{j=1}^n B_j(t)A(\tau_j; \boldsymbol{\theta}) \quad (22)$$

where

$$B_j(t) = \sum_{i=1}^n e_i(t)(V^{-1})_{ij} \quad (23)$$

is the  $j^{\text{th}}$  element of the empirical interpolation operator  $\mathcal{B}$ . In operator form, the empirical interpolant is

$$\mathcal{I}_n[\cdot] = \sum_{j=1}^n B_j(t)(\cdot|_{t=\tau_j}) \quad (24)$$

Notice that  $\mathcal{B}$  in (23) is independent of the parameter  $\boldsymbol{\theta}$  and can be computed off-line once the reduced basis is built in Step 1. The quantities  $\{\tau_j\}_{j=1}^n$  are the corresponding interpolation nodes. In addition, the parameter dependence of the empirical interpolant depends on how the function  $A$  varies only at the nodes  $\{\tau_j\}_{j=1}^n$ . Finally, notice that, given a parameter tuple  $\boldsymbol{\theta}$ , one still needs to know the actual values of the  $\{A(\tau_j; \boldsymbol{\theta})\}_{j=1}^n$  in order to compute the empirical interpolant in (22). We will show how surrogate modeling addresses this issue in Sec. III D.

We applied this greedy algorithm to build an empirical interpolant separately for the waveform amplitude and phase, which are given by

$$\mathcal{I}_{12}^A[A](t; \boldsymbol{\theta}) = \sum_{j=1}^{12} B_j^A(t)A(\tau_j^A; \boldsymbol{\theta}) \quad (25a)$$

$$\mathcal{I}_7^\Phi[\Phi](t; \boldsymbol{\theta}) = \sum_{j=1}^7 B_j^\Phi(t)\Phi(\tau_j^\Phi; \boldsymbol{\theta}) \quad (25b)$$

The interpolating functions for the phase are shown in

Fig. 4. The empirical interpolation nodes  $\{\tau_j^\Phi\}_{j=1}^7$  are not uniformly spaced because they depend on the underlying features of the  $\text{TEOB}_{\text{Resum}}$  waveform family.

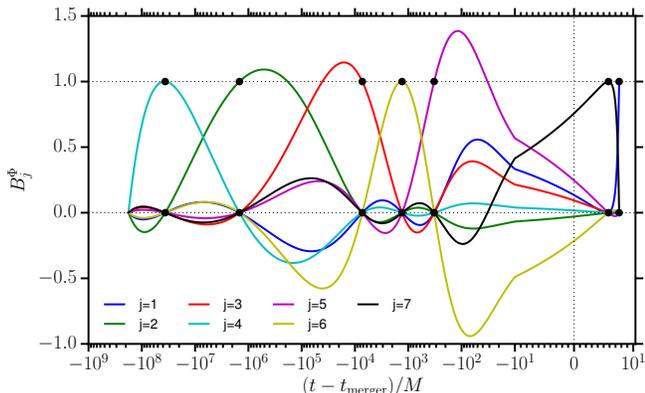


FIG. 4. The seven elements of the empirical interpolation operator  $\{B_j^\Phi(t)\}_{j=1}^7$  for the waveform phase. The empirical interpolation nodes  $\{\tau_j^\Phi\}_{j=1}^7$  are shown as black dots. The  $x$ -axis is linear in the range  $[-10^1, 10^1]$  and logarithmic elsewhere.

The empirical interpolation greedy algorithm we just described and used in this paper is not optimized for speed. As discussed in Appendix B of [45], the original (discrete) empirical interpolation method algorithm proposed in [33], which is the one we discussed above, has a computational cost at the  $m^{\text{th}}$  step of the greedy algorithm that scales as  $\mathcal{O}(m^4)$ . However, this relatively slow evaluation time is immaterial for our surrogate since the sizes of the amplitude ( $n_A = 12$ ) and phase ( $n_\Phi = 7$ ) reduced bases are very small. A faster algorithm was put forward in [45] (see Algorithm 5 in their Appendix A) and was used in describing the reduced-order surrogate modeling strategy in [24]. This faster implementation has a cost that scales as  $\mathcal{O}(m^3)$  [45]. The computational savings with this faster algorithm is particularly useful for problems involving a large reduced basis, such as the reduced-order model built in [44] for the IMRPhenomPv2 waveform family [78] that contained at most 1253 basis elements.

#### D. Step 3: Estimating the parametric variation

The last step in building the surrogate model is to estimate the  $\theta$  dependence at each of the empirical interpolation nodes for both the waveform amplitude and phase data. Because the parameter space is three dimensional, we intentionally selected our training points in Sec. III A to correspond to the nodes of Chebyshev interpolation in three dimensions. Chebyshev interpolation, which for smooth  $C^\infty$  functions, has errors that converge exponentially with the number of Chebyshev polynomials  $T_n(x)$  [72]. The amplitude and phase at each empirical

interpolation node are approximated as a tensor product of Chebyshev polynomials,

$$A(\tau_j^A; \theta) \approx \tilde{A}_j(\theta) := \sum_{l,m,n} a_{j,lmn} T_l(q) T_m(\Lambda_2^A) T_n(\Lambda_2^B), \quad (26a)$$

$$\Phi(\tau_j^\Phi; \theta) \approx \tilde{\Phi}_j(\theta) := \sum_{l,m,n} b_{j,lmn} T_l(q) T_m(\Lambda_2^A) T_n(\Lambda_2^B). \quad (26b)$$

for  $j = 1, \dots, n_X$  and  $X = \{A, \Phi\}$ . Although it would be possible to optimize the number of coefficients for each node  $\tau_j$ , for simplicity we use all  $16 \times 16 \times 16$  coefficients at each node. The summations are efficiently performed using Clenshaw summation [79]. The coefficients  $a_{j,lmn}$  and  $b_{j,lmn}$  of the Chebyshev series are precomputed from the known amplitudes and phases on the training set grid using Gaussian quadrature [72]. This quadrature is efficiently performed using a type-I discrete cosine transform [80].

The required fractional accuracy of the interpolation is the maximum allowed amplitude or phase error divided by the range of values that the amplitude or phase takes at each empirical node  $\tau_j$  over the training set. This is most stringent for the phase where, near the merger, we would like the error to be  $\lesssim 0.1$  radians but the spread is  $\sim 10^3$  radians over the considered parameter space. This requires a fractional accuracy of  $\sim 10^{-4}$  for the interpolation. We assess the errors in the surrogate model due to interpolation in Sec. IV A.

A distinct advantage of reduced-order surrogate models is that the data output from a simulation or code is used *directly* for building the model. The first steps (building a reduced basis and an empirical interpolant) are accomplished off-line using *only* the training data generated by the  $\text{TEOB}_{\text{Resum}}$  code. The only model inputs come in the last step when we estimate the parameter dependence of the data at the empirical interpolation nodes because we implement a choice of fitting functions that can affect the resulting quality of the surrogate predictions. This often results in a surrogate that nearly retains the accuracy of the underlying training data used to build the model. In addition, propagating the training data uncertainties and assessing the surrogate errors is fairly straightforward because of the minimal amount of modeling inputs, which are isolated to Step 3. An example of this is given in [34] where a reduced-order surrogate model is built for the gravitational waveforms of non-spinning BBH coalescences produced by numerical relativity simulations.

#### E. Surrogate waveform evaluation

After Step 3, the surrogate model for the amplitude and phase is defined by evaluating the corresponding empirical interpolant in (25) using the parametric estima-

tions in (26) to predict the values at any new parameter values,

$$A_S(t; \boldsymbol{\theta}) := \sum_{i=1}^{12} B_i^A(t) \tilde{A}_i(\boldsymbol{\theta}) \quad (27a)$$

$$\Phi_S(t; \boldsymbol{\theta}) := \sum_{i=1}^7 B_i^\Phi(t) \tilde{\Phi}_i(\boldsymbol{\theta}) \quad (27b)$$

The surrogate evaluation is performed online once a parameter tuple  $(q, \Lambda_2^A, \Lambda_2^B)$  is given.

In practical gravitational wave data analysis applications, one rescales from geometric units to physical units of time and amplitude. This rescaling also depends on the total mass  $M$  of the compact binary. In addition, one specifies a starting frequency  $f_{\text{start}}$ . The starting time  $t_{\text{start}}$  corresponding to  $f_{\text{start}}$  is calculated by numerically solving  $f_{\text{start}} = f(t_{\text{start}})$ , where  $f(t) = (d\Phi_S/dt)/(2\pi)$  is the frequency once given a tuple of parameter values  $(q, \Lambda_2^A, \Lambda_2^B)$ . We then perform the necessary time and phase shifts and resample the amplitude and phase surrogate predictions in (27). The final expression for the surrogate model waveforms takes the form

$$h_{+S}(t; \boldsymbol{\theta}) = \frac{1}{2}(1 + \cos \iota) \frac{GM}{c^2 d} A_S\left(\frac{c^3 t}{GM}; \boldsymbol{\theta}\right) \times \cos \Phi_S\left(\frac{c^3 t}{GM}; \boldsymbol{\theta}\right), \quad (28a)$$

$$h_{\times S}(t; \boldsymbol{\theta}) = \cos \iota \frac{GM}{c^2 d} A_S\left(\frac{c^3 t}{GM}; \boldsymbol{\theta}\right) \sin \Phi_S\left(\frac{c^3 t}{GM}; \boldsymbol{\theta}\right), \quad (28b)$$

where  $\iota$  is the binary inclination angle and  $d$  is the distance to the gravitational wave source.

In this paper, we have built a surrogate model for the amplitude and phase of  $\text{TEOB}_{\text{Resum}}$  waveforms in the time domain. This surrogate approximates the waveform in a nonlinear representation (because the dependence on phase is nonlinear) and, as such, cannot be used directly for speeding up likelihood computations in parameter estimation studies [42–44] that use reduced-order quadratures [45]. However, it is straightforward to use our surrogate to build a new surrogate for this purpose in the following way. As we will discuss in the next section, our surrogate accurately predicts waveforms output by the  $\text{TEOB}_{\text{Resum}}$  code. As such, we may use the surrogate above to generate a new training set in  $h_+$  and  $h_\times$  form, which is a linear representation. Then, one can repeat the surrogate building Steps 1–3 for the two waveform polarizations directly. As mentioned earlier, this would generate reduced basis sizes that contain a few hundred elements and so would be somewhat larger in physical memory size and slower to evaluate the resulting surrogate. Furthermore, one could do this in the frequency domain by computing Fourier transforms of the time-domain training waveforms generated by the surrogate in (28) so that the transformation is part of the offline

stage. One may still use our time domain surrogate presented here for parameter estimation studies but we expect significant speed-ups could be obtained by following the strategy just outlined for likelihood computations.

## IV. RESULTS

### A. Accuracy

The required accuracy of the surrogate is determined by the smallest effect that we want to model. For BNS systems, this is the tidal interaction that effects the waveform by  $\sim 10$  radians up to merger for typical EOSs. This means that we will require the waveform error to be significantly smaller than  $\sim 10$  radians.

The accuracy of the surrogate can be assessed by comparing it to the training set used to construct the reduced basis as well as to a large set of waveforms with randomly sampled parameters. Fig. 5 shows the fractional error in the amplitude as well as the error in the phase between the surrogate waveform and each of the 4096 training set waveforms. The difference between the surrogate and each training set waveform is maximized over all times. For clarity we suppress the third parameter  $\Lambda_2^B$  and only show the maximum error for the 16 values of  $\Lambda_2^B$  at each grid point. The maximum fractional error in amplitude is  $\Delta A/A = 7.7 \times 10^{-5}$ , and the maximum error in phase is  $\Delta \Phi = 0.014$  radians. Because the interpolated values of  $A$  and  $\Phi$  exactly match the training set at each grid point at the empirical nodes  $\tau_j$ , we see from (27) that the error in reproducing each training set waveform is due almost entirely to the finite number of reduced bases.

To determine how well the surrogate reproduces a generic waveform within the parameter space, we produce  $10^4$  waveforms with parameters randomly sampled in the range  $q \in [0.5, 1]$ ,  $\Lambda_2^A \in [50, 5000]$ , and  $\Lambda_2^B \in [50, 5000]$ . The fractional amplitude and phase errors in reproducing the generic waveforms are shown in Fig. 6. The errors are maximized over time for each waveform. We find a maximum fractional amplitude error of  $\Delta A/A = 0.038$ , but note that, prior to the last  $100M$  before merger, the maximum error is  $\Delta A/A = 4 \times 10^{-4}$ , about two orders of magnitude smaller. The maximum phase error is  $\Delta \Phi = 0.043$  radians. For both the amplitude and phase, the error is largest for small values of the parameters. This results because both the amplitude and phase vary most rapidly for small values of the waveform parameters, so the interpolation of the amplitude  $A(\tau_j^A; q, \Lambda_2^A, \Lambda_2^B)$  and phase  $\Phi(\tau_j^\Phi; q, \Lambda_2^A, \Lambda_2^B)$  at each empirical node  $\tau_j$  is least accurate there.

In Fig. 7, we show the last  $10^3 M$  of the  $\sim 10^8 M$  long waveform that is reproduced by the surrogate with the largest phase error. The phase error typically increases gradually with time. On the other hand, the amplitude error increases dramatically during the last cycle, and is typically two orders of magnitude smaller before the last cycle. This increase in error near the maximum ampli-

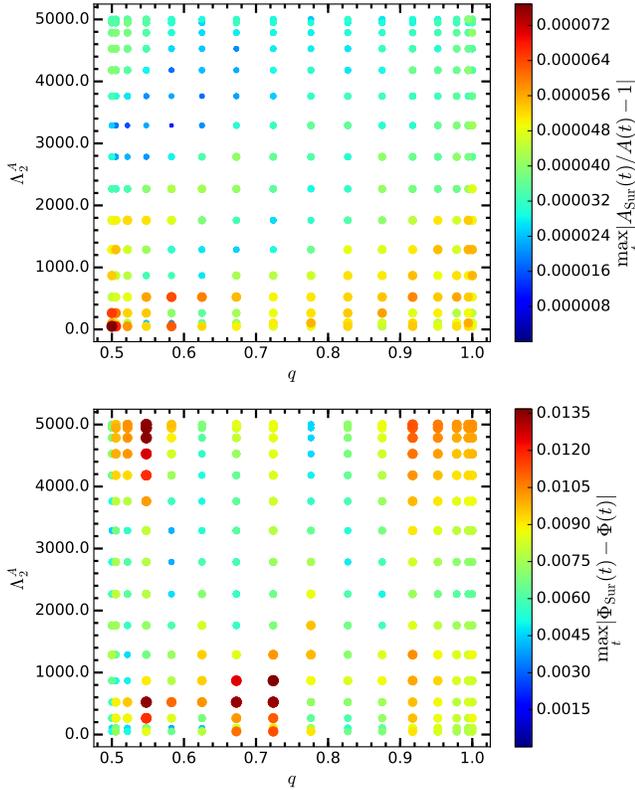


FIG. 5. Error between the surrogate and the  $16^3$  training set waveforms used to construct the reduced basis. Larger points represent larger errors. The fractional amplitude and phase errors, maximized over time and waveform parameters, are  $\Delta A/A = 7.7 \times 10^{-5}$  and  $\Delta\Phi = 0.014$  respectively. These errors are due to the finite number of reduced bases.

tude likely results from the finite accuracy with which the training set waveforms are numerically aligned at maximum amplitude where the amplitude changes more rapidly with time, making accurate interpolation more difficult.

As a final check of the surrogate model accuracy, we examine the mismatch between the surrogate model and the original EOB waveform. The mismatch represents the loss in signal-to-noise ratio that would result from using the surrogate model instead of the original EOB waveform. It is defined by the deviation from a perfect overlap after aligning the two waveforms using the time and phase free parameters  $t_0$  and  $\phi_0$ :

$$\mathcal{M} = 1 - \max_{t_0, \phi_0} \frac{(h_{\text{EOB}}, h_{\text{Sur}})}{\sqrt{(h_{\text{EOB}}, h_{\text{EOB}})(h_{\text{Sur}}, h_{\text{Sur}})}}. \quad (29)$$

The inner product here is the integral of the Fourier transformed waveforms  $\tilde{h}(f)$  weighted by the noise power spectral density (PSD)  $S_n(f)$  of the detector:

$$(h_1, h_2) = 4\Re \int_{f_{\text{low}}}^{f_{\text{high}}} \frac{\tilde{h}_1(f)\tilde{h}_2^*(f)}{S_n(f)} df. \quad (30)$$

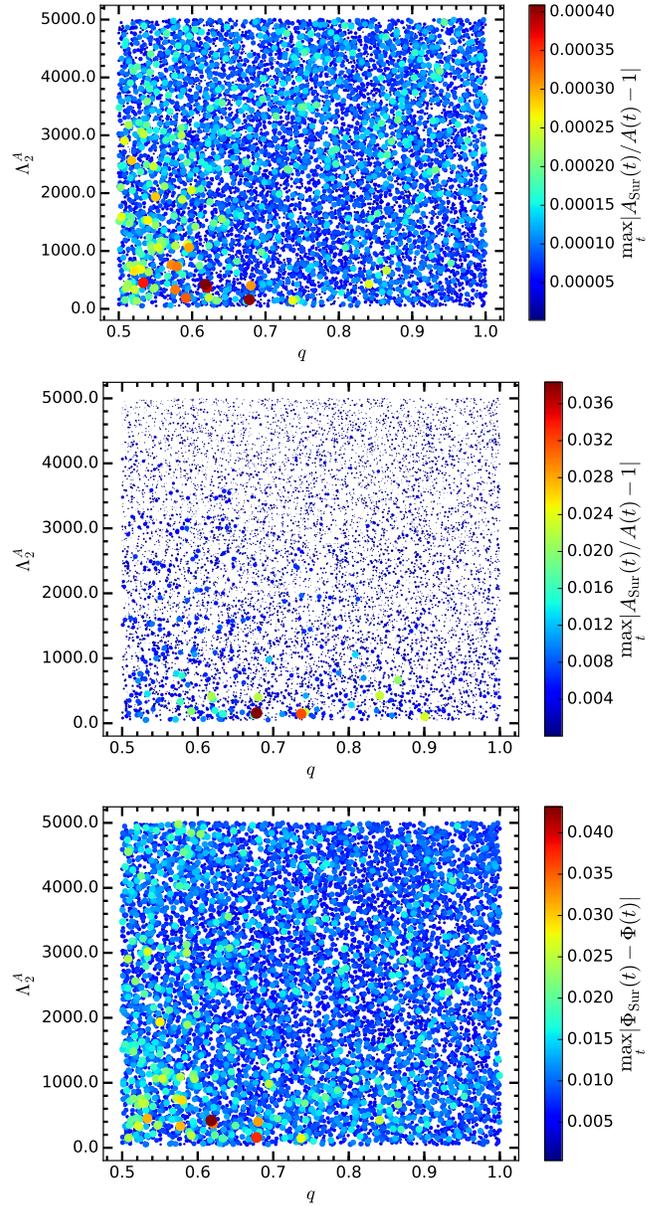


FIG. 6. Error between the surrogate and  $10^4$  waveforms with randomly sampled parameters not in the training set. Larger points represent larger errors. Top: Fractional amplitude error  $\Delta A/A$  maximized over all times except the last  $100M$ . Middle: Fractional amplitude error  $\Delta A/A$  maximized over all times. Bottom: Phase error  $\Delta\Phi$  maximized over all times.

In Fig. 8, we show the distribution of mismatch  $\mathcal{M}$  between our surrogate and the  $10^4$  randomly sampled EOB waveforms. We use the design sensitivity aLIGO PSD [81] and a sampling rate of 4096 Hz. Our integration bounds are  $f_{\text{low}} = 30$  Hz and the Nyquist frequency  $f_{\text{high}} = 2048$  Hz. Because the surrogate can be rescaled with mass, we show results for the smaller mass  $M_B$  fixed at  $1M_\odot$  or fixed at  $2M_\odot$ . The mismatch is larger for the higher mass systems because the frequency where

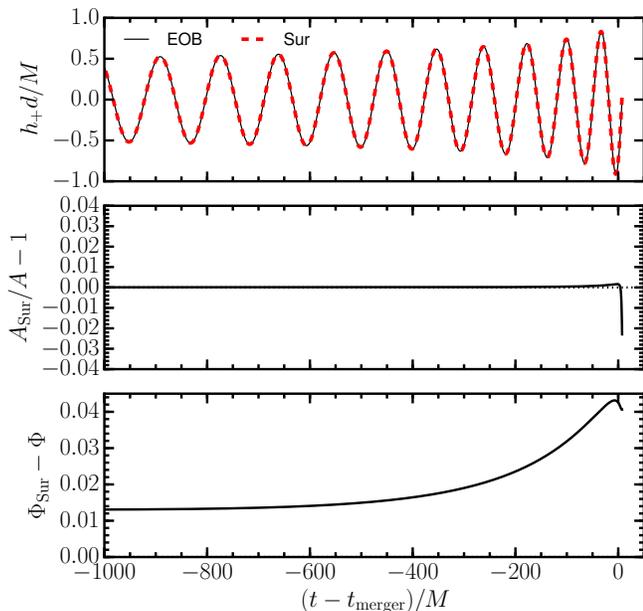


FIG. 7. Final  $\sim 10$  cycles of the waveform with the largest phase error. The parameters are  $\{q, \Lambda_2^A, \Lambda_2^B\} = \{0.618, 420, 81\}$ . The amplitude errors are always largest during the last gravitational wave cycle, and are about two orders of magnitude smaller before the last gravitational wave cycle.

the waveform ends scales inversely with the total mass, resulting in the less accurate end of the waveform occurring at smaller frequencies where the detector is more sensitive. Overall, the mismatch is typically smaller than  $\sim 10^{-4}$  except for systems with large component masses, and the mismatch never exceeds  $7 \times 10^{-4}$ .

## B. Timing

One of the main purposes for generating a surrogate is to make parameter estimation more computationally efficient. In particular, parameter estimation codes, usually based on Markov Chain Monte Carlo or Nested Sampling, typically require  $10^7$ – $10^8$  sequential waveform evaluations, and the computational time is dominated by the waveform evaluation time even for PN waveforms [82, 83]. Recently, a few parallel algorithms have become available [84, 85]. However, the performance of these algorithms still scales linearly with the performance of the waveform generator.

We have produced a prototype Python implementation of our surrogate and a C implementation in the LIGO Algorithm Library (LAL) [86] under the name `TEOBResum_ROM` which is about 2 times faster than the Python version. The performance of the LAL implementation is shown in Fig. 9. For each waveform evaluation, there is a flat cost of  $\sim 0.04$  s to calculate the amplitude and phase of the waveform at the  $\sim 7 \times 10^4$  samples using

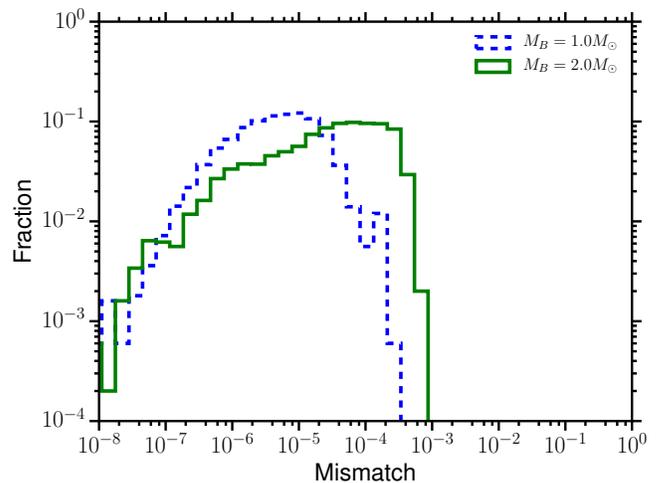


FIG. 8. Histogram of the mismatch between the surrogate and the  $10^4$  randomly sampled EOB waveforms. The smaller mass  $M_B$  is set to either  $1M_\odot$  or  $2M_\odot$ . The PSD corresponds to the aLIGO design sensitivity and the lower and upper frequency integration bounds are  $f_{\text{low}} = 30$  Hz and  $f_{\text{high}} = 2048$  Hz.

Eq. (27). The difference in evaluation time as the starting frequency  $f_{\text{start}}$  is varied results from the resampling of the amplitude and phase at evenly spaced times in Eq. (28). The number of these samples increases rapidly as the starting frequency is decreased. Most parameter estimation is done at 4096 Hz with a starting frequency of 30 Hz down to 10 Hz. The total waveform evaluation time will therefore be about 0.07 s up to 0.8 s. If necessary, the resampling of the amplitude and phase, which is the bottleneck for low starting frequencies, can be parallelized. The original implementation of the EOB code in Matlab can take  $\sim 20$  minutes to evaluate, so this represents a speed-up factor of  $10^3$ – $10^4$  in some cases. However, it is not possible to make a direct comparison since the original Matlab code calculates a non-uniformly sampled waveform with units rescaled by the distance and mass.

## C. Parameter estimation

As an end-to-end test of the surrogate, we inject a waveform into simulated aLIGO data and estimate its parameters using a Nested Sampling algorithm as implemented in LAL [87–90]. Although EOB waveforms have been used in Bayesian parameter estimation for BBH systems which have significantly fewer cycles (e.g. [91]), this is one of the first times EOB waveforms have been used in a Bayesian analysis for BNS systems. Another analysis using a parallelized algorithm [85] is in progress [92].

The synthetic data are taken to be stationary and Gaussian, with a PSD corresponding to the aLIGO final design sensitivity [81]; two detectors are assumed,

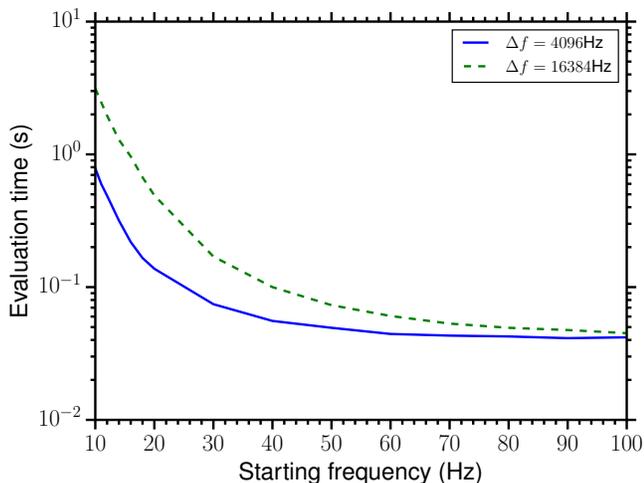


FIG. 9. Performance of the surrogate implemented in LAL for a binary with component masses  $(1.4, 1.4)M_{\odot}$ . We used sampling frequencies of 4096Hz and 16384Hz. The evaluation time at each starting frequency is averaged over 16 evaluations on a 2.7 GHz Intel Xeon CPU using one core.

located at the Hanford and Livingston sites. For the simulated signal we choose an arbitrary sky position and orientation, a distance of 87 Mpc, and component masses  $(M_A, M_B) = (1.4, 1.4)M_{\odot}$ . The EOS is taken to be MS1b, so that for the given masses one has  $\Lambda_2^A = \Lambda_2^B = 1286$  [7]. For the parameter choices made, the optimal signal-to-noise ratio is 26.28.

For the parameter estimation, the prior densities for sky position as well as orientation are chosen to be uniform on the sphere, and the distance prior is uniform in co-moving volume with an upper cut-off at 100 Mpc. Component masses are uniform in the interval  $[1, 2]M_{\odot}$ , and we take  $\Lambda_2^A, \Lambda_2^B$  to be uniform in the interval  $[50, 5000]$ . In the Nested Sampling we use 512 live points and up to 5000 MCMC points [89], leading to  $10^7 - 10^8$  likelihood evaluations; results from 4 different sampling chains are combined. Since the tidal effects we are interested in manifest themselves predominantly at high frequency, for this first exploration we use a lower cut-off frequency of 40 Hz; the sampling rate is 4096 Hz.

Since  $\Lambda_2^A$  and  $\Lambda_2^B$  are highly correlated, after the Nested Sampling algorithm has finished we change parameters to the  $\Lambda, \delta\tilde{\Lambda}$  introduced in [6], which depend on  $(q, \Lambda_2^A, \Lambda_2^B)$ ; note that these have the convenient properties  $\tilde{\Lambda}(q = 1, \Lambda_2^A = \Lambda_2^B = \Lambda) = \Lambda$  and  $\delta\tilde{\Lambda}(q = 1, \Lambda_2^A = \Lambda_2^B = \Lambda) = 0$ .

Posterior density functions for  $\tilde{\Lambda}$  and  $\delta\tilde{\Lambda}$  are shown in Fig. 10; both these parameters are recovered quite well. The parameter estimation code ran for 15 days; without a surrogate it would have taken well over a year.

Finally, as shown in [5, 7, 8], if the functional dependence of  $\Lambda_2^A, \Lambda_2^B$  on component masses is expressed in terms of observables that can be expected to take approx-

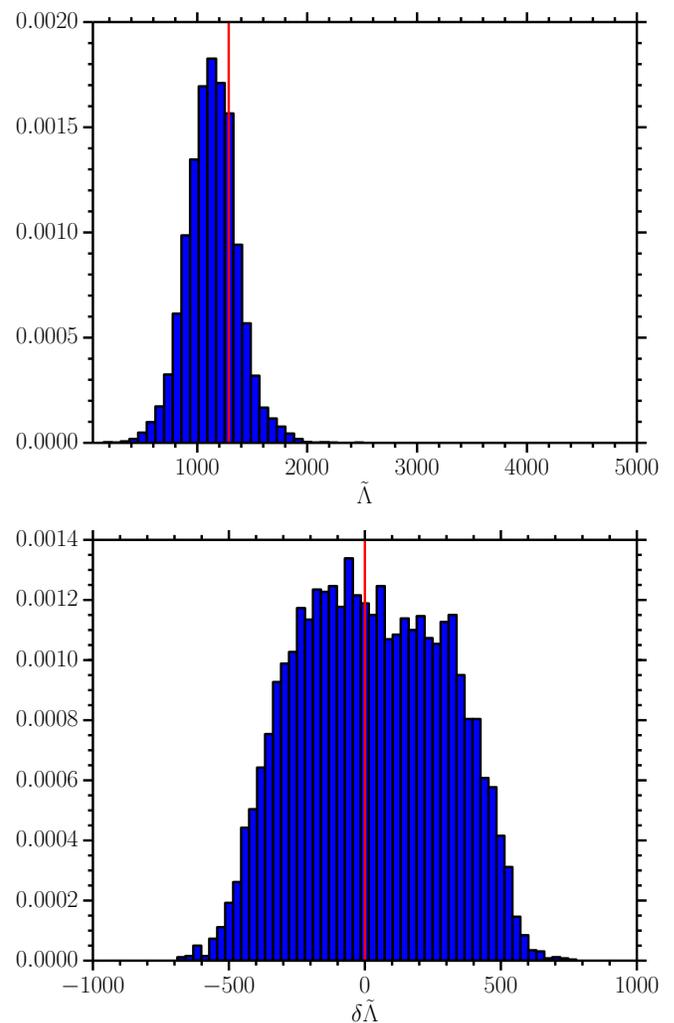


FIG. 10. Posterior densities for the tidal parameters  $\tilde{\Lambda}$  (top) and  $\delta\tilde{\Lambda}$  (bottom). The simulated signal had component masses  $(1.4, 1.4)M_{\odot}$  and EOS MS1b, so that  $\tilde{\Lambda} = 1286$  and  $\delta\tilde{\Lambda} = 0$ , as indicated by the vertical red lines.

imately the same values for all sources, then posterior density functions for the latter can be trivially combined across detections to arrive at a more accurate measurement. These observables could be coefficients in a Taylor expansion [5, 8], or the parameters in a representation of the EOS in terms of piecewise polytropes [7]. An implementation for the surrogate model presented in this paper is left for future work.

## V. DISCUSSION AND FUTURE WORK

We have constructed a surrogate for waveforms from nonspinning BNS systems that includes the  $\ell = 2$  tidal interaction and approximates the  $\ell = 3$  and 4 tidal interactions. The error of the surrogate is small compared to the size of the tidal effect, and we have demonstrated

that this surrogate can be used in generic parameter estimation algorithms.

The implementation of the model in LAL takes  $\sim 0.07$  s to evaluate for  $f_{\text{low}} = 30$  Hz and  $\sim 0.8$  s to evaluate for  $f_{\text{low}} = 10$  Hz, but there remains plenty of room for improvements. For starting frequencies below  $\sim 30$  Hz, the waveform generation time is dominated by straightforward interpolation of the amplitude and phase at uniformly sampled times. This interpolation can be optimized or even parallelized on either CPUs or a GPU if necessary. Most other operations can be parallelized as well. In addition, the number of amplitude and phase bases as well as the the number of time samples used to store the amplitude and phase bases (currently  $\sim 7 \times 10^4$ ) can be optimized, and the number of Chebyshev coefficients for the interpolation can be strategically reduced as well. We expect speed-up factors of a few should be possible without significantly effecting the accuracy of the surrogate. Finally, one could use our time domain surrogate to build a linear frequency domain surrogate that can be used directly in a reduced-order quadratures implementation of likelihood computations for speeding up parameter estimation studies.

Because EOB models that include both spin and tidal interactions are still in progress [93, 94], our surrogate leaves out spin parameters. Once these models are available, adding the two spin magnitudes  $|S_1|$  and  $|S_2|$  for aligned spin systems for a total of five parameters will likely be straightforward using a standard grid-based interpolation scheme. However, incorporating an additional four parameters to account for the spin orientations will likely be significantly more difficult. So far this problem has not been fully solved for BBH systems without tidal interactions.

Finally, we have also left out the post-merger stage for BNS systems. Unlike BBH systems, the post-merger stage can only be modeled by expensive numerical relativity simulations, and it is unlikely that more than 100–1000 simulations could be performed over the course of a few years. However, work by Clark *et al.* has shown that it is possible to reconstruct post-merger waveforms with a small number of orthonormal bases [95]. In future work, we would like to examine the possibility of constructing a surrogate model for the complete inspiral–post-merger waveform for BNS systems.

## ACKNOWLEDGMENTS

BL thanks Larne Pekowsky and Duncan Brown for significant computing help and Rory Smith and Michael Pürerer for helpful discussions at the beginning of this work. SB thanks Paolo Pani for helpful discussions about  $\Lambda_{3,4}^{\text{fit}}(\Lambda_2)$  fits. BL was supported by NSF grant AST-1333142. SB was supported by a Rita Levi Montalcini fellowship of the Italian Ministry of Education, University and Research (MIUR). CRG was supported in part by NSF grant PHY-1404569 to Caltech and by the Sher-

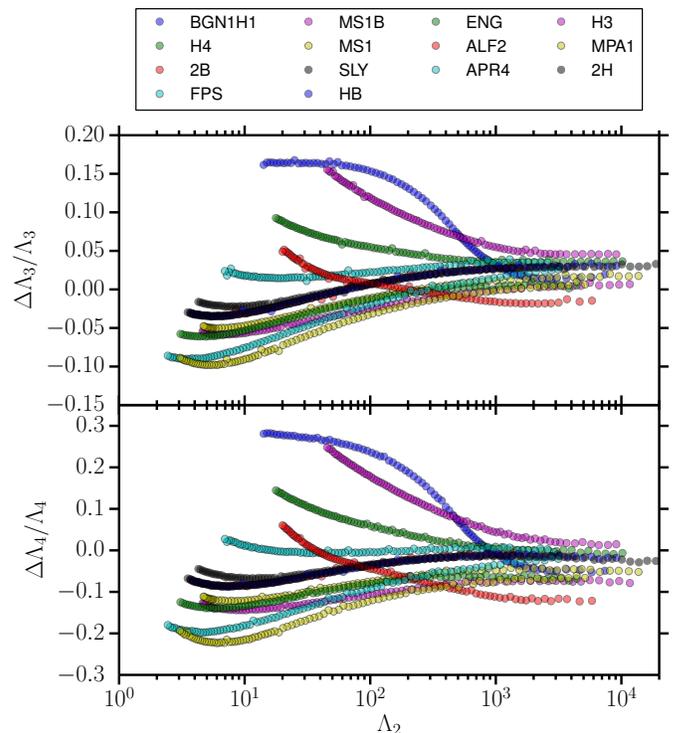


FIG. 11. Relative errors  $\Delta\Lambda_3/\Lambda_3$  and  $\Delta\Lambda_4/\Lambda_4$ , with  $\Delta\Lambda_\ell = \Lambda_\ell - \Lambda_\ell^{\text{fit}}$  (shown only 300 points).

man Fairchild Foundation. JM and CVDB were supported by the research programme of the Foundation for Fundamental Research on Matter (FOM), which is partially supported by the Netherlands Organisation for Scientific Research (NWO). Computations were performed on the Syracuse University Campus Grid which is supported by NSF awards ACI-1341006, ACI-1541396, and by Syracuse University ITS.

## Appendix A: Systematic uncertainties related to $\ell = 3, 4$ tidal polarizability coefficients

In our 3D surrogate the quantities  $\Lambda_{3,4}$  are computed from  $\Lambda_2$  using [67]

$$\ln \Lambda_\ell^{\text{fit}} = \sum_{i=0}^4 a_i^{(\ell)} (\ln \Lambda_2)^i \quad (\ell = 3, 4) \quad (\text{A1})$$

where  $(a_i^{(3)}) = (-1.15, 1.18, 2.51 \times 10^{-2}, -1.31 \times 10^{-3}, 2.52 \times 10^{-5})$  and  $(a_i^{(4)}) = (-2.45, 1.43, 3.95 \times 10^{-2}, -1.81 \times 10^{-3}, 2.80 \times 10^{-5})$ . The accuracy of these fits and the systematic uncertainties that they introduce in the surrogate are tested using a sample of 14 EOS and, for each EOS, about 3000 star configurations spanning the mass range  $M \in [0.9, M_{\text{max}}]M_\odot$ .

Figure 11 and Table I quantify the relative errors of the fits, where by definition  $\Delta\Lambda_\ell = \Lambda_\ell - \Lambda_\ell^{\text{fit}}$ . The largest fit

errors occur for small  $\Lambda_2 \sim 1 - 10$ , which occur when the NS is near its maximum mass. The largest positive error is for EOS BGN1H1 near its maximum mass of  $1.64M_\odot$  and the largest negative error is for MPA1 near its maximum mass of  $2.43M_\odot$ . These fit errors have the range  $\Delta\Lambda_3/\Lambda_3 \in [-0.10, 0.17]$  and  $\Delta\Lambda_4/\Lambda_4 \in [-0.22, 0.29]$ . These high mass configurations, however, are not expected to be found in a BNS system. Comparing with Ref. [67], we obtain a larger range of errors because we use a larger sample of EOS.

Because the tidal effect is largest for equal mass systems and the tidal parameter is largest for smaller masses, the error in the waveform phase due to the fit is largest for equal mass  $q = 1$  systems with smaller masses. In Fig. 1 and Table I, we show the phase error from NSs with more likely masses of  $1.4M_\odot$  using the soft EOS SLY and the stiff EOS MS1b. The phase error grows with time, reaching its maximum near merger, and in general, we find typical phase errors of  $|\Delta\phi| \lesssim 0.01$  radians.

- 
- [1] G. M. Harry and LIGO Scientific Collaboration, *Class. Quantum Grav.* **27**, 084006 (2010).
- [2] F. Acernese, et al. (Virgo Collaboration), “Advanced virgo baseline design, VIR-027A-09,” (2009).
- [3] K. Somiya, *Class. Quantum Grav.* **29**, 124007 (2012), arXiv:1111.7185 [gr-qc].
- [4] B. Iyer, T. Souradeep, C. S. Unnikrishnan, S. Dhurandhar, S. Raja, and A. Sengupta, LIGO Document M1100296-v2 (2011).
- [5] W. Del Pozzo, T. G. F. Li, M. Agathos, C. Van Den Broeck, and S. Vitale, *Phys. Rev. Lett.* **111**, 071101 (2013), arXiv:1307.8338 [gr-qc].
- [6] L. Wade, J. D. E. Creighton, E. Ochsner, B. D. Lackey, B. F. Farr, T. B. Littenberg, and V. Raymond, *Phys. Rev. D* **89**, 103012 (2014).
- [7] B. D. Lackey and L. Wade, *Phys. Rev. D* **91**, 043002 (2015), arXiv:1410.8866 [gr-qc].
- [8] M. Agathos, J. Meidam, W. Del Pozzo, T. G. F. Li, M. Tompitak, J. Veitch, S. Vitale, and C. Van Den Broeck, *Phys. Rev. D* **92**, 023012 (2015), arXiv:1503.05405 [gr-qc].
- [9] K. Chatziioannou, N. Cornish, A. Klein, and N. Yunes, *Astrophys. J. Lett.* **798**, L17 (2015), arXiv:1402.3581 [gr-qc].
- [10] T. Damour, B. R. Iyer, and B. S. Sathyaprakash, *Phys. Rev. D* **63**, 044023 (2001), gr-qc/0010009.
- [11] A. Buonanno, B. R. Iyer, E. Ochsner, Y. Pan, and B. S. Sathyaprakash, *Phys. Rev. D* **80**, 084043 (2009), arXiv:0907.0700 [gr-qc].
- [12] J. Veitch, V. Raymond, B. Farr, W. Farr, P. Graff, S. Vitale, B. Aylott, K. Blackburn, N. Christensen, M. Coughlin, W. Del Pozzo, F. Feroz, J. Gair, C.-J. Haster, V. Kalogera, T. Littenberg, I. Mandel, R. O’Shaughnessy, M. Pitkin, C. Rodriguez, C. Röver, T. Sidery, R. Smith, M. Van Der Sluys, A. Vecchio, W. Voudsen, and L. Wade, *Phys. Rev. D* **91**, 042003 (2015), arXiv:1409.7215 [gr-qc].
- [13] T. Hinderer, B. D. Lackey, R. N. Lang, and J. S. Read, *Phys. Rev. D* **81**, 123016 (2010), arXiv:0911.3535 [astro-ph.HE].
- [14] M. Favata, *Phys. Rev. Lett.* **112**, 101101 (2014), arXiv:1310.8288 [gr-qc].
- [15] K. Yagi and N. Yunes, *Phys. Rev. D* **89**, 021303 (2014), arXiv:1310.8358 [gr-qc].
- [16] A. Buonanno and T. Damour, *Phys. Rev. D* **59**, 084006 (1999), arXiv:gr-qc/9811091.
- [17] T. Damour and A. Nagar, *Phys. Rev. D* **81**, 084016 (2010), arXiv:0911.5041 [gr-qc].
- [18] Y. Pan, A. Buonanno, M. Boyle, L. T. Buchman, L. E. Kidder, H. P. Pfeiffer, and M. A. Scheel, *Phys. Rev. D* **84**, 124052 (2011), arXiv:1106.1021 [gr-qc].
- [19] A. Taracchini, Y. Pan, A. Buonanno, E. Barausse, M. Boyle, T. Chu, G. Lovelace, H. P. Pfeiffer, and M. A. Scheel, *Phys. Rev. D* **86**, 024011 (2012), arXiv:1202.0790 [gr-qc].
- [20] A. Taracchini, A. Buonanno, Y. Pan, T. Hinderer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, A. H. Mroué, H. P. Pfeiffer, M. A. Scheel, B. Szilágyi, N. W. Taylor, and A. Zenginoglu, *Phys. Rev. D* **89**, 061502 (2014), arXiv:1311.2544 [gr-qc].
- [21] T. Damour, A. Nagar, and S. Bernuzzi, *Phys. Rev. D* **87**, 084035 (2013), arXiv:1212.4357 [gr-qc].
- [22] A. Nagar, T. Damour, C. Reisswig, and D. Pollney, *Phys. Rev. D* **93**, 044046 (2016), arXiv:1506.08457 [gr-qc].
- [23] C. Devine, Z. B. Etienne, and S. T. McWilliams, *Classical and Quantum Gravity* **33**, 125025 (2016), arXiv:1601.03393 [astro-ph.HE].
- [24] S. E. Field, C. R. Galley, J. S. Hesthaven, J. Kaye, and M. Tiglio, *Phys. Rev. X* **4**, 031006 (2014), arXiv:1308.3565 [gr-qc].
- [25] T. H. Cormen, C. Stein, R. L. Rivest, and C. E. Leiserson, *Introduction to Algorithms*, 2nd ed. (McGraw-Hill Higher Education, 2001).
- [26] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk, *SIAM Journal on Mathematical Analysis* **43**, 1457 (2011).
- [27] R. DeVore, G. Petrova, and P. Wojtaszczyk, (2012), 1204.2290 [math.FA].
- [28] J. S. Hesthaven, B. Stamm, and S. Zhang, *ESAIM: Mathematical Modelling and Numerical Analysis* **48**, 259 (2014).
- [29] S. E. Field, C. R. Galley, F. Herrmann, J. S. Hesthaven, E. Ochsner, *et al.*, *Phys. Rev. Lett.* **106**, 221102 (2011), arXiv:1101.3765 [gr-qc].
- [30] M. Pürrer, *Class. Quantum Grav.* **31**, 195010 (2014), arXiv:1402.4146 [gr-qc].
- [31] M. Pürrer, *Phys. Rev. D* **93**, 064041 (2016), arXiv:1512.02248 [gr-qc].
- [32] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera, *Comptes Rendus Mathématique* **339**, 667 (2004).
- [33] S. Chaturantabut and D. C. Sorensen, *SIAM Journal on Scientific Computing* **32**, 2737 (2010).
- [34] J. Blackman, S. E. Field, C. R. Galley, B. Szilágyi, M. A. Scheel, M. Tiglio, and D. A. Hemberger, *Phys. Rev. Lett.* **115**, 121102 (2015), arXiv:1502.07758 [gr-qc].

TABLE I. Deviations in waveform phase due to errors in the fits  $\Lambda_3^{\text{fit}}(\Lambda_2)$  and  $\Lambda_4^{\text{fit}}(\Lambda_2)$ . Waveforms are calculated with a mass ratio of  $q = 1$ . Phase differences are computed as  $|\Delta\phi(t)| = |\phi(t) - \phi^{\text{fit}}(t)|$  and the maximum difference up to the amplitude peak of the shorter waveform is listed.

EOS	M ( $M_\odot$ )	$\Lambda_2$	$\Lambda_3$	$\Delta\Lambda_3/\Lambda_3$	$\Lambda_4$	$\Delta\Lambda_4/\Lambda_4$	$ \Delta\phi $
BGN1H1	1.64	14	9.7	0.17	6.5	0.29	0.002
MPA1	2.47	3.05	1.1	-0.089	0.37	-0.19	0.0005
SLY	1.4	307	511	0.024	819	-0.017	0.003
MS1b	1.4	1260	3440	0.0001	9080	-0.075	0.007

- [35] C. R. Galley, F. Herrmann, J. Silberholz, M. Tiglio, and G. Guerberoff, *Class. Quantum Grav.* **27**, 245007 (2010), arXiv:1005.5560 [gr-qc].
- [36] K. Cannon, A. Chapman, C. Hanna, D. Keppel, A. C. Searle, and A. J. Weinstein, *Phys. Rev. D* **82**, 044025 (2010), arXiv:1005.0012 [gr-qc].
- [37] K. Cannon, C. Hanna, and D. Keppel, *Phys. Rev. D* **84**, 084003 (2011), arXiv:1101.4939 [gr-qc].
- [38] S. E. Field, C. Galley, and E. Ochsner, *Phys. Rev. D* **86**, 084046 (2012), arXiv:1205.6009 [gr-qc].
- [39] S. Caudill, S. E. Field, C. R. Galley, F. Herrmann, and M. Tiglio, *Class. Quantum Grav.* **29**, 095016 (2012), arXiv:1109.5642 [gr-qc].
- [40] J. Blackman, B. Szilagyi, C. R. Galley, and M. Tiglio, *Phys. Rev. Lett.* **113**, 021101 (2014), arXiv:1401.7038 [gr-qc].
- [41] K. Cannon, J. D. Emberson, C. Hanna, D. Keppel, and H. Pfeiffer, *Phys. Rev. D* **87**, 044008 (2013), arXiv:1211.7095 [gr-qc].
- [42] P. Canizares, S. E. Field, J. R. Gair, and M. Tiglio, *Phys. Rev. D* **87**, 124005 (2013), arXiv:1304.0462 [gr-qc].
- [43] P. Canizares, S. E. Field, J. Gair, V. Raymond, R. Smith, and M. Tiglio, *Physical Review Letters* **114**, 071104 (2015), arXiv:1404.6284 [gr-qc].
- [44] R. Smith, S. E. Field, K. Blackburn, C.-J. Haster, M. Pürrer, V. Raymond, and P. Schmidt, *Phys. Rev. D* **94**, 044031 (2016), arXiv:1604.08253 [gr-qc].
- [45] H. Antil, S. E. Field, F. Herrmann, R. H. Nochetto, and M. Tiglio, *Journal of Scientific Computing* **57**, 604 (2013).
- [46] K. Cannon *et al.*, *Astrophys. J.* **748**, 136 (2012), arXiv:1107.2665 [astro-ph.IM].
- [47] S. Bernuzzi, A. Nagar, T. Dietrich, and T. Damour, *Physical Review Letters* **114**, 161103 (2015), arXiv:1412.4553 [gr-qc].
- [48] T. Hinderer, A. Taracchini, F. Foucart, A. Buonanno, J. Steinhoff, M. Duez, L. E. Kidder, H. P. Pfeiffer, M. A. Scheel, B. Szilagyi, K. Hotokezaka, K. Kyutoku, M. Shibata, and C. W. Carpenter, *Phys. Rev. Lett.* **116**, 181101 (2016), arXiv:1602.00599 [gr-qc].
- [49] M. Kramer and N. Wex, *Class. Quantum Grav.* **26**, 073001 (2009).
- [50] R. S. Lynch, P. C. C. Freire, S. M. Ransom, and B. A. Jacoby, *Astrophys. J.* **745**, 109 (2012), arXiv:1112.2612 [astro-ph.HE].
- [51] T. Damour and A. Nagar, *Phys. Rev. D* **81**, 084016 (2010), arXiv:0911.5041 [gr-qc].
- [52] D. Bini and T. Damour, *Phys. Rev. D* **90**, 124037 (2014), arXiv:1409.6933 [gr-qc].
- [53] T. Damour and A. Nagar, *Phys. Rev. D* **80**, 084035 (2009), arXiv:0906.0096 [gr-qc].
- [54] S. Bernuzzi, A. Nagar, M. Thierfelder, and B. Bruggmann, *Phys. Rev. D* **86**, 044030 (2012), arXiv:1205.3403 [gr-qc].
- [55] D. Bini and T. Damour, *Phys. Rev. D* **87**, 121501 (2013), arXiv:1305.4884 [gr-qc].
- [56] E. Barausse, A. Buonanno, and A. Le Tiec, *Phys. Rev. D* **85**, 064010 (2012), arXiv:1111.5610 [gr-qc].
- [57] D. Bini and T. Damour, *Phys. Rev. D* **89**, 064063 (2014), arXiv:1312.2503 [gr-qc].
- [58] A. Nagar, T. Damour, C. Reisswig, and D. Pollney, *Phys. Rev. D* **93**, 044046 (2016), arXiv:1506.08457 [gr-qc].
- [59] The tidal interactions of a star (e.g. star A, we omit the labels in this footnote) in an external field are in general parametrized by (i) gravito-electric coefficients  $G\mu_\ell [length]^{2\ell+1}$  measuring the  $\ell$ th-order mass multipolar moment induced in the star by the external  $\ell$ th-order gravito-electric field; (ii) gravito-magnetic coefficients  $G\sigma_\ell [length]^{2\ell+1}$  measuring the  $\ell$ th-order spin multipolar moment induced in the star by the external  $\ell$ th-order gravito-magnetic field; (iii) shape coefficients  $h_\ell$  measuring the distortion of the surface of the star by an external  $\ell$ th-order gravito-electric field. In the literature, it is customary to consider only the dominant gravito-electric interactions (i) and, in most of the cases, only the leading order term  $\ell = 2$ . The dimensionless Love numbers are defined as  $k_\ell = (2\ell - 1)!! G\mu_\ell / (2R^{2\ell+1})$ , with  $R$  the star radius.  $\mu_2$  is simply called  $\lambda$  in [62, 65, 96]. Other notations are employed in [65] ( $\lambda \rightarrow \Lambda$ ) and in [67].
- [60] T. Hinderer, *Astrophys. J.* **677**, 1216 (2008), arXiv:0711.2420 [astro-ph].
- [61] T. Binnington and E. Poisson, *Phys. Rev. D* **80**, 084018 (2009), arXiv:0906.1366 [gr-qc].
- [62] T. Hinderer, B. D. Lackey, R. N. Lang, and J. S. Read, *Phys. Rev. D* **81**, 123016 (2010), arXiv:0911.3535 [astro-ph.HE].
- [63] D. Bini, T. Damour, and G. Faye, *Phys. Rev. D* **85**, 124034 (2012), arXiv:1202.3565 [gr-qc].
- [64] T. Damour, B. R. Iyer, and A. Nagar, *Phys. Rev. D* **79**, 064004 (2009), arXiv:0811.2069 [gr-qc].
- [65] J. E. Vines and E. E. Flanagan, *Phys. Rev. D* **88**, 024046 (2013), arXiv:1009.4919 [gr-qc].
- [66] T. Damour, A. Nagar, and L. Villain, *Phys. Rev. D* **85**, 123007 (2012), arXiv:1203.4352 [gr-qc].
- [67] K. Yagi, *Phys. Rev. D* **89**, 043011 (2014), arXiv:1311.0872 [gr-qc].
- [68] F. Douchin and P. Haensel, *Astron. Astrophys.* **380**, 151 (2001), arXiv:astro-ph/0111092.
- [69] H. Müller and B. D. Serot, *Nuclear Physics A* **606**, 508 (1996), arXiv:nucl-th/9603037.
- [70] J. S. Read, B. D. Lackey, B. J. Owen, and J. L. Friedman, *Phys. Rev. D* **79**, 124032 (2009), arXiv:0812.2163.

- [71] Effective-one-body (EOB) code developed at IHES: <https://eob.ihes.fr/>.
- [72] J. P. Boyd, *Chebyshev and fourier spectral methods* (Dover Publications, Inc., New York, 2000).
- [73] C. R. Galley and P. Schmidt, (in preparation).
- [74] <https://bitbucket.org/chadgalley/romspline>.
- [75] W. Hoffmann, *Computing* **41**, 335 (1989).
- [76] <https://bitbucket.org/sfield83/greedycpp>.
- [77] J. Kaye, *The interpolation of gravitational waveforms*, Thesis, Brown University, 2012, [www.dam.brown.edu/scicomp/reports/2013-8/](http://www.dam.brown.edu/scicomp/reports/2013-8/).
- [78] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, *Phys. Rev. Lett.* **113**, 151101 (2014), arXiv:1308.3271 [gr-qc].
- [79] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes: the art of scientific computing* (Cambridge University Press, Cambridge, UK, 2007).
- [80] “scipy.fftpack.dct,” Accessed 14 September 2016, <http://docs.scipy.org/doc/scipy/reference/generated/scipy.fftpack.dct.html#scipy.fftpack.dct>.
- [81] J. Aasi *et al.* (VIRGO, LIGO Scientific), (2013), 10.1007/lrr-2016-1, [Living Rev. Rel.19,1(2016)], arXiv:1304.0670 [gr-qc].
- [82] C. P. L. Berry, I. Mandel, H. Middleton, L. P. Singer, A. L. Urban, A. Vecchio, S. Vitale, K. Cannon, B. Farr, W. M. Farr, P. B. Graff, C. Hanna, C.-J. Haster, S. Mohapatra, C. Pankow, L. R. Price, T. Sidery, and J. Veitch, *Astrophys. J.* **804**, 114 (2015), arXiv:1411.6934 [astro-ph.HE].
- [83] B. Farr, C. P. L. Berry, W. M. Farr, C.-J. Haster, H. Middleton, K. Cannon, P. B. Graff, C. Hanna, I. Mandel, C. Pankow, L. R. Price, T. Sidery, L. P. Singer, A. L. Urban, A. Vecchio, J. Veitch, and S. Vitale, *Astrophys. J.* **825**, 116 (2016), arXiv:1508.05336 [astro-ph.HE].
- [84] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, *Publications of the Astronomical Society of the Pacific* **125**, 306 (2013), arXiv:1202.3665 [astro-ph.IM].
- [85] C. Pankow, P. Brady, E. Ochsner, and R. O’Shaughnessy, *Phys. Rev. D* **92**, 023002 (2015), arXiv:1502.04370 [gr-qc].
- [86] “LSC algorithm library,” <http://www.lsc-group.phys.uwm.edu/la1>.
- [87] J. Veitch and A. Vecchio, *Phys. Rev.* **D78**, 022001 (2008), arXiv:0801.4313 [gr-qc].
- [88] J. Veitch and A. Vecchio, *Proceedings, 12th Workshop on Gravitational wave data analysis (GWDAW-12): Cambridge, USA, December 13-16, 2007*, *Class. Quant. Grav.* **25**, 184010 (2008), arXiv:0807.4483 [gr-qc].
- [89] J. Veitch and A. Vecchio, *Phys. Rev.* **D81**, 062003 (2010), arXiv:0911.3820 [astro-ph.CO].
- [90] J. Veitch *et al.*, *Phys. Rev.* **D91**, 042003 (2015), arXiv:1409.7215 [gr-qc].
- [91] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, and *et al.*, *ArXiv e-prints* (2016), arXiv:1606.01210 [gr-qc].
- [92] R. O’Shaughnessy and M. Rizzo, *In progress* (2016).
- [93] S. Bernuzzi and *et al.*, *In preparation* (2016).
- [94] T. Hinderer, J. Steinhoff, and A. Taracchini, *In progress* (2016).
- [95] J. A. Clark, A. Bauswein, N. Stergioulas, and D. Shoemaker, *Classical and Quantum Gravity* **33**, 085003 (2016), arXiv:1509.08522 [astro-ph.HE].
- [96] E. E. Flanagan and T. Hinderer, *Phys. Rev.* **D77**, 021502 (2008), arXiv:0709.1915 [astro-ph].