Extending black-hole remnant surrogate models to extreme mass ratios

Matteo Boschini[1,2,*] Davide Gerosa[1,3,4] Vijay Varma[2,5] Cristóbal Armaza,[6] Michael Boyle,[6] Marceline S. Bonilla[7]
Andrea Ceja,[7] Yitian Chen[6] Nils Deppe[8,9] Matthew Giesler[6] Lawrence E. Kidder[6] Prayush Kumar,[10]
Guillermo Lara[2] Oliver Long[2] Sizheng Ma,[8] Keefe Mitman[8] Peter James Nee[2] Harald P. Pfeiffer,[2]
Antoni Ramos-Buades,[2] Mark A. Scheel[8] Nils L. Vu[8] and Jooheon Yoo[6]

[1]*Dipartimento di Fisica "G. Occhialini," Universitá degli Studi di Milano-Bicocca,*
Piazza della Scienza 3, 20126 Milano, Italy
[2]*Max Planck Institute for Gravitational Physics (Albert Einstein Institute),*
Am Mühlenberg 1, Potsdam 14476, Germany
[3]*INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy*
[4]*School of Physics and Astronomy & Institute for Gravitational Wave Astronomy,*
University of Birmingham, Birmingham, B15 2TT, United Kingdom
[5]*Department of Mathematics, Center for Scientific Computing and Data Science Research,*
University of Massachusetts, Dartmouth, Massachusetts 02747, USA
[6]*Cornell Center for Astrophysics and Planetary Science, Cornell University, Ithaca, New York 14853, USA*
[7]*Nicholas and Lee Begovich Center for Gravitational-Wave Physics and Astronomy,*
California State University, Fullerton, Fullerton, California 92831, USA
[8]*TAPIR 350-17, California Institute of Technology,*
1200 E California Boulevard, Pasadena, California 91125, USA
[9]*Department of Physics, Cornell University, Ithaca, New York 14853, USA*
[10]*International Centre for Theoretical Sciences, Tata Institute of Fundamental Research,*
Bangalore 560089, India

Numerical-relativity surrogate models for both black-hole merger waveforms and remnants have emerged as important tools in gravitational-wave astronomy. While producing very accurate predictions, their applicability is limited to the region of the parameter space where numerical-relativity simulations are available and computationally feasible. Notably, this excludes extreme mass ratios. We present a machine-learning approach to extend the validity of existing and future numerical-relativity surrogate models toward the test-particle limit, targeting in particular the mass and spin of postmerger black-hole remnants. Our model is trained on both numerical-relativity simulations at comparable masses and analytical predictions at extreme mass ratios. We extend the gaussian-process-regression model NRSur7dq4Remnant, validate its performance via cross validation, and test its accuracy against additional numerical-relativity runs. Our fit, which we dub NRSur7dq4EmriRemnant, reaches an accuracy that is comparable to or higher than that of existing remnant models while providing robust predictions for arbitrary mass ratios.

I. INTRODUCTION

Accurate modeling of merging black-hole (BH) binaries is crucial to both understanding the two-body problem in General Relativity (GR) and characterizing gravitational-wave (GW) observations. Numerical relativity (NR)

*m.boschini1@campus.unimib.it

simulations [1] are, at present, the most accurate approach to capture the dynamics of merging BHs as well as the emitted gravitational signal. Despite the great computational cost of NR, several groups are compiling extensive catalogs of simulations with ever-increasing parameter-space coverage [2–4].

NR simulations are a crucial ingredient to calibrate approximate and fast-to-evaluate emulators for data analysis and astrophysical modeling (for reviews see Refs. [5–9]). Among these, surrogate models are assuming a prominent role. Compared to other approaches, surrogates interpolate the underlying NR simulations in a purely data-driven fashion without ansatzes based on post-Newtonian (PN) and BH-perturbation theory. NR surrogates for both

waveforms [10–14] and postmerger remnants [11,12,15–19] are now widely available and routinely used in GW astronomy.

While generically more accurate than other approximants, the main limitation of surrogate models is their applicability regime, which is largely restricted to the region of the parameter space used for training. Notably, the cost of NR simulations increases for BH binaries with unequal masses, ultimately becoming prohibitive as one approaches the extreme mass-ratio inspiral (EMRI) limit. At present, NR surrogate models for both waveforms and remnants extrapolate rather poorly in this regime.

NRSur7dq4Remnant [11] is the state-of-the-art surrogate model for the mass, spin, and recoil velocity (or kick) of the postmerger remnant. The training set is composed of 1528 NR simulations from the Simulating eXtreme Spacetimes (SXS) catalog [3] which are interpolated with Gaussian Process Regression (GPR) [20]. The model covers the 7-dimensional parameter space spanned by quasicircular binaries: the mass ratio $q = m_1/m_2$ (here defined to be ≥ 1 for consistency with previous work), and the dimensionless spins of the two BHs $\chi_{1,2}$ (where labels 1 and 2 refer to the heavier and lighter BH, respectively), thus including spin precession. The model predicts remnant mass m_f/M (where $M = m_1 + m_2$ is the total mass), spin vector χ_f, and recoil vector v_f. The training dataset is restricted to binaries with $q \leq 4$ and $\chi_{1,2} \leq 0.8$. While numerical extrapolation to the high-spin limit and up to mass ratios $q = 6$ return reasonable results [11], a naive extension of the model toward the EMRI limit fails dramatically. For $q \gg 4$, NRSur7dq4Remnant often returns predictions that are wildly unphysical, such as remnant masses that are greater than the total mass of the binary and/or spins that exceed the Kerr limit.

In this paper, we develop a procedure to consistently extend NR surrogate models for the remnant mass and remnant spin vector to the extreme mass-ratio region. We make use of analytical limits for the energy and angular momentum carried away by GWs to generate suitable training data in addition to the NR simulations used in NRSur7dq4Remnant. We identify the optimal set of tuning parameters using cross-validation and compare our predictions against an additional test set of NR simulations. Our fits, which we dub NRSur7dq4EmriRemnant, are as accurate as the existing ones in the nominal NR training region while extending their validity to extreme mass ratios with similar residuals.

This paper is organized as follows. In Sec. II we write down the EMRI limit for m_f and χ_f. In Sec. III we presents training and validation of our augmented model. In Sec. IV we test our results against additional NR simulations. In Sec. V we investigate the computational costs of our extended models. In Sec. VI draw our conclusions. We use geometric unit where $c = G = 1$.

II. THE EXTREME-MASS-RATIO LIMIT

Our procedure relies on analytical expressions for the mass and spin of the postmerger BH as a function of the binary parameters that are valid in the EMRI limit.

A. Remnant mass

The remnant mass is equal to the difference between the binary total mass and the energy radiated in GWs. We assume that the secondary BH inspirals adiabatically till the innermost stable circular orbit (ISCO) and neglect energy dissipation from the subsequent plunge onward. This is appropriate because, even though the energy loss rate might be large, the transition from ISCO to merger happens in a short time [21].

In this approximation, the final mass is given by (e.g., Ref. [22])

$$\frac{m_f}{M} = 1 - \frac{1}{q}[1 - E_{\mathrm{ISCO}}(\chi_1 \cos\theta_1)] + O\left(\frac{1}{q^2}\right), \quad (1)$$

where $\theta_1 = \arccos\hat{\chi}_1 \cdot \hat{\mathbf{L}}_{\mathrm{ISCO}}$ is the angle between the spin of the primary and the orbital angular momentum at the ISCO. The separation and energy at the ISCO in units of m_1 and m_2, respectively, are [23]

$$E_{\mathrm{ISCO}}(\chi) = \sqrt{1 - \frac{2}{3r_{\mathrm{ISCO}}(\chi)}}, \quad (2)$$

$$r_{\mathrm{ISCO}}(\chi) = 3 + Z_2 - \mathrm{sgn}(\chi)\sqrt{(3 - Z_1)(3 + Z_1 + 2Z_2)}, \quad (3)$$

$$Z_1 = 1 + (1 - \chi^2)^{1/3}[(1 - \chi)^{1/3} + (1 + \chi)^{1/3}], \quad (4)$$

$$Z_2 = \sqrt{3\chi^2 + 2Z_1^2}. \quad (5)$$

The additional terms of $\mathcal{O}(1/q^2)$ in Eq. (1) account for the modification of the ISCO due to the nonzero mass of the secondary BH [24,25] as well as deviations from the adiabatic inspiral near plunge [26].

The expression above is equivalent[1] to that of Barausse *et al.* [22] to first order in $1/q$. Those authors then augment the EMRI ansatz with coefficients that are calibrated on the NR simulations available at the time. They also include information on the spin of the secondary by inserting a suitable weighted combinations of the two spins into the expression for the ISCO energy. In the interest of avoiding uncontrolled systematics from a different NR dataset, here we implement the simple expression from Eq. (1).

---

[1]Note that Refs. [22,27] adopt a convention where $q \leq 1$ while here we use $q \geq 1$ for compatibility with Ref. [11].

B. Remnant spin

We model the spin of the remnant as the sum of the two BH spins and orbital angular momentum at the ISCO:

$$m_f^2 \chi_f = m_1^2 \chi_1 + m_2^2 \chi_2 + m_1 m_2 \mathbf{L}_{\mathrm{ISCO}}, \qquad (6)$$

thus neglecting once more the contribution to the GW flux from the plunge onward. Taylor expanding Eq. (6) to first order in $1/q$ yields

$$\chi_f = \chi_1 + \frac{1}{q}[\mathbf{L}_{\mathrm{ISCO}}(\chi_1 \cos\theta_1)$$
$$- 2\chi_1 E_{\mathrm{ISCO}}(\chi_1 \cos\theta_1)] + O\left(\frac{1}{q^2}\right), \qquad (7)$$

where [23]

$$L_{\mathrm{ISCO}}(\chi) = \frac{2}{3\sqrt{3}}\left[1 + 2\sqrt{3 r_{\mathrm{ISCO}}(\chi) - 2}\right]. \qquad (8)$$

is the orbital angular momentum at the ISCO in units of $m_1 m_2$.

This expression agrees with that presented by Hofmann *et al.* [27] (see also Ref. [28]) as long as terms of $\mathcal{O}(1/q^2)$ are neglected. Much like the mass case described above, we opt for a cleaner ansatz that only contains information on the test-particle limit without further assumption on the role of the secondary spin and/or calibrated coefficients.

III. MODEL TRAINING

Our machine-learning model is trained on a joint dataset made of NR simulations covering the comparable-mass regime and analytical predictions in the EMRI limit. We wish to provide a representation of the mapping between the 7-dimensional space (q, χ_1, χ_2) and the 4-dimensional space (m_f, χ_f).

A. Constructing the training dataset

The NR training set is made of the same 1528 SXS simulations used in Ref. [11]; see therein for details and further references.

For the EMRI datapoints, we consider n_{EMRI} binaries with mass ratios distributed between $q_{\min}$ and $q_{\max}$. We distribute spin magnitudes $\chi_{1,2}$ uniformly in [0, 1], polar angles (measured from the orbital angular momentum) $\theta_{1,2}$ uniformly in $[0, \pi]$, and azimuthal angles (measured in the orbital plane) $\phi_{1,2}$ uniformly in $[0, 2\pi]$. We sample $\theta_{1,2}$ uniformly (and not uniformly in cosine) in order to increase accuracy for aligned spins; this is a configuration that is astrophysically relevant and is also overrepresented in the test set, cf. Sec IV. We sample the mass ratio from a log-uniform distribution, which is the same input passed to the GPR algorithm [11]. Note that the EMRI training set

includes values of the spin of the secondary, even though this is ignored at $\mathcal{O}(1/q)$, cf. Sec. II.

All input quantities in NRSur7dq4Remnant are specified at a time $t = -100M$ before merger [11] in a frame where the z-axis is along the orbital angular momentum, the x-axis is along the line of separation from the less to the more massive BH, and the y-axis completes the right-handed triad. We refer to this frame as the "wave frame" at $t = -100M$ and note it is defined using the waveform at future null infinity [11]. The choice of $t = -100M$ is well suited for modeling the merger remnant for comparable-mass binaries, as this time falls within 2–4 GW cycles before the peak GW amplitude for comparable-mass binaries [29]. We use this same reference frame for the NR binaries of the training set for the new model NRSur7dq4EmriRemnant.

For the additional EMRI inputs, we instead use the wave frame at the ISCO. Defining $t = -100M$ before peak-strain amplitude is challenging for EMRIs as it would require a detailed spin-evolution procedure which is valid in this regime. On the other hand, the ISCO naturally indicates a point near merger where conservation laws can be applied (Sec. II), which implies the resulting datapoints can be (at least approximately) put together with the NR dataset at $t = -100M$. In practice this means that when evaluating the EMRI limits of m_f and χ_f we are neglecting the evolution of the angular momenta between $t = -100M$ and the ISCO for our EMRI data, which is a reasonable assumption [30].

The quantities returned by our model would then also be defined in the wave frame at $t = -100M$ in the NR regime and in the wave frame at ISCO in the EMRI regime, with a transition between the two frames in the intermediate regimes. This is reasonable for this work as the remnant mass and spin are relatively insensitive to small changes in the spins near merger. A more careful approach may be necessary to model the remnant kick, which instead highly depends on the spin directions before merger [15].

B. Fit optimization

Training is performed using GPR [20]. We use the same kernel choices and parameter settings of Refs. [11,16], which appear to work well also in this case.

When estimating the final mass, we enforce the physical constraint $m_f/M \le 1$ by appropriately transforming the input data. In particular, we fit for $\mathrm{arctanh}(m_f/M)$, thus mapping the interval [0, 1) to $[0, \infty)$. When evaluating the trained model, we then reverse the operation [$\tanh: [0, \infty) \to [0, 1)$], which ensures that the remnant mass is properly limited. Furthermore, we empirically find that this choice enhances the overall performance of the model because it simplifies the fitting process: the function tanh resembles the expected behavior of m_f/M as

a function of q, so the training data become closer to linear in that dimension.

As for the remnant spin, we use the GPR strategy of Ref. [11] which requires the cartesian components of χ_f and fits them separately. In the following, we report results in terms of magnitude χ_f, polar angle θ_f measured from the pre-merger binary orbital angular momentum and azimuthal angle ϕ_f measured in the premerger orbital plane. For simplicity, we focus on the first two of these quantities because they are more relevant for astrophysical applications.

For extremely large mass ratios $q \gg q_{max}$, the GPR error on χ_f becomes comparable to the residual, which implies we are better off assuming the EMRI limit at the evaluation stage instead of asking GPR to learn it from the training simulations. We use a squared sinusoidal transition function to smoothly connect fit outcomes and the exact EMRI limit. For mass ratios $q_{max} < q < 2q_{max}$ we correct each cartesian component $\chi_{f,i}$ for $i = \{x, y, z\}$ using the affine transformation

$$\frac{\chi_{f,i}(q) - \chi_{f,i}^{fit}(q_{max})}{\chi_{f,i}^{EMRI}(2q_{max}) - \chi_{f,i}^{fit}(q_{max})} = \sin^2\left(\pi \frac{q - q_{max}}{2q_{max}}\right), \quad (9)$$

where superscripts "fit" and "EMRI" indicate GPR and analytical predictions, respectively. Note we only correct along the q direction of the fit. This output corrections is not necessary for the remnant mass because the zeroth order EMRI limit $m_f/M \to 1$ is already imposed by the tanh transformation described above.

C. Cross validation and parameter tuning

When adding EMRI training points, we set the lower end of the mass-ratio interval to $q_{min} = 100$. We verified numerically that at this value of the mass ratio the discrepancy between our first-order relations from Sec. II and the analytic formulas from Refs. [22,27] is of the same order of the 95th percentile residuals in the NRSur7dq4Remnant model [11], see their Fig. 7. We optimize the number of EMRI binaries n_{EMRI} and the upper edge of the mass ratio regime q_{max} with a grid search. For each of the 13 parameter-space locations shown in the left panels of Fig. 1, we perform a 20-fold cross validation [31] and record the 90th percentile of the absolute differences ΔP_{90} between the true and fitted values of m_f/M, χ_f, and θ_f. For this hyperparameters search, the variations on the final mass are of $\mathcal{O}(10^{-5})$ while those on the final spin are of $\mathcal{O}(10^{-4})$.

An optimal choice of the hyperparameters requires a balance between accuracy and computational costs. Considering the grid search results, the test performance (Sec. IV), and the computational cost (Sec. V), we set $q_{max} = 10^3$ and $n_{EMRI} = 250, 1250$ for the m_f and χ_f fit, respectively. The corresponding models show high accuracy

in both validation and test with a limited increase of computational evaluation time compared to NRSur7dq4-Remnant. Although some of the alternative fits from Fig. 1 returns marginally better results (see, e.g., the $q_{max} = 10^{2.5}$ case for m_f), we believe setting the same mass-ratio interval for both fits (mass and spin) ensures a smoother procedure in light of future retraining with more NR simulations, when available. It is also worth stressing that the difference between these two models $\Delta P_{90} \sim \mathcal{O}(10^{-5})$ is an order of magnitude smaller than the test errors, see Sec. IV.

The right panel of Fig. 1 shows the distributions of the validation residuals for these optimal fits. Considering the full validation estimates, our extended fits NRSur7dq4Emri-Remnant have performances that are comparable to those of NRSur7dq4Remnant from Ref. [11] (see their Fig. 7) while extending its validity to the extreme-mass ratio regime. This is made evident by separating the two subsets of EMRI and NR training data: the NR data are fitted as accurately as NRSur7dq4Remnant and validation errors for the EMRI datapoints are smaller by several orders of magnitude.

IV. MODEL PERFORMANCE

The cross-validation procedure presented above is an internal consistency check, where the model is tested in the same mass-ratio region used for training. More ambitiously, we also test our models against an independent NR dataset. Our test set is made of 1228 NR simulations from the SXS group. These were recently performed for building a new waveform approximant [32] and will be released publicly together with that model. Our test dataset contains numerous systems with $4 < q < 100$ that allows us to test the behavior of the extended fits in between the NR and EMRI training regions. We compare the performance of our extended model NRSur7dq4EmriRemnant against those from NRSur7dq4Remnant as published in Ref. [11] as well as the analytic formulas from Refs. [22,27] (hereafter "HBMR") with $n_M = 3$ and $n_J = 4$. In the following, we do not show HBMR predictions for θ_f because that quantity is not readily accessible from their fits as a function of q, χ_1, and χ_2. NR surrogates have been extensively tested against other fits used in GW astronomy [33–35] with similar results, cf. Ref. [16].

A. Test example

An example is shown in Fig. 2, where we illustrate remnant predictions for a series of BH binaries with $\chi_1 = [0.50, -0.49, -0.31]$ and $\chi_2 = [-0.37, 0.37, 0.42]$ as a function of the mass ratio. For these spin values, the test set contains a simulation with $q = 7.90$.

In the NR training region where $q \leq 4$, our new model NRSur7dq4EmriRemnant is essentially equivalent to the previous version of NRSur7dq4Remnant. The differences between the two are of are comparable to or smaller than
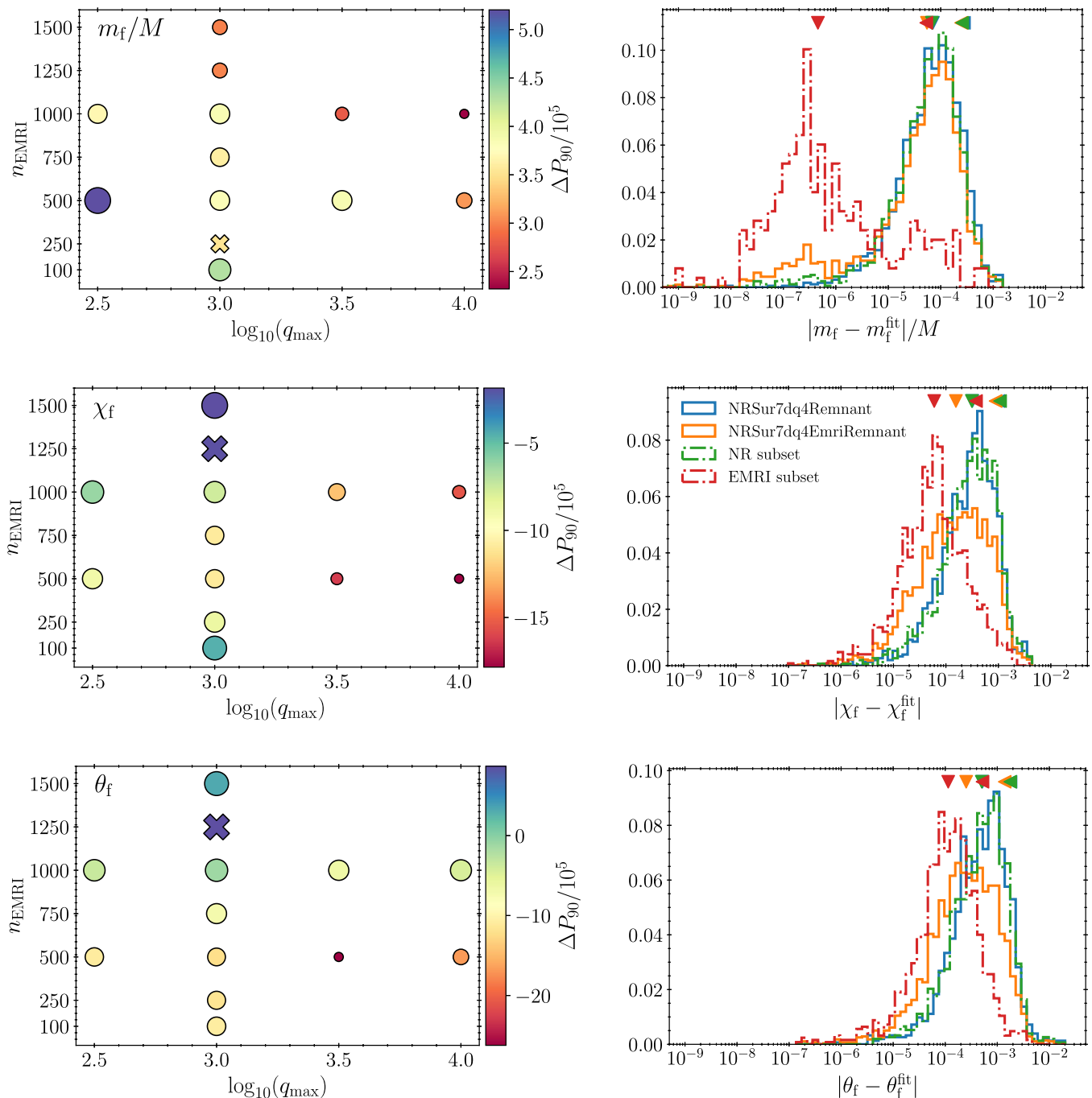
FIG. 1. Outcome of a 20-fold cross validation on the combined NR and EMRI training dataset. Top, middle, and bottom panels show results for the final mass m_f/M, the final spin magnitude χ_f, and the polar angle θ_f. The left panels show the 90th percentile ΔP_{90} of the validation residuals varying over the number of added EMRI systems n_{EMRI} and the upper bound of the EMRI training region q_{max} while keeping the lower bound fixed to $q_{min} = 10^2$. The mass (spin) fit with $n_{EMRI} = 250$ ($n_{EMRI} = 1250$) and $q_{max} = 1000$ are marked with crosses and selected as optimal, see Sec. III. The right panels show the full distribution of the validation residuals for these cases. The validation results computed on the full training dataset are shown in orange; these are then broken down by selecting only the NR (green) and EMRI (red) samples. Residuals from the NRSur7dq4Remnant model as published in Ref. [11] are shown in blue. Left- (down-) pointing triangles indicate the 90th percentile (median) of the corresponding distributions.

the NR residuals (cf. the values reported in Fig. 2 against those in Fig. 7 from Ref. [11]). At higher mass ratios, the new fit converges to the imposed EMRI limit with similar residuals. In contrast, the previous surrogate model departs

significantly from the expected limit, in some cases even returning widely unphysical predictions $m_f > M$. The new fit NRSur7dq4EmriRemnant provides accurate predictions in the intermediate mass ratio region $4 < q < 100$ even if
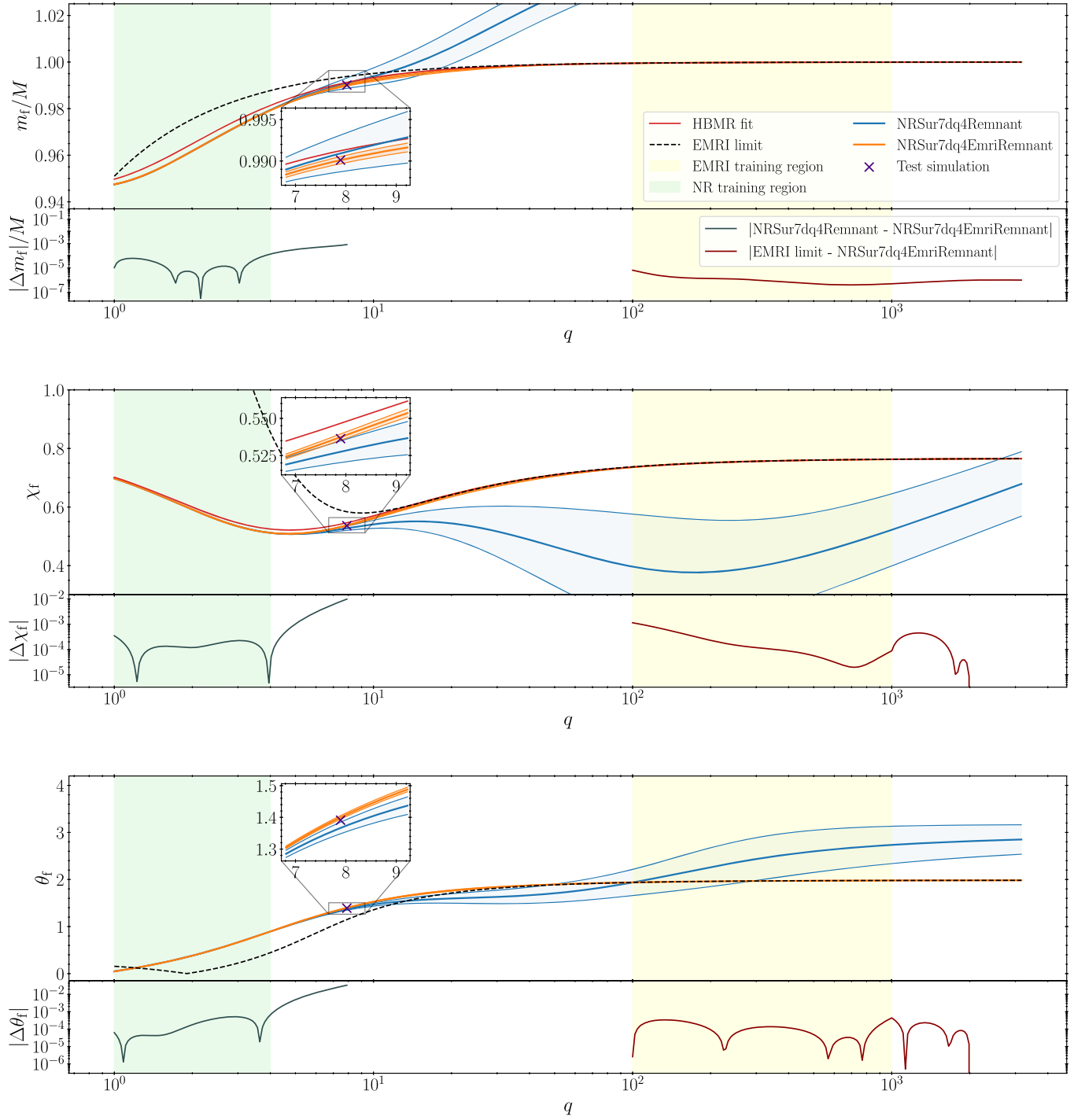
FIG. 2.   Predictions for the remnant mass m_f/M (top panel), spin magnitude χ_f (middle panel) and spin polar angle θ_f (bottom panel). Systems shown in this figure have $\chi_1 = [0.50, -0.49, -0.31]$, $\chi_2 = [-0.37, 0.37, 0.42]$ and different values of q. Orange, blue, and red curves refer to our extended fits NRSur7dq4EmriRemnant, the previously published NRSur7dq4Remnant model [11], and the HBMR analytic expressions [22,27], respectively. For the GPR models, solid curves indicate the returned means and shaded areas indicate 1-σ intervals. Black dotted curves show the EMRI limit from Sec. II. Our model is trained on binaries with mass ratios in the green (NR) and yellow (EMRI) shaded regions. Purple crosses mark the remnant properties extracted from a external test simulation that was not used for either training or internal validation. The bottom subpanels of each panel show residual between our model and NRSur7dq4Remnant at low mass ratios and between our model and the EMRI limit at large mass ratios.
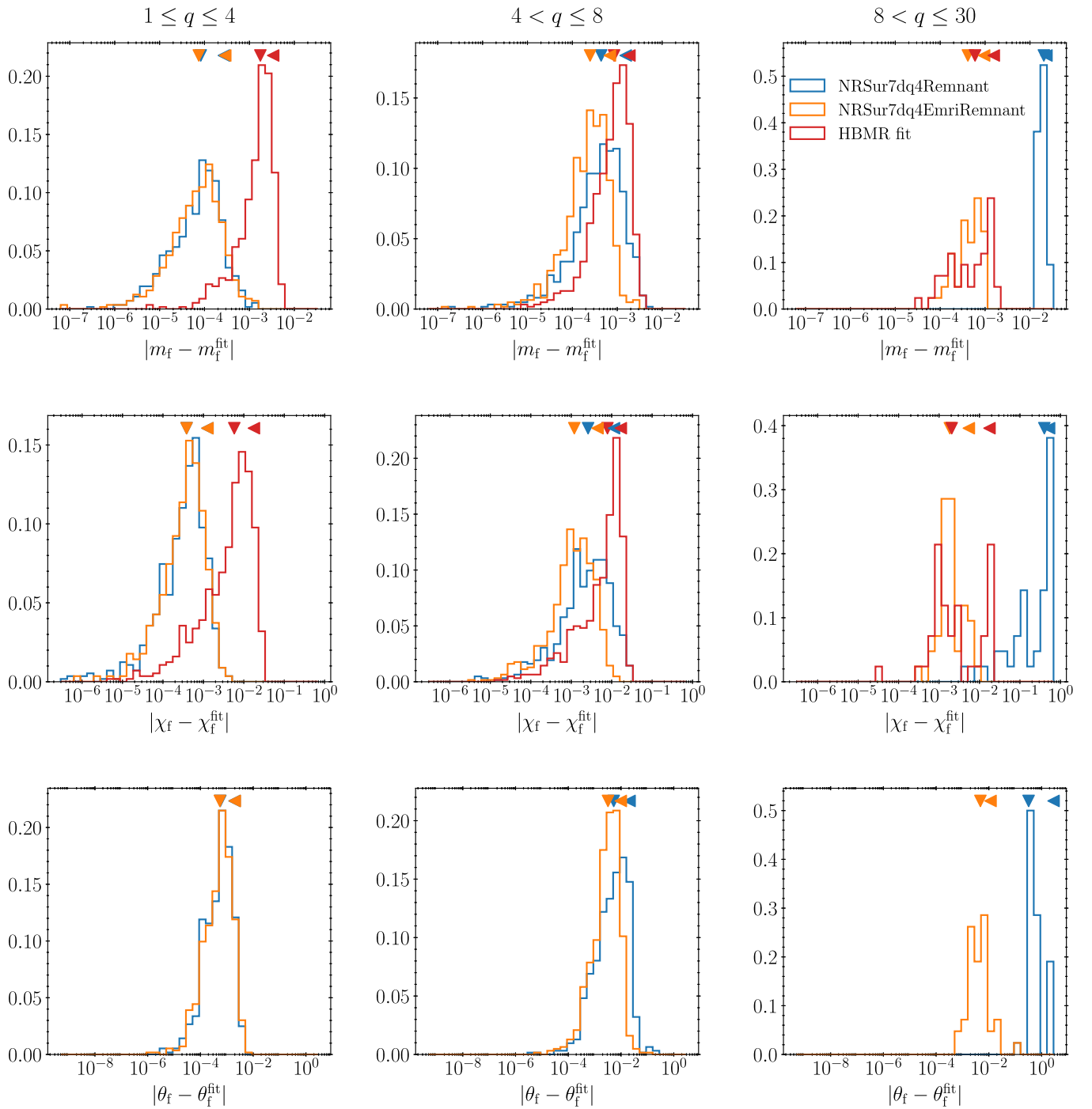
FIG. 3. Test residuals for the extended surrogate (orange) against the previous version of NRSur7dq4Remnant [11] (blue), and the HBMR analytic fits [22,27]. Top, middle, and bottom panels show results for remnant mass m_f/M, spin magnitude χ_f, and polar angle θ_f, respectively. Test simulations are divided into three mass ratio bins: NR ($q < 4$, left), near intermediate ($4 \leq q < 8$, middle), and far intermediate ($8 \leq q < 30$, right). The left- and down-pointing triangles on top show the 90th percentiles and the median for each of the histograms; the former are also reported in Table I.

we do not have training datapoints. The remnant properties extracted from the $q = 7.90$ test case are within the uncertainties returned by the GPR algorithm; we report residuals $|\Delta m_f| \simeq 6.9 \times 10^{-6} M$, $|\Delta \chi_f| \simeq 1.1 \times 10^{-3}$, and $|\Delta \theta_f| \simeq 8.9 \times 10^{-3}$. The extrapolation at $q > 1000$ is also well behaved; for the case of the final spin, the truncation of the residual curve at $q > 2000$ is a consequence of the regularization procedure described in Sec. III.

TABLE I. 90th percentiles on the test residuals for remnant mass m_f/M, spin magnitude χ_f, and spin polar angle θ_f computed for the NRSur7dq4EmriRemnant model presented here, the NRSur7dq4Remnant model from Ref. [11], and the HBMR fits [22,27]. The column labeled "All" indicates percentiles computed over the entire test set while for the last three columns we consider subsets of the test simulations in three mass-ratio bins. The second row indicates the fraction of test simulations in each of these bins.

		All	$1 \leq q \leq 4$	$4 < q \leq 8$	$8 < q \leq 30$
		100%	45.8%	50.7%	3.4%
m_f/M	NRSur7dq4EmriRemnant	6.0×10^{-4}	2.9×10^{-4}	6.9×10^{-4}	9.5×10^{-4}
	NRSur7dq4Remnant	1.4×10^{-3}	2.7×10^{-4}	1.5×10^{-3}	2.3×10^{-2}
	HBMR fit	2.8×10^{-3}	3.4×10^{-3}	1.9×10^{-3}	1.6×10^{-3}
χ_f	NRSur7dq4EmriRemnant	3.4×10^{-3}	1.3×10^{-3}	4.5×10^{-3}	5.4×10^{-3}
	NRSur7dq4Remnant	9.3×10^{-3}	1.3×10^{-3}	1.1×10^{-2}	5.3×10^{-1}
	HBMR fit	1.7×10^{-2}	1.8×10^{-2}	1.7×10^{-2}	1.8×10^{-2}
θ_f	NRSur7dq4EmriRemnant	7.3×10^{-3}	2.0×10^{-3}	9.8×10^{-3}	1.2×10^{-2}
	NRSur7dq4Remnant	1.9×10^{-2}	2.0×10^{-3}	2.1×10^{-2}	2.8×10^{0}

B. Test summary

We now present some summary results using all the NR simulations from the test set. For convenience, these are separated into three disjoint sets:

(i) "NR regime" $1 \leq q \leq 4$ (563 simulations). These are simulations covering the same parameter space of NRSur7dq4Remnant which were not used to train either that model or ours.

(ii) "Near intermediate regime" $4 < q \leq 8$ (623 simulations). At these moderate mass ratio, the parameter space is reasonably well covered by simulations in the test set [32].

(iii) "Far intermediate regime" $8 < q \leq 30$ (42 simulations). These NR runs are computationally very expensive, resulting in limited number of simulations and a sparse coverage of the parameter space. In particular, many test simulations in this regime have aligned, nonprecessing spins, preventing us from fully testing the precession sector.

Note that the largest mass ratio in the test set $q = 30$ is still relatively far from the lowest mass ratio $q_{min} = 100$ covered by our EMRI training data.

Figure 3 shows test residuals for m_f/M, χ_f, and θ_f in each of the tree subsets. The test simulations are compared against the new model NRSur7dq4EmriRemnant, the previous NRSur7dq4Remnant model from Ref. [11], and the HBMR analytic fits [22,27]. The qualitative conclusion is that our fit has the best performances in all three mass ratio regimes. In particular, it is comparable to the previous NRSur7dq4Remnant model and superior to HBMR when in the NR regime ($1 \leq q \leq 4$). In the near intermediate regime ($4 < q \leq 8$), our results is mildly more accurate than both previous fits. As expected, the previous NRSur7dq4Remnant model behaves poorly in the far intermediate regime ($8 < q \leq 30$) while including EMRI information as in this paper returns performance that are similar to, if not better than, those of the analytical HBMR expressions. Quantitative results are presented in Table I, where we report the 90th percentile of our test residuals for each of the three mass ratio bins and well considering the entire test set.

V. COMPUTATIONAL COST

The computational cost of fitting a GPR model scales as $\mathcal{O}(n^3)$, where n is the size of the training dataset. On the other hand, evaluating the fit is an $\mathcal{O}(n^2)$ problem [20]. Compared to NRSur7dq4Remnant from Ref. [11], the new model NRSur7dq4EmriRemnant requires n_{EMRI} additional binaries and thus take longer to evaluate. We test the performance of the new model by generating 10^4 binaries from a broad distribution and evaluating the time needed to compute m_f and χ_f. The execution times reported below refer to a single thread on an Intel Xeon Gold 5220R processor.

The previous NRSur7dq4Remnant model requires $\sim 2.5 \times 10^{-3}$ s to evaluate m_f and $\sim 7.4 \times 10^{-3}$ s to evaluate χ_f. The latter is about three times more expensive than the former because the spin is a vector quantity with three cartesian components while the mass is a single scalar. Compared to this baseline, our new model NRSur7dq4Emri-Remnant increases the computational time by ~ 1.2 times for m_f and ~ 2.6 for χ_f. As expected, this corresponds to a complexity that is roughly quadratic in the size of the training set $n = n_{NR} + n_{EMRI}$ when considering that $n_{NR} = 1528$ and $n_{EMRI} = 250$ (1250). The evaluation time is independent of the mass ratio and spins, with the exception of the spin fit at $q > 2q_{max} = 2000$ where we simply return the EMRI analytical expression.

Overall, this additional computational cost is an acceptable price to pay given the extended parameter space covered by the augmented model presented in this paper.

As the number of available NR simulations increases and better surrogate models are built, the $\mathcal{O}(n^2)$ complexity of GPR evaluations will become critical and alternative regression algorithm will need to be explored.

VI. CONCLUSIONS

We have presented a strategy to augment existing and future NR surrogate models for the mass and spin of the postmerger BH remnant, extending their regime of validity to the test-particle limit. Our approach consists of adding training data points for binaries with extreme-mass ratios using analytic expressions valid at $\mathcal{O}(1/q)$.

We applied this procedure to the GPR fit NRSur7dq4-Remnant [11] which models precessing binary BHs. We tested our augmentation both internally via a cross-validation approach (which was also used to select some model parameters) and externally against a new set of NR simulations. We report excellent performances:

   (i) at comparable masses, our new model behaves like the previous NRSur7dq4Remnant from Ref. [11] which, in turn, was shown to be as accurate as the NR simulations used to train it;

  (ii) at extreme mass ratios, our new model reproduces the test-particle analytic limit with similar residuals; and

 (iii) in between the two regimes, our new model returns regular and accurate values when compared against test NR runs, outperforming both NRSur7dq4Remnant and the HBMR fits.

In summary, we are releasing a single data-driven model able to predict postmerger masses and spins across the entire mass-ratio range, from equal mass binaries to EMRIs. The one drawback is a moderately higher computational cost; which we quantified and found to be acceptable given the extended reliability of the model.

Another limitation of our NR + EMRI approach compared to NR-only surrogates is the time dependence of the spin evolution. NRSur7dq4Remnant allows users to specify the time to merger (or orbital frequency) where spins are specified; it then evolves the spins using the surrogate dynamics to $t = -100M$ where the GPR fits can be consistently evaluated [11,16]. In the absence of a spin-evolution interpolant that captures both comparable masses

and extreme mass ratios, the model presented here can only provide remnant predictions given premerger quantities specified at the reference time $t = -100M$ before merger (which we approximate with the ISCO in the EMRI case, cf. Sec. III). Inputting values (from e.g. GW posterior distributions) that respect this assumption is left to the user.

Future work will tackle the modeling of the remnant kick velocity, which was captured in previous NR-only surrogates [11,15,16]. This is a more challenging task because BH kicks depend on the orbital phase at plunge. Building waveform surrogate models spanning the entire mass ratio range is a considerably more complicated problem and current attempts are restricted to nonspinning sources [36,37].

The model presented in this paper, which we dub NRSur7dq4EmriRemnant, is made publicly available through the Python module surfinBH [38]. More broadly, the augmentation procedure is being integrated in the SXS surrogate codebase and we expect it be valuable for building future BH-remnant models [32].

[1] U. Sperhake, Classical Quantum Gravity **32**, 124011 (2015).

[2] K. Jani, J. Healy, J. A. Clark, L. London, P. Laguna, and D. Shoemaker, Classical Quantum Gravity **33**, 204001 (2016).

[3] M. Boyle, D. Hemberger, D. A. B. Iozzo, G. Lovelace, S. Ossokine, H. P. Pfeiffer, M. A. Scheel, L. C. Stein, C. J. Woodford, A. B. Zimmerman *et al.*, Classical Quantum Gravity **36**, 195006 (2019).

[4] J. Healy and C. O. Lousto, Phys. Rev. D **105**, 124010 (2022).

[5] F. Ohme, Classical Quantum Gravity **29**, 124002 (2012).

[6] L. Blanchet, Living Rev. Relativity **17**, 2 (2014).

[7] T. Damour and A. Nagar, in *Lecture Notes in Physics*, edited by F. Haardt, V. Gorini, U. Moschella, A. Treves, and M. Colpi (Springer Verlag, Berlin, 2016), Vol. 905, p. 273, 10.1007/978-3-319-19416-5_7.

[8] A. Pound and B. Wardell, in *Handbook of Gravitational Wave Astronomy* (Springer Singapore, 2022), p. 38.

[9] M. Tiglio and A. Villanueva, Living Rev. Relativity **25**, 2 (2022).

[10] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, C. D. Ott, M. Boyle, L. E. Kidder, H. P. Pfeiffer, and B. Szilágyi, Phys. Rev. D **96**, 024058 (2017).

[11] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, D. Gerosa, L. C. Stein, L. E. Kidder, and H. P. Pfeiffer, Phys. Rev. Res. **1**, 033015 (2019).

[12] T. Islam, V. Varma, J. Lodman, S. E. Field, G. Khanna, M. A. Scheel, H. P. Pfeiffer, D. Gerosa, and L. E. Kidder, Phys. Rev. D **103**, 064022 (2021).

[13] J. Yoo, V. Varma, M. Giesler, M. A. Scheel, C.-J. Haster, H. P. Pfeiffer, L. E. Kidder, and M. Boyle, Phys. Rev. D **106**, 044001 (2022).

[14] J. Yoo, K. Mitman, V. Varma, M. Boyle, S. E. Field, N. Deppe, F. Hébert, L. E. Kidder, J. Moxon, H. P. Pfeiffer, M. A. Scheel, L. C. Stein, S. A. Teukolsky, W. Throwe, and N. L. Vu, arXiv:2306.03148 [Phys. Rev. D (to be published)].

[15] D. Gerosa, F. Hébert, and L. C. Stein, Phys. Rev. D **97**, 104049 (2018).

[16] V. Varma, D. Gerosa, L. C. Stein, F. Hébert, and H. Zhang, Phys. Rev. Lett. **122**, 011101 (2019).

[17] L. Reali, M. Mould, D. Gerosa, and V. Varma, Classical Quantum Gravity **37**, 225005 (2020).

[18] T. Islam, S. E. Field, and G. Khanna, arXiv:2301.07215 [Phys. Rev. D (to be published)].

[19] L. Haegel and S. Husa, Classical Quantum Gravity **37**, 135005 (2020).

[20] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, Adaptive Computation and Machine Learning (MIT Press, Cambridge, MA, 2006).

[21] D. Kennefick and A. Ori, Phys. Rev. D **53**, 4319 (1996).

[22] E. Barausse, V. Morozova, and L. Rezzolla, Astrophys. J. **758**, 63 (2012).

[23] J. M. Bardeen, W. H. Press, and S. A. Teukolsky, Astrophys. J. **178**, 347 (1972).

[24] L. Barack and N. Sago, Phys. Rev. Lett. **102**, 191101 (2009).

[25] A. Le Tiec, E. Barausse, and A. Buonanno, Phys. Rev. Lett. **108**, 131103 (2012).

[26] M. Kesden, Phys. Rev. D **83**, 104011 (2011).

[27] F. Hofmann, E. Barausse, and L. Rezzolla, Astrophys. J. Lett. **825**, L19 (2016).

[28] E. Barausse and L. Rezzolla, Astrophys. J. Lett. **704**, L40 (2009).

[29] V. Varma, M. Isi, S. Biscoveanu, W. M. Farr, and S. Vitale, Phys. Rev. D **105**, 024045 (2022).

[30] T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, Phys. Rev. D **49**, 6274 (1994).

[31] M. Stone, J. R. Stat. Soc. Ser. B Stat. Methodol. **36**, 111 (1974).

[32] V. Varma *et al.* (to be published).

[33] X. Jiménez-Forteza, D. Keitel, S. Husa, M. Hannam, S. Khan, and M. Pürrer, Phys. Rev. D **95**, 064024 (2017).

[34] J. Healy and C. O. Lousto, Phys. Rev. D **95**, 024037 (2017).

[35] N. K. Johnson-McDaniel, A. Gupta, P. Ajith, D. Keitel, O. Birnholtz, F. Ohme, and S. Husa, LIGO Document No. T070247, 2016, https://dcc.ligo.org/LIGO-T1600168/public.

[36] T. Islam, S. E. Field, S. A. Hughes, G. Khanna, V. Varma, M. Giesler, M. A. Scheel, L. E. Kidder, and H. P. Pfeiffer, Phys. Rev. D **106**, 104025 (2022).

[37] N. E. M. Rifat, S. E. Field, G. Khanna, and V. Varma, Phys. Rev. D **101**, 081502 (2020).

[38] V. Varma *et al.*, 10.5281/zenodo.1418522 (2023).