Enhancing gravitational waveform models through dynamic calibration

Yoshinta Setyawati,[*] Frank Ohme, and Sebastian Khan

Max Planck Institute for Gravitational Physics (Albert Einstein Institute),
Callinstr. 38, 30167 Hannover, Germany
and Leibniz Universität Hannover, D-30167 Hannover, Germany

Current gravitational-wave observations made by Advanced LIGO and Advanced Virgo use theoretical models that predict the signals generated by the coalescence of compact binaries. Detections to date have been in regions of the parameter space where systematic modeling biases have been shown to be small. However, we must now prepare for a future with observations covering a wider range of binary configurations, and ever increasing detector sensitivities placing higher accuracy demands on theoretical models. Strategies to model the inspiral, merger and ringdown of coalescing binaries are restricted in parameter space by the coverage of available numerical-relativity simulations, and when more numerical waveforms become available, substantial efforts to manually (re)calibrate models are required. The aim of this study is to overcome these limitations. We explore a method to combine the information of two waveform models: an accurate, but computationally expensive *target* model, and a fast but less accurate *approximate* model. In an automatic process we systematically update the basis representation of the approximate model using information from the target model. The result of this process is a new model which we call the *enriched basis*. This new model can be evaluated anywhere in the parameter space jointly covered by either the approximate or target model, and the enriched basis model is considerably more accurate in regions where the sparse target signals were available. Here we show a proof-of-concept construction of signals from nonprecessing, spinning black-hole binaries based on the phenomenological waveform family. We show that obvious shortcomings of the previous PhenomB being the approximate model in the region of unequal masses and unequal spins can be corrected by combining its basis with interpolated projection coefficients derived from the more recent and accurate PhenomD as the target model. Our success in building such a model constitutes an major step towards dynamically combining numerical relativity data and analytical waveform models in the computationally demanding analysis of LIGO and Virgo data.

I. INTRODUCTION

The dawn of the gravitational-wave (GW) era began with the first detection of a binary-black-hole (BBH) merger on September 14, 2015 [1] by the Advanced Laser interferometer Gravitational-wave Observatory aLIGO [2]. More BBH [3–6] and one binary neutron star (BNS) merger on August 17, 2017 [7] have been observed by aLIGO and Virgo [8] during their first two observing runs.

The search for GWs requires coincident signals in at least two instruments. In order to uncover signals of astrophysical origin hidden behind the instruments' noise, their data are filtered with a large number of waveform templates [9]. More than one hundred thousand templates of coalescing compact binaries were employed in aLIGO GW searches during each of the first two observing runs.

An order of magnitude more modelled waveforms are then used to estimate the source parameters and their uncertainties. More accurate and efficient follow ups of GW detections and their parameters will be needed for the following aLIGO observing runs. This implies the need for waveform models covering a wide range of parameter space that can be generated quickly.

The GW signal emitted by coalescing binaries depends upon many different parameters that are often grouped into intrinsic and extrinsic parameters. Intrinsic parameters are astrophysical parameters of the binary. These are two mass parameters: the chirp mass ($\mathcal{M}_c$) and the symmetric mass ratio η; eccentricity; tidal parameters for neutron stars; and the spin components of the two objects ($\vec{\chi}_1, \vec{\chi}_2$) that are often represented by the dominant, effective spin parameter (χ_{eff}) in the case of non-precessing binaries. The exact definition of these parameters will be introduced in Sec. II.

In this study we focus on nonprecessing BBHs for which the spins are (anti)aligned with the binary's orbital

---

[*]Corresponding author.
yoshinta.setyawati@aei.mpg.de

angular momentum. The dimensionless tidal parameters are set to zero. Eccentricity has also been neglected in all aLIGO GW searches that employ modelled templates so far, mainly because for most plausible astrophysical formation scenarios, the binary is expected to have circularized by the time its GW signal enters the aLIGO frequency range. However, future waveform developments might include the eccentricity of the binary.

In addition to these properties, extrinsic parameters define the location and orientation of the source relative to the observer, such as the luminosity distance (D_L), inclination angle (ι), sky position (RA, Dec), polarization angle (Ψ), time of coalescence (t_c), and phase of coalescence (ϕ_c) [10]. For nonprecessing systems, modifications in these parameters simply shift the waveform in time, phase or amplitude, and they are much simpler to model than changes in intrinsic parameters.

In order to predict GW signals from binaries, one needs to solve the Einstein equation in general relativity (GR). Analytical approximations have been established in form of post-Newtonian (PN) expansions. These are asymptotic expansions in a small parameter such as the ratio of the characteristic velocity of the binary to the speed of light [11,12]. By the nature of the approximation, PN expansions become increasingly inaccurate as the two bodies move closer to each other and faster, entering the strong gravity regime. At this stage, numerical relativity (NR) simulations provide the only viable approach to solve the Einstein equation [13]. In general, NR waveforms can in principle be very accurate and the accuracy can be tested through different types of convergence tests, but they are computationally extremely expensive [14–17].

Hence, many efforts in the past focused on bridging PN and NR [18,19], leading to a variety of effective-one-body (EOB) and phenomenological waveform models that are used in aLIGO's analyses. EOB is an analytical method proposed by Buonanno and Damour [20–27] which substantially reformulates PN results into a new description of the binary coalescence beyond the inspiral phase. A different approach was developed to build phenomenological models (see II A) that essentially model coalescing binaries using analytical fits of PN-NR hybrids.

However, both approaches depend on a number of tunable parameters and fits whose optimal form and values are determined through complex procedures that typically require a fair amount of human input. Therefore, updated models that incorporate new NR data and improved analytical descriptions typically take years to develop.

A different method to generate an accurate waveform model is based on sophisticated interpolation methods to create a *surrogate* model [28–30] of NR waveforms. These surrogate models have a high accuracy to the original NR waveforms, however, they are limited to the parameter space covered by the original simulations. Although boundaries are constantly being expanded in parameter space, this modeling strategy relies on large amounts of computational power. At the time of writing this article, the latest precessing surrogate model [29] is limited in mass ratio and dimensionless spin magnitude to $q \leq 2$ and $|\chi| \leq 0.8$, respectively.

Here we explore a complementary method of constructing a waveform model that combines the information of an existing (computationally efficient) model with more accurate waveforms that are only available in a limited set of points in the parameter space. A future application of our method would be a dynamical (i.e., fully automized) update of an analytical model with NR waveforms to produce a new waveform model that can be evaluated continuously and has a better accuracy than the original model.

To develop our method, here we employ two analytic phenomenological models: PhenomB [31] being the *approximate, less accurate* model and PhenomD [32,33] being the *target, more accurate* model.

We use singular value decomposition (SVD) to decompose the approximate model into an orthogonal basis and update the basis coefficients using information from the more accurate model. Similar ideas of using SVD to improve waveform models have been presented by Cannon *et al.* [34–36] and Pürrer [37,38].

Cannon *et al.* explore the use of reduced-order SVD in time domain. However, they only use one-dimensional interpolation in mass components and consider a restricted parameter space with no spin. We use a similar technique, but consider frequency-domain waveforms, and we extend the method to a much greater parameter space including spin.

Pürrer discusses the use of SVD to build computationally more efficient reduced-order models (ROMs) of existing spinning, nonprecessing EOB models. ROMs are now a standard tool to reduce the time taken to generate a waveform, but the resulting accuracy is that of the original model, or slightly less in challenging points of the parameter space.

Throughout this article geometric units are used by setting $G = c = 1$.

II. METHODOLOGY

A. Waveform models

The constantly increasing sensitivity of GW interferometers demands ever more accurate models. Updating and improving models is a major tasks entering the era of GW astronomy, and we present a first end-to-end test of a fully automatic tuning that in future will use NR simulations to improve analytical models. Here, however, we start with a proof-of-concept using two phenomenological waveform models.

The phenomenological family is a set of approximate waveform models, written as closed-form analytical expressions in the frequency domain [31,33,39–42]. These models have been calibrated to NR waveforms that

naturally cover a limited region of the intrinsic parameter space. However, the most recent models [33,41] have been shown to be perfectly suited for current BBH observations with mass ratios close to unity.

As explained in Sec. I, we use PhenomB as the *approximate, less accurate model* that we aim to update with information from PhenomD as the *accurate target model*. PhenomB [31] was the first (anti)aligned spin model of this family, released almost simultaneously with an alternative description of the same parameter space, called PhenomC [42]. Both models were calibrated up to mass ratios of 4 and black hole (BH) spins up to 0.75. They have known shortcomings when extrapolating beyond the region of calibration, especially towards more extreme mass ratios. PhenomD is the most recent and most sophisticated version of aligned-spin phenomenological models. It has been calibrated to 19 NR waveforms from the SXS collaboration [43] and the BAM code [44,45] that span mass ratios from unity up to 18 and dimensionless spin magnitudes up to 0.85 (0.98 for equal-mass systems) [32,33].

The intrinsic parameters of relevance in the nonprecessing case are the mass ratio q, or equivalently the symmetric mass ratio η,

$$q = \frac{m_1}{m_2} \geq 1, \qquad \eta = \frac{m_1 m_2}{(m_1 + m_2)^2}, \qquad (1)$$

as well as the dimensionless spin projections along the orbital angular momentum χ_{1z}, χ_{2z} (nonvanishing spin components perpendicular to the orbital angular momentum cause precession effects that we leave for future work). For vacuum solutions of Einstein's equation, the total mass $M = m_1 + m_2$ is a simple scaling factor.

We emphasize that the spin degrees of freedom in a binary are commonly reduced in phenomenological models to the observationally relevant dominant parameter combinations. Following the analysis in [31,42] for aligned-spin binaries, the dominant spin effect in GW phase can be expressed as the weighted combination of individual BH spins,

$$\chi_{\text{eff}} = \frac{m_1 \chi_{1z} + m_2 \chi_{2z}}{m_1 + m_2}. \qquad (2)$$

Apart from an overall time and phase, PhenomB exclusively depends on χ_{eff} and η. PhenomD uses χ_{eff} for the coefficients that were tuned to NR simulations, however, through the inspiral and the final state portion of PhenomD inherits two-spin dynamics. In Sec. III B we will apply our method to the 3D problem (η, χ_1, χ_2) and express our results in terms of the symmetric (χ_{eff}) and antisymmetric (χ_a) spin parameters where (χ_a) is defined as

$$\chi_a = \frac{\chi_{1z} - \chi_{2z}}{2}. \qquad (3)$$

In the following sections, we present the details of how to update PhenomB with the more accurate PhenomD waveforms in frequency domain for a given range of η, χ_{eff} and scaled by the total mass M. The end result of this computation is a new waveform model that is closer to its target waveforms. We call this new family as the *enriched basis (EB)* waveforms.

B. Parameter ranges

This exploratory study is designed to test our method across a wide range in parameter space. Here, we essentially consider the range in mass ratio and spins where PhenomD has been calibrated to NR waveforms (see Sec. II A),

$$\eta \in [0.05, 0.25], \qquad \chi_{\text{eff}} \in [-1, 1]. \qquad (4)$$

We stress that this region in the parameter space includes a considerable part where PhenomB has not been calibrated, e.g., mass ratios above 4 ($\eta < 0.16$). What we are going to show is that despite the fact that the underlying approximate model does not accurately describe signals in certain regions, using accurate signals to update the approximate basis representation can entirely fix that problem.

In order to fully determine the signals for our test case, we fix the following additional parameters,

$$M = 50 \text{ M}_\odot, \qquad f_{\text{low}} = 30 \text{ Hz},$$
$$\Delta f = 0.1 \text{ Hz}, \qquad M f_{\text{high}} = 0.2, \qquad (5)$$

where f_{low} and f_{high} are the values of the lowest and the highest frequency we consider, respectively, and Δf defines the numerical discretization of the signal. $M f_{\text{high}} = 0.2$ is chosen to be slightly higher than the signal with the largest ringdown frequency in our dense grid.[1] For $M = 50 \text{ M}_\odot$, f_{high} corresponds to 812 Hz.

Following the above choice of parameter ranges, we create two two-dimensional (2D) uniform grids in η and χ_{eff}. We build a *dense grid* of approximate PhenomB waveforms, and a *sparse grid* of accurate PhenomD waveforms (see Fig. 1 for visual representation). Our dense grid contains $N = 65 \times 65 = 4225$ signals, and the sparse grid has $S = 33 \times 33 = 1089$ signals. Thus, about 25% of the approximate waveforms have the same η and χ_{eff} as the target waveforms.

On each point of each grid, we generate the GW polarizations, $h_{+/\times}$. In this work, we only consider non-precessing signals and their $(\ell, |m|) = (2, 2)$ multipoles

---

[1]The system with the highest ringdown frequency will be the equal-mass, maximally spinning case ($\chi_{\text{eff}} = 1$) which has dimensionless ringdown frequency of ∼0.13.
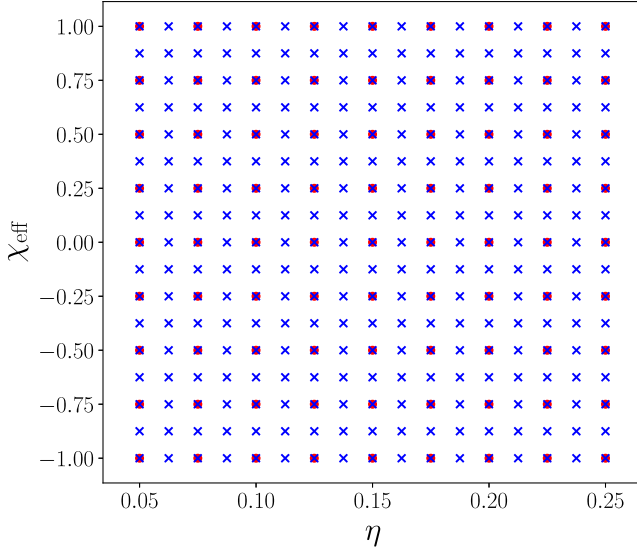
FIG. 1. Illustration of the two uniform grids we consider in η-χ_{eff} parameter space. The blue crosses illustrate the *dense grid* of approximate signals that we use to build an SVD basis, and the red circles are the *sparse grid* of accurate signals we use to update the model.

which means that the extrinsic parameters, such as the orientation and location of the binary, simply scale the amplitude of the signal and introduce a constant phase shift. We can treat these trivial dependencies independently and, at this stage, normalize all waveforms to have the same extrinsic parameters. We then use the software library LALSUITE [46] to generate the GW polarizations.

In this study, both approximate and target models are inexpensive to compute, so we can test our method for large numbers of target-model waveforms. In the future, we will use target waveforms that come from computationally expensive methods such as NR simulations. In that situation we may not have access to signals at arbitrary points in parameter space and we will have fewer waveforms. Here we first choose a reasonably high number of target waveforms, and later discuss how low this number can become to still produce satisfactory results.

C. Waveform matrices

In our method we will represent the waveform manifold of the *approximate* model with a set of orthogonal basis functions computed using SVD. First, we prepare our data set in appropriate matrix form which we can factorize subsequently. The procedure is explained by the following steps.

1. Waveform decomposition

The frequency-domain strain $\tilde{h}(f)$ is the combination of both GW polarizations, where f is defined for positive frequencies. Here we assume the circular polarizations of GW and describe $\tilde{h}(f)$ as follows[2]:

$$\tilde{h}(f) = \tilde{h}_+(f) + i\tilde{h}_\times(f). \tag{6}$$

We note that if we express $\tilde{h}_{+/\times}(f)$ in terms of their amplitudes, $\mathcal{A}_{+,\times}$, and phases, $\Psi_{+,\times}$, factoring out the dependency on the inclination angle, ι, we obtain the following expressions [10]:

$$\tilde{h}_+(f) = \mathcal{A}_+(f)e^{i\Psi_+(f)}\left(\frac{1+\cos^2\iota}{2}\right), \tag{7}$$

$$\tilde{h}_\times(f) = \mathcal{A}_\times(f)e^{i\Psi_\times(f)}\cos\iota. \tag{8}$$

The nonprecessing signals we consider further satisfy a simple relation between the polarizations,

$$\mathcal{A}_+ = \mathcal{A}_\times, \qquad \Psi_\times = \Psi_+ - \frac{\pi}{2}. \tag{9}$$

While (9) is exactly valid only in the limit of large separations, assuming it through merger and ringdown is a commonly made approximation that does not introduce inaccuracies relevant to today's analyses.

By computing $\tilde{h}_+$ and $\tilde{h}_\times$ for $\iota = 0$, we can now decompose $\tilde{h}(f)$ into amplitude and phase components,

$$\tilde{h}(f) = 2\mathcal{A}_+(f)e^{i\Psi_+(f)}. \tag{10}$$

In this form, we can focus on two real-valued functions: the strain's amplitude and phase (we drop the "+" subscript henceforth). This decomposition is convenient because amplitude and phase are simpler, real-valued, nonoscillatory functions which are better suited to perform SVD than the oscillating strain.

Once we have constructed the improved EB amplitude $\mathcal{A}_{\text{EB}}(f)$ and phase $\Psi_{\text{EB}}(f)$, we can combine them again into the EB strain $\tilde{h}^{\text{EB}}(f)$, as well as individual polarizations, using Eqs. (7)–(10).

2. Phase alignment

Time and phase shifts enter the frequency-domain waveform through the GW phase $\Psi(f)$ according to

$$\Psi'(f) = \Psi(f) + 2\pi f t + \psi, \tag{11}$$

where t is the amount of time shifted and ψ is the phase shift.

We use (11) to align the phases in our approximate waveform grid by determining the time and phase shift individually for each configuration such that the square phase difference with one fiducial case is minimized.

---

[2]In other literature, the strain is sometimes defined as $h_+ - ih_\times$, owing to a different convention of the Fourier transform. Here we adopt the definition used in LIGO algorithm library (LAL), $\tilde{h}(f) = \int h(t)\exp(-i2\pi ft)dt$

Specifically, we align the phases against the first case in our grid ($\eta = 0.05$, $\chi_{\text{eff}} = -1$), although any other choice yields comparable results. By aligning the phases before performing the SVD we remove variations between the phases that are purely due to time and phase shifts. These variations can always be reintroduced analytically via (11). The shifted phase is denoted by $\Psi_B(f)$.

3. Matrix form

We generate N signals from the approximate model PhenomB, each discretized at $L > N$ points in frequency domain between f_{low} and f_{high}. After computing the strain $\tilde{h}(f)$ from the two polarizations as explained in the previous section, we decompose them into amplitude and phase then align the phases. We then pack all amplitude and phase arrays into a matrix, respectively (see Fig. 2) Specifically, the rows of the matrices are arranged from the lowest (η, χ_{eff}) to the highest (η, χ_{eff}).

We repeat the above procedure and generate S target waveforms, using PhenomD, on the sparse grid, where $S < N < L$. At this point, we have four matrices: two amplitude matrices and two phase matrices; one of each type for each approximant. The matrices of the approximate model PhenomB have the dimensions $\mathbb{R}^{N \times L}$ while the target PhenomD model matrices are $\in \mathbb{R}^{S \times L}$. With this prepared, we perform an SVD as discussed in subsection II D.

D. The singular value decomposition

Our goal is to generate a new waveform family that can be evaluated for arbitrary parameters from interpolating a set of sparse target waveforms. To do this, we project our target model onto a basis of the approximate model, generated from a grid that is as dense as possible and computationally feasible. As a first step, the basis is built by an appropriate factorization of the grid of the approximate waveforms.

There are two main strategies to factorize sets of waveforms. One uses a Gram-Schmidt orthogonalization to
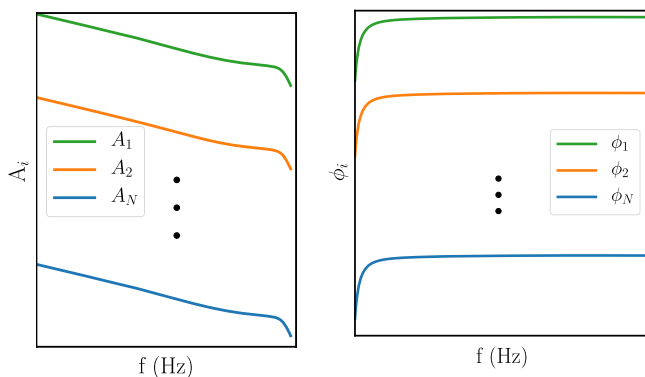

FIG. 2. Illustration of N signals, each of length L, decomposed into amplitudes (left) and phases (right) packed into two matrices.

obtain the basis from a first set of approximate waveforms followed by a greedy algorithm to extend the basis until an acceptable error limit is reached [29,47,48]. The second strategy uses the SVD as in Cannon *et al.* [34–36] and Pürrer [37,38,49] to factorize each matrix into two unitary matrices and one diagonal matrix with elements sorted in descending order. The comparison between the two strategies has been discussed in [29]. Here we use the SVD because it produces smoother result, and because it is elegant and convenient given that it sorts the contribution from the dominant basis vector to the least important ones. This ensures that the error caused by SVD truncation is generally small.

We adopt SVD to individually factorize amplitude and phase matrices (P) of PhenomB into two unitary matrices (U and V) and one diagonal matrix Σ [50],

$$P = U\Sigma V^T. \tag{12}$$

Here, $U = [u_1|\ldots|u_p] \in \mathbb{R}^{N \times p}$ and $V = [v_1|\ldots|v_p] \in \mathbb{R}^{L \times p}$ are orthogonal matrices and the superscript T denotes the transpose of the corresponding matrix. The vectors u_i and v_i are left and right singular vectors of P respectively. The singular values $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_p) \in \mathbb{R}^{p \times p}$ is a diagonal matrix sorted in descending order, where $p = \min(N, L)$, which in our setup yields $p = N$. The diagonal elements σ_i^2 are the eigenvalues of $P^T P$.

SVD can be interpreted as matrix decomposition into a weighted sum of separable matrices, meaning that a matrix P can be written as an outer product of two vectors $P = \bar{u} \otimes v^T$ ($\bar{u}$ denote the u vectors weighted by the singular values). The rank of this outer product depends on how many singular values are involved in the sum. The index notation of the above reads

$$P_{ij} = \sum_{k=1}^{p} u_{ik}\sigma_k v_{kj}^T. \tag{13}$$

E. Projection coefficients and reduced order

In our study, we use Eq. (13) in the following way. Every row of the matrix P_{ij} represents a Fourier-domain series of either amplitude or phase; the index j represents individual frequency samples. Every one of those Fourier-domain series is expressed on the right-hand side as a linear combination of orthogonal basis vectors $(V^T)_{kj}$ (k is the index of the basis, j specifies the frequency) multiplied with coefficients $c_{ik} = u_{ik}\sigma_k$ (k corresponds to the associated basis, i specifies the frequency series that is reconstructed in this way). We call c_{ik} the projection coefficients. The projection coefficients can be interpreted as updating the left singular vectors u_{ik} weighted by the rank of singular value σ_k.

In order to build an analytical model that can be evaluated continuously across the parameter space, the

projection coefficients need to become functions that interpolate in the parameter space between the discrete points given in the rows of P_{ij}. We emphasize this below by replacing the index i with the explicit functional dependency on η and $\chi_{\rm eff}$, leading to

$$c_k(\eta, \chi_{\rm eff}) = \sum_{j=1}^{L} P_j(\eta, \chi_{\rm eff}) V_{jk}. \qquad (14)$$

The sum now describes a discretized inner product $\langle \cdot, \cdot \rangle$, so that (14) becomes

$$c_k(\eta, \chi_{\rm eff}) = \langle P(\eta, \chi_{\rm eff}), v_k \rangle. \qquad (15)$$

Again, P in this expression represents either the amplitude or phase for the parameters $(\eta, \chi_{\rm eff})$, v_k are the basis vectors calculated via SVD.

Different from standard practice in SVD and ROM, we now proceed by calculating coefficients from projecting the *target* waveforms' amplitude and phase onto the basis representation of the *approximate* waveform, respectively. In addition, we study a reduction of the basis size that is achieved by only considering the first K coefficients. K then reduces the rank of the singular values matrices [36], and it enters (13) as the upper limit of the sum instead of p. This reduced order is introduced to increase computational efficiency and to decrease memory requirements when building the EB in comparison to the full basis $k = N$.

By updating the approximate (less accurate) waveforms basis coefficients with information from the (more accurate) target waveforms we have manipulated the basis representation of approximate waveforms to be closer to target waveforms. Hence, we name this process *enriching the basis*.

F. Interpolation

To construct our enriched basis model, we calculate the approximate SVD basis and project the target amplitude and phase onto the respective basis vectors, giving us projection coefficients according to (14) on the sparse grid in parameter space (recall, the sparse grid is where we have access to accurate target signal). We then interpolate the projection coefficients and calculate their values on all points on the dense grid, so that we can compare with all approximate signals that we needed to start this process.

We stress that the dimensionality of the interpolation depends on the target model. For equal-spin case, we use two-dimensional interpolation $(\eta, \chi_{\rm eff})$ in parameter space, and later we consider two independent spins, where we need three-dimensional interpolation. Here we employ cubic spline interpolation as the most efficient and easy method for this project. However, different interpolation methods such as Chebyshev polynomials [36], tensor product interpolation [37], Gaussian interpolation [51]

and empirical interpolation [28] have been used in different studies. For the future, it will be beneficial to compare all these methods systematically, evaluating computational efficiency, accuracy and generalizability to higher dimensions.

Once the target waveform's coefficients, that we denote by $c'(\eta, \chi_{\rm eff})$, have been obtained, we combine them with the basis vectors to calculate the EB's amplitude and phase,

$$P_j^{\rm EB}(\eta, \chi_{\rm eff}) = \sum_{k=1}^{K} c'_k(\eta, \chi_{\rm eff}) v_{kj}^T. \qquad (16)$$

Having amplitude and phase, we can build $\tilde{h}^{\rm EB}(\eta, \chi_{\rm eff})$ using Eq. (10).

G. Match and improvement evaluation

Once the EB strains $\tilde{h}^{\rm EB}$ have been calculated, we evaluate their accuracy and improvement of EB model relative to its approximate and target models. We then test the accuracy of EB model both at points where the target model was used to update the projection coefficients, as well as at points where no target signals were available and we use the interpolated projection coefficients. To perform the evaluation, we compute *matches* between PhenomB and PhenomD and compare them to the matches between EB and PhenomD.

The match is defined as the normalized, noise-weighted inner product between two waveforms h_1 and h_2 [10], maximized over relative time and phase shifts between them,

$$\mathcal{O} = \frac{\langle h_1, h_2 \rangle}{\|h_1\| \|h_2\|}$$
$$= \max_{\phi_0, t_0} \left[4 {\rm Re} \int_{f_1}^{f_2} \frac{\tilde{h}_1(f) \tilde{h}_2^*(f)}{S_n(f)} \frac{df}{\|h_1\| \|h_2\|} \right]. \qquad (17)$$

Here, ϕ_0 and t_0 are relative phase and time shifts between the waveforms, respectively, and $\|h\|^2 = \langle h, h \rangle$. $S_n(f)$ is the noise spectral density of the detector, $\tilde{h}^*$ denotes the complex conjugation of $\tilde{h}$, and (f_1, f_2) is a suitable integration range which corresponds to $f_{\rm low}$ and $f_{\rm high}$ respectively. We use two noise spectra in our analysis, flat noise $(S_n \equiv 1)$ and the aLIGO zero detuned high power density (AZDHP) which is the anticipated design sensitivity of aLIGO in 2020 or later [52]. The motivation behind using a flat power spectral density (PSD) is to evaluate the signal agreement with equal weight on all frequencies independent of an assumed instrument, whereas using AZDHP allows us to relate our results to GW analysis applications.

Matches are close to unity where waveforms agree (see Sec. III), so it is easier to compare the difference between two models by quoting the mismatch, defined as

$$\mathcal{M}(h_1, h_2) = 1 - \mathcal{O}(h_1, h_2). \qquad (18)$$

Finally, to quantify the accuracy improvement of the EB model over the approximate model PhenomB, we define improvement, $\mathcal{I}$, as the mismatch of the approximate waveform with the target model divided by the mismatch of EB model with the target,

$$\mathcal{I}(h_1, h_2) = \frac{\mathcal{M}(h_1, h_3)}{\mathcal{M}(h_2, h_3)}, \qquad (19)$$

where in this study, h_1, h_2 and h_3 correspond to PhenomB, EB, and PhenomD respectively.

III. RESULTS

We have outlined a technique to build a more accurate waveform model in the above section. Here we present results and analyses based on two different assumptions about the spins in the target parameter space, *equal-spin* ($\chi_{\mathrm{eff}} = \chi_1 = \chi_2$) and *double-spin*, where χ_1 and χ_2 are varied independently (i.e., χ_a does not necessary vanish).

A. Two dimensions: Equal-spin systems

Following the above procedure, we evaluate the match between the EB model and the target model under flat noise and AZDHP. We also compare the mismatch between the approximate model against the target model to calculate the improvement we gain.

Figure 3 shows the original match of PhenomB against PhenomD. It is evident that PhenomB has not been calibrated to mass ratios above 4, and the agreement between the two models deteriorates quickly, especially for high spins.

Figure 4 presents the matches of EB against PhenomD without invoking any interpolation. Recall, EB here is based on basis vectors derived from PhenomB that do not accurately represent high-mass ratio systems. However, by
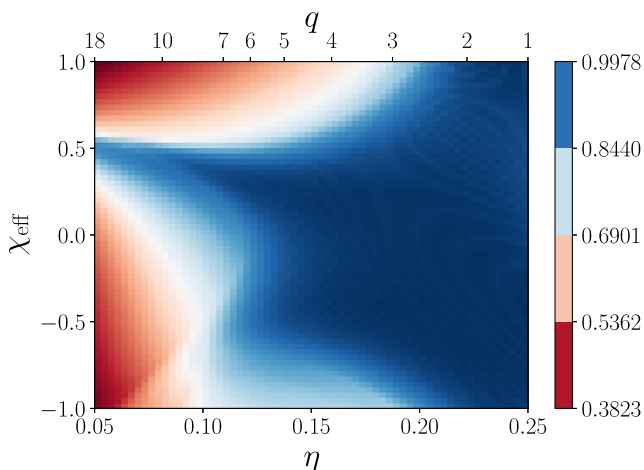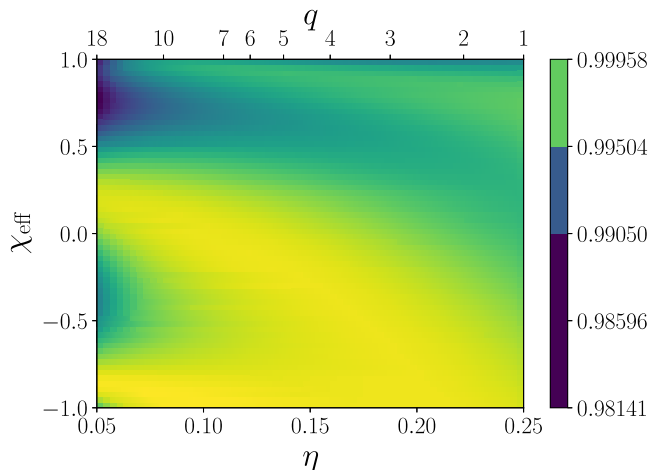

FIG. 4. Matches of EB against PhenomD without interpolation and under flat PSD. In this figure, we generated target model in the same grid as the approximate model, and run our method in full bases (without reduced order). This plot is used as comparison to interpolation and reduced order result as explained in the text.

projecting $N = 65 \times 65$ PhenomD waveforms onto the basis derived from N PhenomB signals on the same points in parameter space, we see that there is enough extra freedom in the basis such that updated projection coefficients can correct for the inaccuracies of the approximate model. Put differently, the space spanned by the approximate PhenomB basis vectors does contain more accurate signals, also for higher mass ratios, if the coefficients in front of the basis vectors are adapted appropriately. This might not be a surprising result, given the fairly large number of basis vectors we use; it is not a trivial result either.

Of course, this is not a useful application of the method we develop. If one has access to N accurate waveforms, there is no need build an approximate basis first. Now we reduce the number of accurate waveforms to $S \approx N/4$, and interpolate the projection coefficients to calculate EB signals on all N grid points. The mismatch result is shown in Fig. 5. In most parts of the parameter space, the accuracy of EB is only very slightly lower than what was achieved in the ideal scenario shown in Fig. 4. Interpolation therefore does not introduce significant errors for the grids chosen here. Only at the boundaries of the parameter space we find higher mismatches in Fig. 5.

We note that interpolation will likely become a major source of error when the number of available target waveforms is decreased significantly and when the dimensionality of the parameter space increases. We shall return to discussing both issues later in this paper.

We have repeated the study with the AZDHP noise curve and find qualitatively the same behavior. A summary of mismatches (in $\log_{10}$ scale) and improvements are given in Table I. We present the minimum, maximum and median

FIG. 3. Matches of PhenomB against PhenomD under flat PSD.
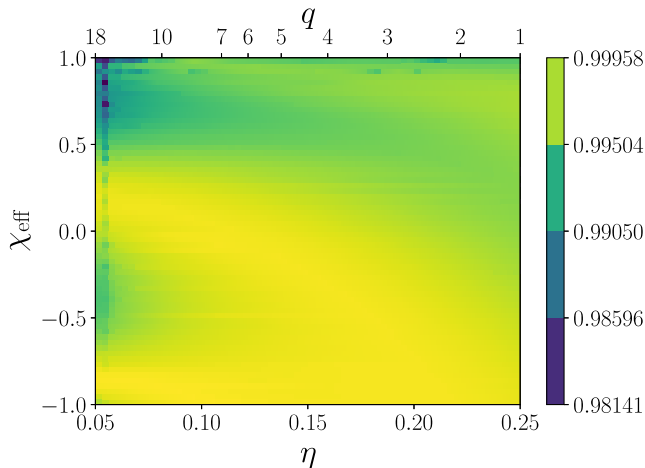
FIG. 5. Matches between EB against PhenomD with interpolation and under flat PSD. In this figure, we generated both target and approximate models in regular grid. The number of target model is about 25% of the approximate model as explained in subsection II B. We perform two dimensional interpolation (see subsection II F) over the projection coefficients. To make the comparison easier, we set the range of match equal as that on Fig. 4.

mismatches across the dense grid, as well as the improvements defined by (19).

Table I shows that overall the difference of mismatches using one or the other noise spectrum is relatively small. Full histograms are shown in Fig. 6. Because results are so similar, we only show the figures for the flat PSD.

So far, we have generated our target model, PhenomD, on a regular grid in the parameter space as illustrated in Fig. 1. We also investigate how the choice of positions of target signals affect our result. For that reason, we distribute the same number of PhenomD waveforms randomly, drawn uniformly from the parameter space of η and χ_{eff}. These target waveforms are then projected onto basis vectors coming from the dense regular grid of PhenomB signals.

TABLE I. Mismatches between PhenomB and PhenomD as well as mismatches between EB and PhenomD in $\log_{10}$ scale. The improvement, $\mathcal{I}$, is defined by (19). Here we compare the results using two different PSD, flat PSD ($S_n = 1$) and AZDHP. We also compare results that interpolate from the sparse to the dense grid with calculations entirely carried out on the dense grid (no interpolation).

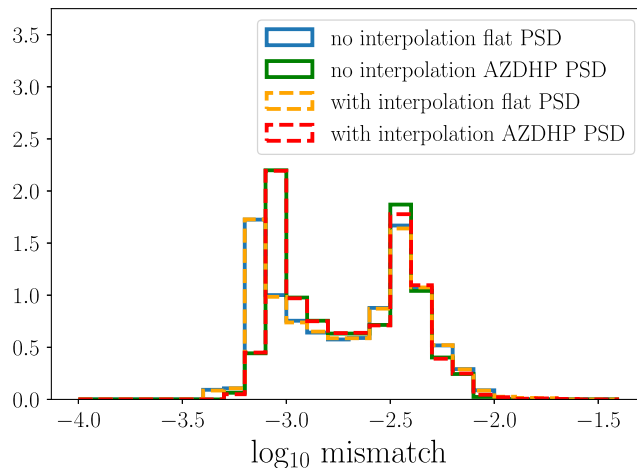| PSD		No interpolation			Interpolation		
		Min	Max	Med	Min	Max	Med
Flat	PhenomB	−2.67	−0.001	−0.03			
	EB	−3.37	−1.99	−2.69	−3.38	−1.73	−2.68
	$\mathcal{I}$	1.42	1201	40	1.42	1195	39
AZDHP	PhenomB	−2.50	−0.10	−1.14			
	EB	−3.23	−1.95	−2.71	−3.23	−1.68	−2.71
	$\mathcal{I}$	1.17	1082	39	1.20	1082	39

FIG. 6. Mismatches between EB and PhenomD target signals for different configurations and PSDs. The histograms are normalized so that the sum of area under each line are set equal to unity. The dashed lines represent the result using fewer target signals and interpolation, whereas the respective solid lines show results using more target signals and no interpolation (see text).

We follow the above procedure to build the EB coefficients and interpolate them onto the dense regular grid to evaluate mismatches between EB and PhenomD waveforms with the same parameters. Since the results for flat and AZDHP PSDs are relatively close, we evaluate the mismatch assuming a flat PSD. We find that $\log_{10}\mathcal{M}$ of random uniform grid ranges between −1.39 and −3.39. For direct comparison, the mismatch of the regular grid of target waveforms is between −1.73 and −3.38 as presented in Table. I. From this simple study, we argue that different positions will not affect the result significantly, so long as the number and distribution of parameters are similar.

1. Accuracy of the reduced basis

Here we examine the accuracy of EB when restricting ourselves to the first K bases. The advantage of a reduced basis is mainly to optimize computational power.

Figure 7 shows the mismatches and improvements as a function of the number of bases that are kept from the SVD of PhenomB. To obtain the result, we projected $S \approx N/4$ PhenomD signals onto the PhenomB basis and performed interpolation as explained in previous section. For very small numbers of bases we observe a rapid drop in mismatches. After the first 25 bases are included, however, the improvement of EB is much more gradual when more bases are used. We speculate that the most important variations in PhenomB signals are already well described with 25 basis vectors, but we do need a lot more basis vectors to accommodate additional features present in PhenomD that are not captured accurately by PhenomB (most notably, the high mass ratio, high spin regime).

If our goal is that the EB signals are at least as accurate as the approximate model, and in most points of parameter

FIG. 8.   Only 144 (12×12) PhenomD signals on a uniform grid are needed to guarantee that all EB waveforms perform better than PhenomB (assuming flat noise). The blue line is the value of the minimal improvement using $r \times r$ target waveforms.
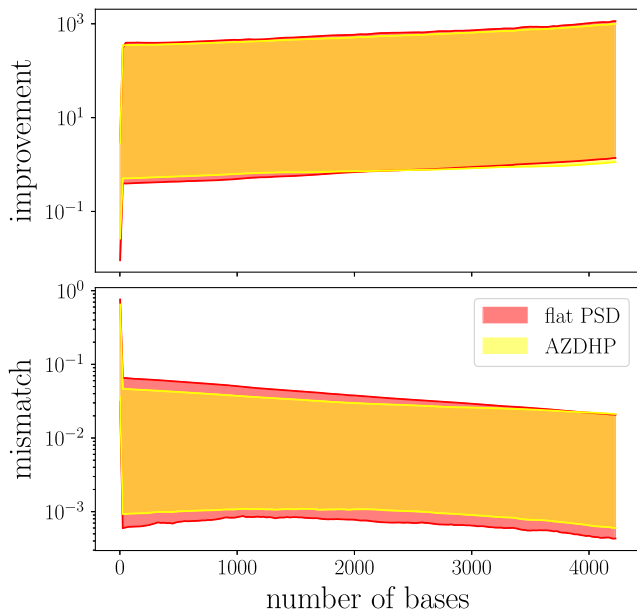
FIG. 7.   The accuracy of a reduced-order model. The top plot is the mismatch between EB and PhenomD waveforms as a function of the number of reduced bases. The bottom plot shows the improvement, cf. (19). The shaded areas are bounded by the minimum and maximum mismatches. The red area is obtained with a flat PSD while the yellow area uses AZDHP. Results with different PSDs overlap well. From this plot, using the minimum of 3375 bases, we can guarantee that all the EB waveforms are more accurate than their approximate waveforms.

space significantly more accurate, then we find that 3375 of 4225 bases are needed to guarantee that the improvement $\mathcal{I} \geq 1$.

We might expect that higher parameter-space dimension ($\mathcal{D} + 1$) require a larger number of bases to obtain at least the same mismatch as lower dimension ($\mathcal{D}$). Naïve intuition would be that the increase of dimensionality in parameter space requires an exponential growth of the basis size. This is called "curse of dimensionality." However, a study by Field *et al.* 2012 [53] shows that one may only need a small number of additional bases in higher dimension to obtain comparable result as in lower dimension. Therefore, the number of reduced bases is not exponentially proportional to the number of dimensions in parameter space. Higher parameter-space dimensions, however, affect computational time as we generate more waveforms covering a greater space.

2. Minimum target waveforms

In the analysis above, we used a uniform grid for target and approximate waveforms with the ratio of PhenomD to PhenomB signals of about $1/4$. In this section, we explore the minimum number of target waveforms needed to obtain a computational efficient EB model that improves the approximate model significantly.
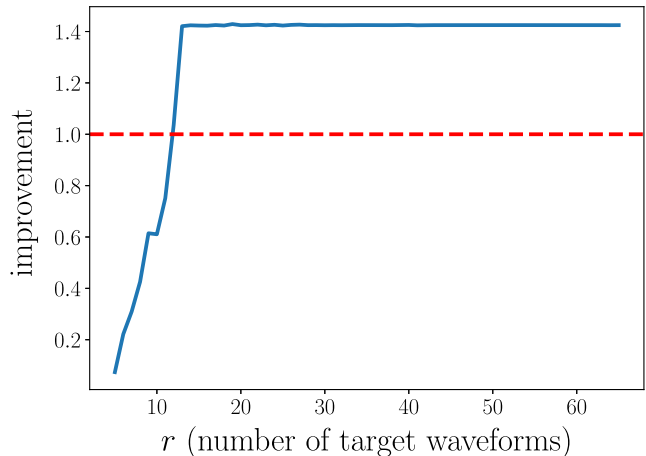
We projected various numbers of target signals given on a sparse, uniform grid with $S = r \times r$ points onto the basis derived from the full N approximate waveforms. We evaluate the improvement $\mathcal{I}$ on the dense grid (after interpolating the projection coefficients from the sparse onto the dense grid) and show the minimum in Fig. 8. We find that $12 \times 12 = 144$ target waveforms guarantee that all EB results are better than PhenomB. This number is almost 30 times smaller than the number of PhenomB signals we use, and more than 95% of the signals generated on the dense grid to compute mismatches are now interpolated and have not been used as target waveforms in the construction of EB.

In fact, we find that the EB model built with $S = 144$ accurate PhenomD signals performs in large parts of the parameter space comparable to the previous case of 33×33 target signals. Only the problematic boundary regions that were visible already in Fig. 5 become more pronounced, both in size and mismatch. Better results, even with this relatively small number of target signals, can be achieved by the iteration procedure we will introduce below.

3. Phase and amplitude contributions

In order to identify the dominant contribution to the inaccuracies that we reported for our EB model, we now evaluate mismatches for individual components. In particular, we can apply the definition of the overlap (17) and mismatch (18) to the amplitude alone, without maximizing over time and phase shifts.

We find that the PhenomB amplitude has relatively high overlap against PhenomD that ranges from 90.78% to 99.98%. As we show in Fig. 9, the EB amplitude also has extremely small mismatches with the target signal PhenomD. Because the strain mismatches, also included in the figure, are orders of magnitude higher, we conclude
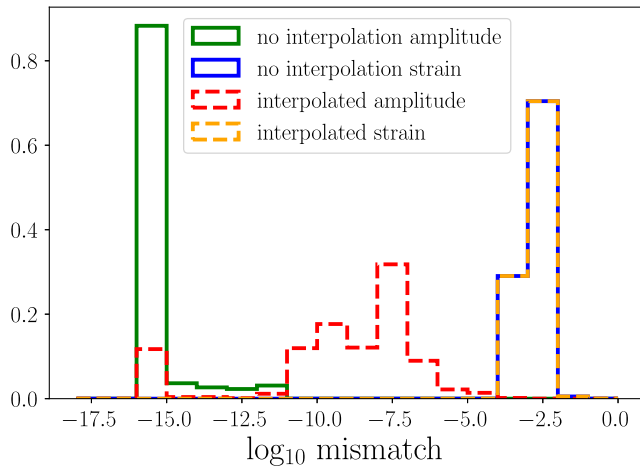
FIG. 9. Normalized histogram of EB amplitude and strain mismatches against PhenomD in flat noise spectrum. The dashed curves are the result from interpolating fewer target signals; the solid line did not employ interpolation (cf. Fig. 6).

that they are dominated by modeling inaccuracies in the phase.

We note that one could in principle calculate mismatches of the pure phase functions as well, but these numbers are less meaningful because they are not invariant under the physical degrees of freedom: phase and time shifts applied to both functions simultaneously. A geometric interpretation relates the overlap to "the angle" between two functions, but because the phases appear in the complex exponential of the strain, the relevant measure is the phase difference instead of their angle.

4. Mass scaling

So far, we have fixed the total mass of the systems in consideration to $M = 50\ M_\odot$. This choice is almost irrelevant for the actual waveform construction as vacuum spacetimes include the system's total mass as a simple scaling factor. As a result, the signal models are actually a function of the dimensionless product Mf. This degeneracy between total mass and frequency is broken when we need to consider physical, full-dimension frequencies that enter the AZDHP noise curve. We also specified our lower cutoff frequency as 30 Hz. Hence, scaling the total mass means appropriately setting $f_{\rm low}$ and $f_{\rm high}$.

Binaries with higher total mass merge at lower frequencies. Therefore, as we have constructed a signal model for $M = 50\ M_\odot$ starting at 30 Hz, we can use the same model also for more massive systems with the same $f_{\rm low}$. The higher mass system then has a shorter frequency range.

Assuming we have carried out the model construction for a total mass M_1, we can scale the frequency of a system with a different total mass M_2, but otherwise the same intrinsic parameters, as follows

$$f_2 = f_1 \left(\frac{M_1}{M_2} \right). \tag{20}$$

As a consequence of the Fourier transform, the strain $\tilde{h}(f)$ also needs to be scaled by the total mass. Putting it all together, the strain for M_2 can be obtained through the following relation,

$$\tilde{h}(f; M_2, \eta, \chi_{\rm eff}) = \left(\frac{M_2}{M_1} \right)^2 \tilde{h}\left(\frac{M_2 f}{M_1}; M_1, \eta, \chi_{\rm eff} \right). \tag{21}$$

Without reconstructing the EB model, we can evaluate the mismatch between EB and the target model PhenomD in frequency range between $f_{\rm low}$ and $f_{\rm high}$ for total masses between 50 to 200 $M_\odot$. We assume the AZDHP PSD. The results are shown in Fig 10. In this plot, we show that the change of mismatches are relatively small for different total masses under the AZDHP noise spectrum. With the same $f_{\rm low}$ (30 Hz) and $f_{\rm high}$ scaled by the total mass as explained above, higher total mass systems produce shorter waveforms. Since the AZDHP noise spectrum is most sensitive in range of early hundred Hz and begin to drop gradually, the agreement between different parts of the waveforms are affected by different sensitivity ranges. Hence the matches are not perfectly uniform for various total masses.

5. SVD iteration

In the previous sections, we found that our method is effective in producing a more accurate waveform model compared to the approximate model we started with,
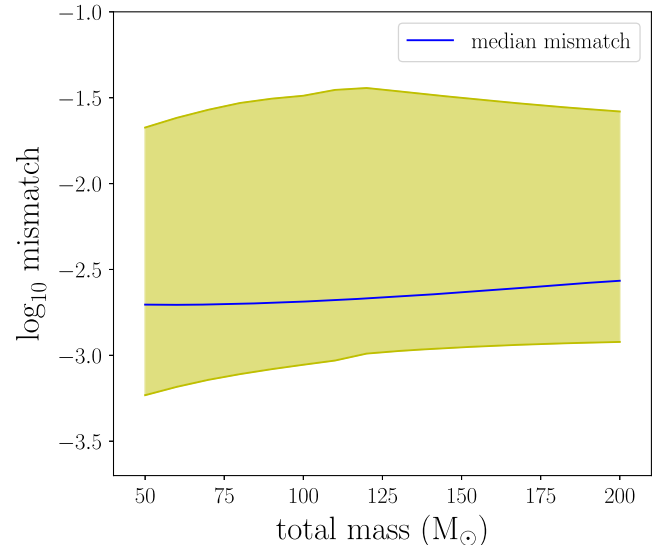

FIG. 10. Mismatches of EB against PhenomD for various total masses as explained in the text. The shaded area is the range of $\log_{10}$ mismatch of the respective total masses, and the blue line is its median. This figure compares mass scaling using AZDHP PSD from 30 Hz to $f_{\rm high}$ of the corresponding total mass.

PhenomB. The mismatches of the resulting EB family are better than PhenomB's mismatches against the target model, PhenomD. This section explores a method to iterate the above steps to produce an even more accurate version of EB, using the same number of approximate and target waveforms.

The basic idea is that we can employ the EB model as the *approximate waveform* of the subsequent iteration and derive a basis from N EB signals interpolated on the dense grid. We then project the same PhenomD signals onto the new basis. We repeat this iterative procedure until the median does not improve significantly.

We first use the minimum number of target waveforms discussed in Sec. III A 2 and later compare the results obtained with more target waveforms. We run the SVD iteration using 12×12 PhenomD signals projected onto 65×65 approximate models without reducing the basis. The reported mismatches employ a flat noise spectrum.

The first EB improves upon PhenomB in mismatch between 1.04 and 860 with median of 23.5. This corresponds to $\log_{10}$ mismatches between -3.36 to -3.57. We then use the EB signals to construct a new SVD basis and run the same process iteratively. After 35 iteration the median $\log_{10}$ mismatch of EB decreases to -4.463 while the median improvement raises to 1254. The mismatch and improvement results are shown in Fig. 11. On a standard laptop, one iteration of this process took about 10 minutes using a single node (no parallelization).

For comparison, we also used PhenomD signals on a 33×33 grid and ran the same iterative process. Using more target waveforms, we achieved a median mismatches below 10^{-6} and an improvement of more than 1750 over PhenomB.
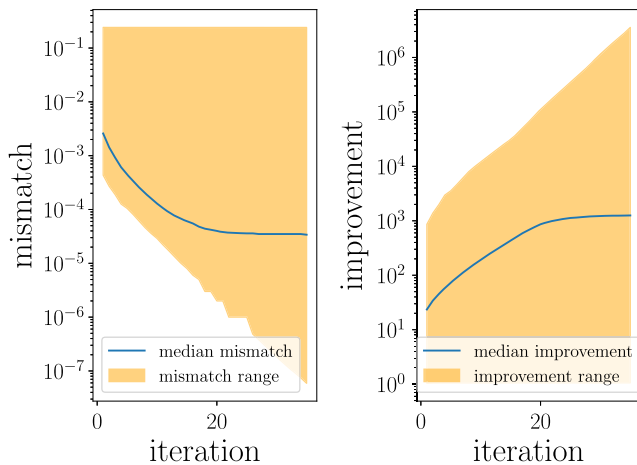

FIG. 11. Mismatches and improvement between EB and PhenomD after iterations. Left: mismatch range over iteration number shown in shaded area with median indicated by the blue curve. Right: improvement range that corresponds to the same iteration is shown in shaded area, where the blue curve shows the median improvement.

In conclusion, we can reduce the mismatch of EB using an iterative process, but of course this will not be as effective as using more target waveforms.

B. Increased dimensionality: Two spins

We have shown that our method can successfully be applied to the aligned equal-spin case, in which both the approximate and target waveforms were varied across an effectively two-dimensional space of intrinsic parameters. Here we expand the dimensionality such that the new EB waveforms are built from a *higher-dimensional* target model projected onto a *lower-dimensional* approximate model. We therefore investigate to what extent the basis can represent a greater parameter space than what it originated from.

Although the case we study here is not yet a practical scenario for actual applications, we argue that in principle one should be able to apply this method for future projections of higher dimensional target models onto lower dimensional basis models.

Specifically, here we consider the case where the target waveforms PhenomD vary in η, χ_{1z} and χ_{2z} individually, so that χ_a [see Eq. (3) for its definition] does not necessarily vanish. We remind the reader that PhenomD is indeed sensitive to these changes, both in the inspiral and in predicting the ringdown signal of the remnant. In contrast, the approximate model PhenomB only depends on $\chi_{\rm eff}$ and not χ_a, hence we keep generating those signals choosing $\chi_{\rm eff} = \chi_{1z} = \chi_{2z}$. Below we discuss results and challenges of this method.

First, we generate the approximate PhenomB waveforms on the same grid of $N = 65 \times 65$ points in the η-$\chi_{\rm eff}$ parameter space that we used before. See Sec. II B for details. Second, we give ourselves $S = 33 \times 33 \times 33 = 35937$ target waveforms on regular grid η, χ_{1z} and χ_{2z}. The parameter ranges are the same as for the approximate signals, except that here $\chi_{1z}, \chi_{2z} \in [-1, 1]$ individually. The procedure we then follow is the same as before. The SVD basis is in fact unchanged compared to what we have used in previous sections, but we now project a much larger number of target signals onto that basis to see if we can accurately represent variations in a parameter that was of no relevance in the approximate model.

Let us emphasize that in this study, we only analyze the errors caused by the projection onto a (lower-dimensional) approximate SVD basis. Therefore, our comparison does not include any interpolation. Instead, we calculate mismatches between PhenomD and either PhenomB or EB on all S points of the parameter space. The results are shown as histograms in Fig. 12. The $\log_{10}$ mismatches of the EB model range from -1.89 to -3.34 (which corresponds to matches between 0.987 to 0.999). Compared to the two-dimensional, equal-spin case of Sec. III A, the matches we find here are slightly lower. This is not surprising, as here we have introduced many more PhenomD waveforms that we know are not accurately captured by PhenomB.
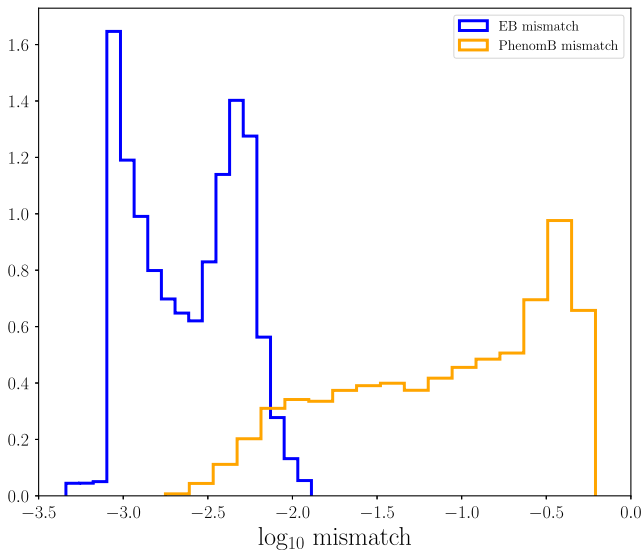
FIG. 12. Normalized histograms of double-spin EB against PhenomD without interpolation, i.e., the EB model was built from $33 \times 33 \times 33$ PhenomD waveforms projected onto an SVD basis of 65×65 PhenomB signals. For comparison, the disagreement between PhenomB and PhenomD is also included.

For comparison, we also show the histogram of mismatches between PhenomB and PhenomD in the same Fig. 12. Evidently, EB achieves a much better accuracy than PhenomB, for which the $\log_{10}$ mismatches range between -2.75 to -0.2 (matches between 0.380 to 0.998). We note that this range is similar to Fig. 3 that is restricted to the equal-spin case.

For completeness, Fig. 13 also illustrates the location of the highest mismatches of between PhenomB and
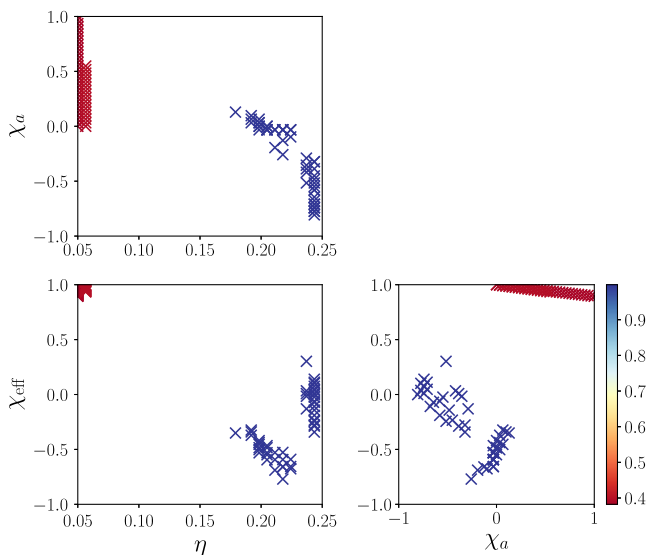

FIG. 13. The location of 50 lowest matches and 50 highest matches of PhenomB against PhenomD. The color bar shows the match. The matches between the worst and the best waves are coloured white.

PhenomD in the parameter space. In this plot, we show the location of the 50 lowest and 50 highest mismatches. The largest disagreement indeed occurs for high mass ratios and asymmetric spins.

From this study, we conclude that one can in principle project a set of higher-dimensional signals onto a basis derived from a lower-dimensional model. However, interpolating across a high-dimensional parameter space becomes much more challenging, especially if a large number of bases has to be included in the EB model. We leave a detailed analysis and discussion of this problem to future work.

IV. CONCLUSION AND FUTURE PERSPECTIVES

The development of accurate GW models is a crucial task to support future detections and the correct interpretation of GWs from merging compact objects. With higher detector sensitivity in the upcoming science runs of LIGO and Virgo, more detections are expected, increasing the chance for an unusually loud, or in other ways special, observation that will require more accurate models than ever before.

Previous work on the development of GW models either targeted a fairly restricted part of the parameter space or required substantial computational as well as human resources. Here we have developed a method to dynamically update an approximate waveform model in a given parameter range. We accomplished this by projecting a set of a more accurate signals onto a larger set of a less accurate waveforms that can be evaluated efficiently and continuously across the parameter space.

We worked in frequency domain and decomposed both waveform models into amplitude and phase that are updated separately. Following earlier studies with a similar goal [36], we employed SVD matrix factorization to split the approximate model's data into two unitary matrices and one diagonal matrix. We used the appropriate unitary matrix as a basis representation of the approximate model, the other two matrices are updated by projecting the accurate model onto that basis. We then interpolated the projection coefficients and combined them with the approximate basis to obtain a new waveform family that we call enriched basis. This model has a higher accuracy than the approximate model and can be evaluated continuously in parameter space.

In this first exploratory study, we restricted ourselves to the nonprecessing parameter space of BBHs. We showed that the EB model is considerably more faithful to its target model (PhenomD) than the approximate model (PhenomB) that we employed. This is true both for flat and AZDHP noise spectra. Let us highlight that especially in regions of the parameter space that were not accurately described by the approximate model because it had not been calibrated there, the improvement of EB can be dramatic. This also holds if an extra physical dependence is

introduced by the target model that was not present in the approximate model.

There are a number of procedural parameters that can be tuned in this approach to achieve optimal results. Among those, we tested the following.

(A) How many basis vectors need to be kept in the EB model? As expected, we found that very few SVD basis vectors are needed to describe the basic parameter dependence of the approximate model. However, as the success of our updating method relies on accurately representing effects beyond what was included in the approximate model, we also found that a wide parameter space such as the one tested here may require a basis of several thousand vectors. We expect this number to sensitively depend on the size of the parameter space and the accuracy of the approximate model.

(B) How many target signals are required?

In the study presented here, we often used a large number of target signal to first test the efficacy of the basic principle. In Sec. III A 2, we reduced that number systematically and analyzed the result. While the edges of the parameter space suffer increasingly from interpolation issues when number of target signals was reduced, we found that even a regular grid of 12×12 target signals showed overall satisfactory improvement. We note that this number is larger than the number of NR waveforms that were used to calibrate PhenomD [32,33] which is not surprising given that more physical insight and intuition went into the original construction while here we test an agnostic, fully automatic approach. We also note that we successfully tested uniform random placement of target waveforms instead of a regular grid, but designing more refined methods of placing target waveforms is an active research topic that can lead to a further reduction of the number of signals required to build an EB model.

(C) Can the process be iterated to achieve better results?

Once a fast and efficient EB model has been built, it can and should be used as an approximate model for the next refinement. While this approach is obvious when more (or different) target waveforms become available, we also showed in Sec. III A 5 that such an iterative procedure can further improve the EB when using the same set of target signals again. This might be counter-intuitive as the same target waveforms seem to be projected onto the same N-dimensional space of amplitude and phase functions in each iterative step. However, it turns out that performing a second SVD on EB data restructures the basis vectors such that the number of irrelevant vectors with vanishingly small σ values increases (i.e., the EB is represented with fewer bases). It is these basis vectors that are not needed to represent the

approximate model, but there are useful in each iterative step to slightly change the vector space toward a more faithful representation of the target. Further studies need to show whether such a procedure also introduces more irregularities and interpolation issues that might counter the gain we report here.

Overall, the results we present here are very promising. One important application that we work toward is actually using the best available analytical models as approximate signals and NR data as the target model. In order for this to be feasible, however, we need to develop additional methods in the immediate future. In particular, the parameter space of most interest include precessing systems, and for those, we eventually need to deal with interpolating over a possibly seven-dimensional parameter space (given by two three-dimensional BH spin vectors and the mass ratio). Interpolating a sparse set of projection coefficients (given by the available NR simulations) may require much more sophisticated interpolation techniques than the ones we have employed here. In fact, we expect interpolation to be the most challenging step in more realistic applications of our procedure.

In addition, a likely scenario where our method could be extremely useful is when a large parameter space needs to be accessible for a signal model to be useful, but targeted NR simulations only cover a reasonable small portion of that space. In that case, our EB model could be updated only where new information is available. This can be achieved by implementing a more flexible interpolation approach that smoothly bridges coefficients based on the approximate model with information from a targeted and localized set of NR data. Such a "hybrid"[3] approach would allow updating established models locally, and it would complement, for instance, parameter estimation methods that take advantage of models that can be generated for arbitrary sets of parameters [54] and alternative methods that use discrete NR data sets [55,56].

We intend to develop solutions for the above-described use cases of EB in the near future. Codes will then be fully integrated in existing analysis suites [46] to guarantee direct impact on the analysis of GW observations. We view this as an important step toward further fostering the integration of numerical and analytical modeling techniques in an era of frequent GW observations.

---

[3]We emphasize that "hybrids" in GW modeling commonly refer to the combination of analytical inspiral signals with numerical merger data. Here, however, we mean it in the sense that parameter space portions where only an approximate model is available could be smoothly combined with new information localized in certain regions of the parameter space.

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

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

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

[4] B. P. Abbott *et al.*, GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2, Phys. Rev. Lett. **118,** 221101 (2017).

[5] B. P. Abbott *et al.*, GW170608: Observation of a 19 solar-mass binary black hole coalescence, Astrophys. J. Lett. **851,** L35 (2017).

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

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

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

[9] B. P. Abbott *et al.*, GW150914: First results from the search for binary black hole coalescence with Advanced LIGO, Phys. Rev. D **93,** 122003 (2016).

[10] L. S. Finn and D. F. Chernoff, Observing binary inspiral in gravitational radiation: One interferometer, Phys. Rev. D **47,** 2198 (1993).

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

[12] L. Blanchet, S. Detweiler, A. L. Tiec, and B. F. Whiting, Post-Newtonian and numerical calculations of the gravitational self-force for circular orbits in the Schwarzschild geometry, Phys. Rev. D **81,** 064004 (2010).

[13] T. Baumgarte and S. Shapiro, *Numerical Relativity: Solving Einstein's Equations on the Computer*, 3rd ed. (Cambridge University Press, Cambridge, England, 2010), vol. 1 of 10.

[14] P. Ajith, L. T. Buchman, T. Chu, C. Reisswig, L. Santamaría, M. A. Scheel, U. Sperhake, B. Szilagyi, and N. W. Taylor, The NINJA-2 catalog of hybrid post-Newtonian/numerical-relativity waveforms for non-precessing black-hole binaries, Classical Quantum Gravity **29,** 124001 (2012); Erratum, Classical Quantum Gravity**30,** 199401(E) (2013).

[15] I. Hinder *et al.*, Error-analysis and comparison to analytical models of numerical waveforms produced by the NRAR Collaboration, Classical Quantum Gravity **31,** 025012 (2014).

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

[17] L. Pekowsky, R. O'Shaughnessy, J. Healy, and D. Shoemaker, Comparing gravitational waves from non-precessing and precessing black hole binaries in the corotating frame, Phys. Rev. D **88,** 024040 (2013).

[18] M. Hannam, S. Husa, B. Brügmann, and A. Gopakumar, Comparison between numerical-relativity and post-Newtonian waveforms from spinning binaries: The orbital hang-up case, Phys. Rev. D **78,** 104007 (2008).

[19] M. Hannam, S. Husa, J. A. González, U. Sperhake, and B. Brügmann, Where post-Newtonian and numerical-relativity waveforms meet, Phys. Rev. D **77,** 044020 (2008).

[20] A. Buonanno and T. Damour, Effective one-body approach to general relativistic two-body dynamics, Phys. Rev. D **59,** 084006 (1999).

[21] A. Buonanno, Y. Pan, H. P. Pfeiffer, M. A. Scheel, L. T. Buchman, and L. E. Kidder, Effective-one-body waveforms calibrated to numerical relativity simulations: Coalescence of nonspinning, equal-mass black holes, Phys. Rev. D **79,** 124028 (2009).

[22] T. Damour and A. Nagar, Comparing effective-one-body gravitational waveforms to accurate numerical data, Phys. Rev. D **77,** 024043 (2008).

[23] T. Damour, A. Nagar, and S. Bernuzzi, Improved effective-one-body description of coalescing nonspinning black-hole binaries and its numerical-relativity completion, Phys. Rev. D **87,** 084035 (2013).

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

[25] Y. Pan, A. Buonanno, L. T. Buchman, T. Chu, L. E. Kidder, H. P. Pfeiffer, and M. A. Scheel, Effective-one-body wave-forms calibrated to numerical relativity simulations: Coalescence of nonprecessing, spinning, equal-mass black holes, Phys. Rev. D **81,** 084041 (2010).

[26] Yi Pan, A. Buonanno, A. Taracchini, M. Boyle, L. E. Kidder, A. H. Mroué, H. P. Pfeiffer, M. A. Scheel, B. Szilágyi, and A. Zenginoglu, Stability of nonspinning effective-one-body model in approximating two-body dynamics and gravitational-wave emission, Phys. Rev. D **89,** 061501 (2014).

[27] A. Taracchini, Y. Pan, A. Buonanno, E. Barausse, M. Boyle, T. Chu, G. Lovelace, H. P. Pfeiffer, and M. A. Scheel, Prototype effective-one-body model for nonprecessing spinning inspiral-merger-ringdown waveforms, Phys. Rev. D **86,** 024011 (2012).

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

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

[30] J. Blackman, S. E. Field, C. R. Galley, B. Szilágyi, M. A. Scheel, M. Tiglio, and D. A. Hemberger, Fast and Accurate Prediction of Numerical Relativity Waveforms from Binary Black Hole Coalescences Using Surrogate Models, Phys. Rev. Lett. **115,** 121102 (2015).

[31] P. Ajith, M. Hannam, S. Husa, Y. Chen, B. Brügmann, N. Dorband, D. Müller, F. Ohme, D. Pollney, C. Reisswig, L. Santamaría, and J. Seiler, Inspiral-Merger-Ringdown Waveforms for Black-Hole Binaries with Nonprecessing Spins, Phys. Rev. Lett. **106,** 241101 (2011).

[32] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. J. Forteza, and A. Bohé, Frequency-domain gravitational waves from nonprecessing black-hole binaries. I. New numerical waveforms and anatomy of the signal, Phys. Rev. D **93,** 044006 (2016).

[33] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era, Phys. Rev. D **93,** 044007 (2016).

[34] K. Cannon, A. Chapman, C. Hanna, D. Keppel, A. C. Searle, and A. J. Weinstein, Singular value decomposition applied to compact binary coalescence gravitational-wave signals, Phys. Rev. D **82,** 044025 (2010).

[35] K. Cannon, C. Hanna, and D. Keppel, Efficiently enclosing the compact binary parameter space by singular-value decomposition, Phys. Rev. D **84,** 084003 (2011).

[36] K. Cannon, J. D. Emberson, C. Hanna, D. Keppel, and H. P. Pfeiffer, Interpolation in waveform space: Enhancing the accuracy of gravitational waveform families using numerical relativity, Phys. Rev. D **87,** 044008 (2013).

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

[38] M. Pürrer, Frequency domain reduced order model of aligned-spin effective-one-body waveforms with generic mass-ratios and spins, Phys. Rev. D **93,** 064041 (2016).

[39] P. Ajith *et al.*, A template bank for gravitational waveforms from coalescing binary black holes. I. Non-spinning binaries, Phys. Rev. D **77,** 104017 (2008); Erratum **79,** 129901 (2009).

[40] P. Ajith *et al.*, A phenomenological template family for black-hole coalescence waveforms, Classical Quantum Gravity **24,** S689 (2007).

[41] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, Simple Model of Complete Precessing Black-Hole-Binary Waveforms, Phys. Rev. Lett. **113,** 151101 (2014).

[42] L. Santamaría, F. Ohme, P. Ajith, B. Brügmann, N. Dorband, M. Hannam, S. Husa, P. Mösta, D. Pollney, C. Reisswig, E. L. Robinson, J. Seiler, and B. Krishnan, Matching post-Newtonian and numerical relativity waveforms: Systematic errors and a new phenomenological model for nonprecessing black hole binaries, Phys. Rev. D **82,** 064016 (2010).

[43] http://www.black-holes.org/waveforms.

[44] B. Brügmann, J. A. González, M. Hannam, S. Husa, U. Sperhake, and W. Tichy, Calibration of moving puncture simulations, Phys. Rev. D **77,** 024027 (2008).

[45] S. Husa, J. A. Gonzalez, M. Hannam, B. Brügmann, and U. Sperhake, Reducing phase error in long numerical binary black hole evolutions with sixth-order finite differencing, Classical Quantum Gravity **25,** 105006 (2008).

[46] https://git.ligo.org/lscsoft/lalsuite.

[47] S. E. Field, C. R. Galley, F. Herrmann, J. S. Hesthaven, E. Ochsner, and M. Tiglio, Reduced Basis Catalogs for Gravitational Wave Templates, Phys. Rev. Lett. **106,** 221102 (2011).

[48] R. Smith, S. E. Field, K. Blackburn, C.-J. Haster, M. Pürrer, V. Raymond, and P. Schmidt, Fast and accurate inference on gravitational waves from precessing compact binaries, Phys. Rev. D **94,** 044031 (2016).

[49] M. Pürrer, Frequency domain reduced order model of aligned-spin effective-one-body waveforms with generic mass ratios and spins, Phys. Rev. D **93,** 064041 (2016).

[50] M. Galassi, J. Davies, J. Theiler, B. Gough, G. Jungman, P. Alken, M. Booth, F. Rossi, and R. Ulerich, *GNU Scientific Library: Reference Manual*, 3rd ed. (Network Theory Limited, United Kingdom, 2009).

[51] Z. Doctor, B. Farr, D. E. Holz, and M. Pürrer, Statistical gravitational waveform models: What to simulate next?, Phys. Rev. D **96,** 123011 (2017).

[52] https://dcc.ligo.org/LIGO-P1200087/public.

[53] S. E. Field, C. R. Galley, and Evan Ochsner, Towards beating the curse of dimensionality for gravitational waves using reduced basis, Phys. Rev. D **86,** 084046 (2012).

[54] J. Veitch *et al.*, Parameter estimation for compact binaries with ground-based gravitational-wave observations using the lalinference software library, Phys. Rev. D **91,** 042003 (2015).

[55] R. O'Shaughnessy, J. Blackman, and Scott E Field, An architecture for efficient gravitational wave parameter estimation with multimodal linear surrogate models, Classical Quantum Gravity **34,** 144002 (2017).

[56] C. Pankow, P. Brady, E. Ochsner, and R. O'Shaughnessy, Novel scheme for rapid parallel parameter estimation of gravitational waves from compact binary coalescences, Phys. Rev. D **92,** 023002 (2015).