High-Accuracy Mass, Spin, and Recoil Predictions of Generic Black-Hole Merger Remnants

Varma, Vijay; Gerosa, Davide; Stein, Leo C.; Hébert, François; Zhang, Hao

DOI:
10.1103/PhysRevLett.122.011101

High-accuracy mass, spin, and recoil predictions of generic black-hole merger remnants

Vijay Varma,[1,*] Davide Gerosa,[1,†] Leo C. Stein,[1,2,‡] François Hébert,[1,§] and Hao Zhang[1,3,¶]

[1] *TAPIR 350-17, California Institute of Technology, 1200 E California Boulevard, Pasadena, CA 91125, USA*
[2] *Department of Physics and Astronomy, The University of Mississippi, University, MS 38677, USA*
[3] *Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA*

(Dated: January 11, 2019)

We present accurate fits for the remnant properties of generically precessing binary black holes, trained on large banks of numerical-relativity simulations. We use Gaussian process regression to interpolate the remnant mass, spin, and recoil velocity in the 7-dimensional parameter space of precessing black-hole binaries with mass ratios $q \leq 2$, and spin magnitudes $\chi_1, \chi_2 \leq 0.8$. For precessing systems, our errors in estimating the remnant mass, spin magnitude, and kick magnitude are lower than those of existing fitting formulae by at least an order of magnitude (improvement is also reported in the extrapolated region at high mass ratios and spins). In addition, we also model the remnant spin and kick directions. Being trained directly on precessing simulations, our fits are free from ambiguities regarding the initial frequency at which precessing quantities are defined. We also construct a model for remnant properties of aligned-spin systems with mass ratios $q \leq 8$, and spin magnitudes $\chi_1, \chi_2 \leq 0.8$. As a byproduct, we also provide error estimates for all fitted quantities, which can be consistently incorporated into current and future gravitational-wave parameter-estimation analyses. Our model(s) are made publicly available through a fast and easy-to-use Python module called *surfinBH*.

Introduction– As two black holes (BHs) come together and merge, they emit copious gravitational waves (GWs) and leave behind a BH remnant. The strong-field dynamics of this process are analytically intractable and must be simulated using numerical relativity (NR). However, from very far away, the merger can be viewed as a scattering problem, depicted in Fig. 1. The complicated dynamics of the near zone can be overlooked in favor of the gauge-invariant observables of the in- and out-states: the initial BH masses and spins, the outgoing GWs, and the final BH remnant. This final BH is fully characterized by its mass, spin, and recoil velocity; all additional complexities ("hair") of the merging binary are dissipated away in GWs [1–3].

All GW models designed to capture the entire inspiral-merger-ringdown (IMR) signal from BH binary coalescences need to be calibrated to NR simulations (e.g., [4–12]). In particular, the BH ringdown emission is crucially dependent on the properties of the BH remnant — properties obtained from NR simulations. Accurate modeling of the merger remnant is therefore vital for construction of accurate IMR templates.

Besides waveform building, accurate knowledge of the remnant properties is also instrumental to fulfill one of the greatest promises of GW astronomy: testing Einstein's general relativity (GR) in its strong-field, highly dynamical regime. Current approaches to test the Kerr hypothesis attempt to measure the properties of the inspiralling BHs from the low frequency part of the GW signal,

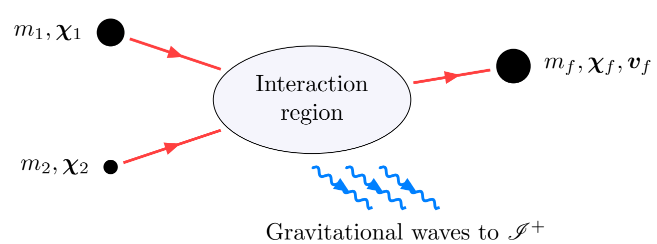

FIG. 1. Quasi-circular binary BH merger problem viewed as a scattering process via a "Feynman" diagram. Time flows to the right. All quantities are well defined in the asymptotically flat region far from the interaction (merger).

then use NR fits to predict the corresponding remnant mass and spin; this final-state prediction is compared to the properties inferred from the high frequency part of the GW signal [13, 14]. Inaccuracies in remnant models therefore directly propagate to the final fundamental-physics test.

The importance of building fits for the remnant properties was realized soon after the NR breakthrough [15–17] and has been periodically revisited by several groups since then [18–39]. There are two important shortcomings in all existing fitting formulae. First, they enforce analytic ansätze (with NR-calibrated coefficients) that are physically motivated, but lack a rigorous mathematical justification. Therefore, current fits can be prone to systematic errors, especially in regions of parameter space where the intuition used to design the formulae become less accurate. Second, current expressions for remnant mass and spins are calibrated on aligned-spin simulations and therefore fail to fully capture the rich physics of precessing systems (but see e.g. [34] where a non-generic subspace of precessing configurations is considered). For example, current LIGO/Virgo parameter-estimation pipelines [40, 41] rely

* vvarma@caltech.edu
† Einstein Fellow; dgerosa@caltech.edu
‡ lcstein@olemiss.edu
§ fhebert@caltech.edu
¶ zhangphy@sas.upenn.edu

on ad-hoc corrections to partially account for precession effects [42]. Aligned fits applied to precessing systems are inevitably ambiguous, as the outcome will depend on *where* (in time, separation, or frequency) the spins are defined and inserted into the fits (e.g., [43]).

In this *Letter* we tackle both these issues for the first time. We construct surrogate models that fit the remnant properties from a large sample of generic, precessing, quasi-circular binary BH simulations performed with the Spectral Einstein Code (SpEC) [44]. Surrogates are trained directly against the NR simulations, using Gaussian process regression (GPR) without any phenomenological ansätz, and achieve accuracies comparable to those of the NR simulations themselves. In their regime of validity, the models presented here are at least an order of magnitude more accurate than previous fits.

In particular, we present two models:

1. *surfinBH7dq2*: a fit trained against precessing systems with mass ratios $q \leq 2$ and dimensionless spin magnitudes $\chi_1, \chi_2 \leq 0.8$.

2. *surfinBH3dq8*: an aligned-spin model trained against systems with mass ratios up to $q \leq 8$ and (anti-)aligned spin magnitudes $\chi_1, \chi_2 \leq 0.8$.

Both these models can be easily accessed using the publicly available Python module *surfinBH* [45], and are ready to be incorporated in both waveform constructions and GW parameter-estimation studies.

Fitting procedure– We construct fits for the BH remnant mass m_f, spin vector $\boldsymbol{\chi}_f$, and recoil kick vector $\boldsymbol{v}_f$ as functions of the binary mass ratio q and spin vectors $\boldsymbol{\chi}_1, \boldsymbol{\chi}_2$. Our fits for *surfinBH7dq2* (*surfinBH3dq8*) map a 7- (3-)dimensional input parameter space to a 7- (4-)dimensional output parameter space. The fits are performed in the coorbital frame at $t=-100M$, with $t=0$ at the peak of the total waveform amplitude (cf. Ref. [12] for details). The coorbital frame is defined such that the z-axis lies along the direction of the orbital angular momentum, the x-axis runs from the smaller BH to the larger BH, and the y-axis completes the triad.

All fits are performed using GPR [46]; details are provided in the supplemental material [47]. Notably, GPR naturally returns estimates of the errors of the fitted quantities across the parameter space.

The values of spins, masses, and kicks used in the training process are extracted directly from the NR simulations. We use the simulations presented in Ref. [12] for *surfinBH7dq2* and those of Ref. [48] for *surfinBH3dq8*. Both spins and masses are evaluated on apparent horizons [49]; the dimensionful spin $\boldsymbol{S}$ solves an eigenvalue problem for an approximate Killing vector, and the mass is determined from the spin and area A following the Christodoulou relation $m^2 = m_{\rm irr}^2 + S^2/(4m_{\rm irr}^2)$, where $m_{\rm irr}^2 = A/16\pi$ is the irreducible mass. The masses $m_{1,2}$ are determined close to the beginning of the simulation at the "relaxation time" [50], whereas the spins $\boldsymbol{\chi}_{1,2} \equiv \boldsymbol{S}_{1,2}/m_{1,2}^2$ are measured at $t=-100M$. The remnant mass m_f and spin $\boldsymbol{\chi}_f$ are determined long after ringdown, as detailed in [50]. All masses are in units of the total mass $M = m_1 + m_2$ at relaxation. The remnant kick velocity is derived from conservation of momentum, $\boldsymbol{v}_f = -\boldsymbol{P}^{\rm rad}/m_f$ [51]. The radiated momentum flux $\boldsymbol{P}^{\rm rad}$ is integrated [52] from the GWs extrapolated to future null infinity [50, 53]. Before constructing the fits, $\boldsymbol{\chi}_f$ and $\boldsymbol{v}_f$ are transformed into the coorbital frame at $t=-100M$.

Besides the GPR error estimate, we further address the accuracy of our procedure using "k-fold" cross validations with $k=20$. First, we randomly divide our training dataset into k mutually exclusive sets. For each set, we construct the fits using the data in the other $k-1$ sets and then test the fits by evaluating them at the data points in the considered set. We thus obtain "out-of-sample" errors which conservatively indicate the (in)accuracies of our fits. We compare these errors against the intrinsic error present in the NR waveforms, estimated by comparing the two highest resolutions available. We also compare the performance of our fits against several existing fitting formulae for remnant mass, spin, and kick which we denote as follows: HBMR ([30, 35] with $n_M=n_J=3$), UIB [37], HL [38], HLZ [33], and CLZM ([21, 22, 27, 31, 32] as summarized in [36]). To partially account for spin precession, fits are corrected as described in Ref. [42] and used in current LIGO/Virgo analyses [40, 41]: spins are evolved from relaxation to the Schwarzschild innermost stable circular orbit, and final UIB and HL spins are post-processed adding the sum of the in-plane spins in quadrature. We note these fitting formulae were calibrated against different sets of simulations. Fitting methods, number of simulations, their quality, and their distribution in parameter space all contribute to the accuracy of the fits.

Aligned-spin model– We first present our fit *surfinBH3dq8*, which is trained against 104 aligned-spin simulations [48] with $q \leq 8$ and $-0.8 \leq \chi_{1z}, \chi_{2z} \leq 0.8$. Symmetry implies that the kick lies in the orbital plane while the final spin is orthogonal to it [54]. We therefore only fit for four quantities: m_f, χ_{fz}, v_{fx}, and v_{fy}.

Figure 2 shows the out-of-sample errors of *surfinBH3dq8*. Our fits are as accurate as the NR simulations used in the training process. 95$^{\rm th}$ percentile errors lie at $\Delta m_f \!\sim\! 4\times10^{-4}M$, $\Delta\chi_f \!\sim\! 10^{-4}$, and $\Delta v_f \!\sim\! 5\times10^{-5}c$. The kick direction is predicted with an accuracy of ~ 0.5 radians, which is the inherent accuracy of the NR simulations. Our errors for the remnant mass and kick magnitude are comparable to the most accurate existing fits. On the other hand, for the final spin, our procedure outperforms all other formulae by at least a factor of 5.

Precessing model– We now present *surfinBH7dq2*, a remnant model trained on 890 simulations [12] of generic, fully precessing BH binaries with mass ratios $q \leq 2$ and spin magnitudes $\chi_1, \chi_2 \leq 0.8$. Out-of-sample errors are shown in Fig. 3. 95$^{\rm th}$ percentiles are $\sim 5\times10^{-4}M$ for mass, $\sim 2\times10^{-3}$ for spin magnitude, $\sim 4\times10^{-3}$ radians for spin direction, $\sim 4\times10^{-4}\,c$ for kick magnitude, and

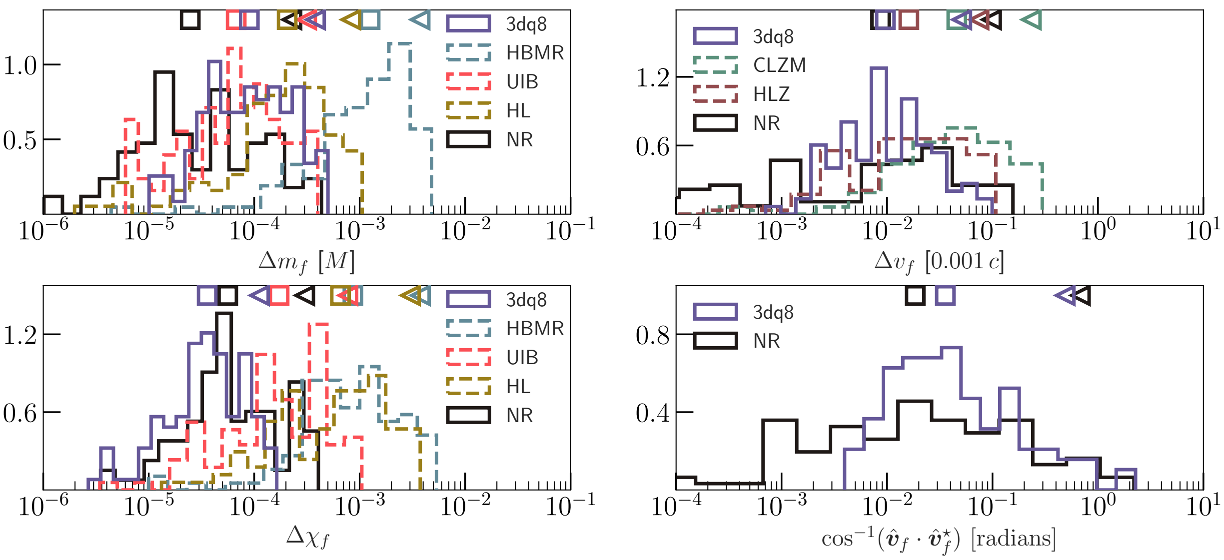

FIG. 2. Errors in predicting remnant mass, spin, kick magnitude, and kick direction for non-precessing binary BHs with mass ratios $q \leq 8$, and spin magnitudes $\chi_1, \chi_2 \leq 0.8$. The direction error is the angle between the predicted vector and a fiducial vector, taken to be the high-resolution NR case and indicated by a *. The square (triangle) markers indicate median (95$^{\text{th}}$ percentile) values. Our model *surfinBH3dq8* is referred to as 3dq8. The black histogram shows the NR resolution error while the dashed histograms show errors for different existing fitting formulae.

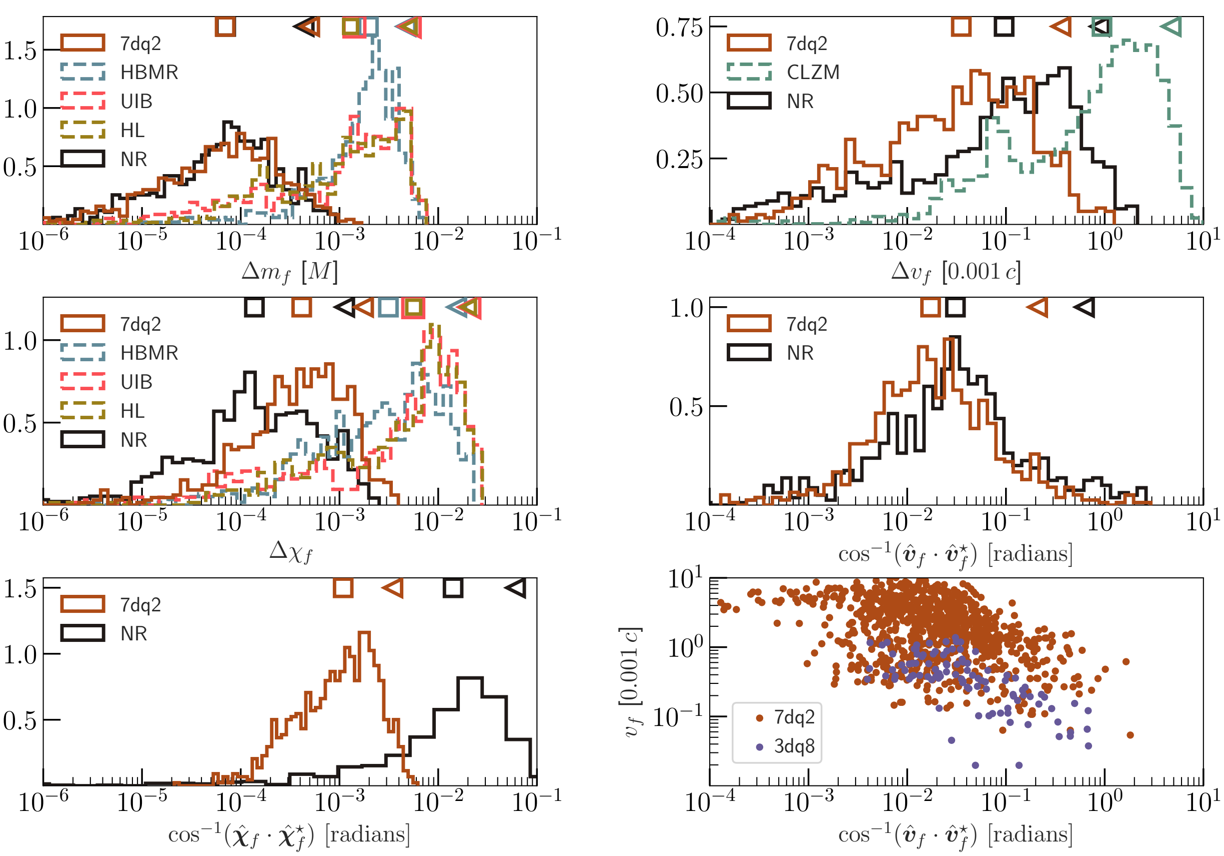

FIG. 3. Errors in predicting the remnant mass, spin magnitude, spin direction, kick magnitude, and kick direction for precessing binary BHs with mass ratios $q \leq 2$, and spin magnitudes $\chi_1, \chi_2 \leq 0.8$. Our model, *surfinBH7dq2* is referred to as 7dq2. The black histogram shows the NR resolution error while the dashed histograms show errors for different existing fitting formulae. In the bottom-right panel we show the distribution of kick magnitude vs. error in kick direction.

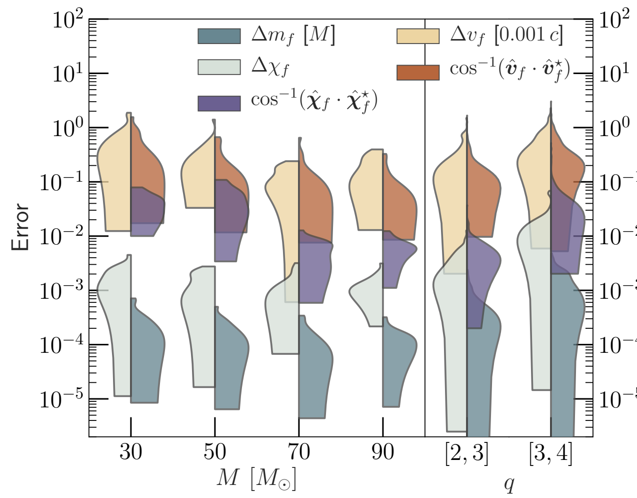

FIG. 4. Left panel: Errors for *surfinBH7dq2* in predicting remnant properties when the spins are specified at an orbital frequency of $f_0 = 10\,\mathrm{Hz}$, for different total masses. Right panel: Errors for *surfinBH7dq2* when extrapolating to higher mass ratios, with the spins specified at $t=-100M$. The labels on the horizontal axis indicate the range of mass ratios being tested. Note that the distributions in these plots are normalized to have a fixed height, not fixed area.

~ 0.2 radians for kick direction. As in the aligned-spin case above, our errors are at the same level as the NR resolution error, thus showing that we are not limited by our fitting procedure but rather by the quality of the training dataset. Our fits appear to outperform the NR simulations when estimating the spin direction, which suggests this quantity has not fully converged in the NR runs, and that the difference between the two highest resolution simulations is an overestimate of the NR error in this quantity.

Figure 3 shows that our procedure to predict remnant mass, spin magnitude, and kick magnitude for precessing systems is more precise than all existing fits by at least an order of magnitude. These existing fits presented significantly lower errors when applied to aligned binaries (cf. Fig. 2), which suggests that they fail to fully capture precession effects despite the augmentation of Ref. [42]. Some impact of precession effects on the final spin and recoil is expected, since both of these quantities have been found to depend strongly on the in-plane orientations of the spins of the merging BHs [43, 51, 55]. More surprisingly, we find that spin precession significantly affects the energy radiated as well, which was expected to depend mostly on the aligned-spin components via the orbital hang-up effect [56–58].

The largest errors in the kick direction can be of order ~ 1 radian. The bottom-right panel of Fig. 3 shows the joint distribution of kick magnitude and kick direction error for both *surfinBH7dq2* and *surfinBH3dq8*, showing that errors are larger at low kick magnitudes. Our error in kick direction is below ~ 0.1 radians whenever $v_f \gtrsim 10^{-3}c$.

Regime of validity– The errors in Fig. 3 are obtained by evaluating fits using input spins specified at $t=-100M$, i.e., where the GPR interpolation is performed. The input spins can also be specified at earlier times; this case is handled by two additional layers of time evolution. Given the spins at an initial orbital frequency f_0, we first evolve the spins using a post-Newtonian (PN) approximant — 3.5PN SpinTaylorT4 [59–61]— until the orbital frequency reaches a value of $0.018\,\mathrm{rad}/M$. At this point, we are in the range of validity of the (more accurate) NRSur7dq2 approximant [12], which we use to evolve the spins until $t=-100M$. Thus, spins can be specified at any given orbital frequency and are evolved consistently before estimating the final BH properties. This is a crucial improvement over previous results, which, being calibrated solely to non-precessing systems, suffer from ambiguities regarding the separation/frequency at which spins are defined.

The left panel of Fig. 4 shows the errors when the spins are specified at an orbital frequency $f_0 = 10\,\mathrm{Hz}$. These errors are computed by comparing against 20 long NR simulations [50] with mass ratios $q \leq 2$ and generically oriented spins with magnitudes $\chi_1, \chi_2 \leq 0.5$. None of these simulations were used to train the fits. Longer PN evolutions are needed at lower total masses, and the errors are therefore larger. These errors will decrease with an improved spin evolution procedure. Note, however, that our predictions are still more accurate (and, crucially, unambiguous) than those of existing fitting formulae (cf. Fig. 3).

Finally, the right panel of Fig. 4 shows the the performance of *surfinBH7dq2* when extrapolating to more extreme mass ratios. We compare against 175 (225) NR simulations [62] with $2 \leq q \leq 3$ ($3 \leq q \leq 4$), and generically oriented spins with magnitudes $\chi_1, \chi_2 \leq 0.8$ specified at $t=-100M$. The error distribution broadens, but our fits still provide a reasonable estimate of the final remnant properties even far out of the training parameter space. Detailed results on extrapolation accuracy are provided in the supplemental materials [47].

Conclusion– We have presented two highly accurate surrogate models for the remnant properties of BH binaries. *surfinBH7dq2* (*surfinBH3dq8*) is trained against 890 (104) NR simulations with mass ratios $q \leq 2$ ($q \leq 8$) and precessing (aligned) spins with magnitude $\chi_1, \chi_2 \leq 0.8$. Both models use GPR to provide fits for the remnant mass, spin, and kick velocity (both magnitudes and directions). Our findings are implemented in a public Python module named *surfinBH* (details are provided in the supplemental materials [47]).

For aligned spins, errors in *surfinBH3dq8* are comparable to existing fitting formulae for the final mass and kick magnitude, while the spin is predicted about 5 times more accurately. For precessing systems, errors in *surfinBH7dq2* for final mass, spin magnitude, and kick magnitude are lower than all existing models by at least an order of magnitude. Crucially, our fits are free from ambiguities regarding the time/frequency at which precessing quantities are specified. This is a point of major improvement over previous models, which all fail to fully capture precession effects.

Is this increased accuracy necessary? For current events like GW150914, the estimated error in the remnant properties are $\Delta m_f \sim 0.1M$ and $\Delta\chi_f \sim 0.1$ [40]. These measurements are currently dominated by statistical errors, as the systematics introduced by existing fits used in the analysis are $\Delta m_f \sim 5\times 10^{-3}M$ and $\Delta\chi_f \sim 2\times 10^{-2}$ (see 95^{th} percentile values in Fig. 3). Because statistical errors scale approximately linearly with the detector sensitivity [63], we estimate that systematic errors in current models for χ_f will start dominating over statistical uncertainties at signal-to-noise ratios which are ~ 5 times larger than that of GW150914. This will happen sooner rather than later, with current interferometers expected to reach their design sensitivity in a few years [64], and future instruments already being scheduled [65] or planned [66, 67]. Our fits, being an order of magnitude more accurate (see Fig. 3), introduce systematic errors which are expected to be relevant only at SNRs ~ 50 times larger than that of GW150914. As shown above, errors are largely dominated by the underlying NR resolution, not by our fitting procedure. The inclusion of self-force evolutions alongside NR in the training dataset might also be exploited to improve extrapolation performance at $q \gg 1$; we leave this to future work.

Moreover, the GPR methods employed here naturally provide error estimates along with the fitted values (some results are provided in the supplemental material [47]). This constitutes a further key application of our results: when performing, e.g., consistency tests of GR [13, 14], systematic uncertainties introduced by remnant fits can be naturally incorporated into the statistical analysis and marginalized over (cf. Ref. [68] for a similar application of GPR and Refs. [69–73] for other applications to GW science).

As GW astrophysics turns into a mature field, increasingly accurate tools such as those presented here will become crucial to uncover more hidden secrets in this new field of science.

Acknowledgments– We thank Jonathan Blackman, Stephen Taylor, David Keitel, Anuradha Gupta, and Serguei Ossokine for useful discussions. We made use of the public LIGO Algorithm Library [74] in the evaluation of existing fitting formulae and to perform PN evolutions. We thank Nathan Johnson-McDaniel for useful discussions, comments on the manuscript, and for sharing his code to evaluate the HLZ kick fits. V.V. and F.H. are supported by the Sherman Fairchild Foundation and NSF grants PHY–1404569, PHY–170212, and PHY–1708213 at Caltech. D.G. is supported by NASA through Einstein Postdoctoral Fellowship Grant No. PF6–170152 awarded by the Chandra X-ray Center, which is operated by the Smithsonian Astrophysical Observatory for NASA under Contract NAS8-03060. L.C.S. acknowledges support from NSF grant PHY–1404569 and the Brinson Foundation. H.Z. acknowledges support from the Caltech SURF Program and NSF Grant No. PHY–1404569. Computations were performed on NSF/NCSA Blue Waters under allocation NSF PRAC–1713694 and on the Wheeler cluster at Caltech, which is supported by the Sherman Fairchild Foundation and by Caltech.

[1] W. Israel, Communications in Mathematical Physics **8**, 245 (1968).
[2] B. Carter, PRL **26**, 331 (1971).
[3] M. Heusler, *Black hole uniqueness theorems, Cambridge University Press, 1996.* (1996).
[4] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, PRL **113**, 151101 (2014), arXiv:1308.3271 [gr-qc].
[5] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, PRD **93**, 044007 (2016), arXiv:1508.07253 [gr-qc].
[6] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. J. Forteza, and A. Bohé, PRD **93**, 044006 (2016), arXiv:1508.07250 [gr-qc].
[7] A. Buonanno, Y. Pan, H. P. Pfeiffer, M. A. Scheel, L. T. Buchman, and L. E. Kidder, PRD **79**, 124028 (2009), arXiv:0902.0790 [gr-qc].
[8] A. Bohé, L. Shao, A. Taracchini, A. Buonanno, S. Babak, I. W. Harry, I. Hinder, S. Ossokine, M. Pürrer, V. Raymond, T. Chu, H. Fong, P. Kumar, H. P. Pfeiffer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, M. A. Scheel, and B. Szilágyi, PRD **95**, 044028 (2017), arXiv:1611.03703 [gr-qc].
[9] S. Babak, A. Taracchini, and A. Buonanno, PRD **95**, 024010 (2017), arXiv:1607.05661 [gr-qc].
[10] S. E. Field, C. R. Galley, J. S. Hesthaven, J. Kaye, and M. Tiglio, PRX **4**, 031006 (2014), arXiv:1308.3565 [gr-qc].
[11] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, D. A. Hemberger, P. Schmidt, and R. Smith, PRD **95**, 104023 (2017), arXiv:1701.00550 [gr-qc].
[12] 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, PRD **96**, 024058 (2017), arXiv:1705.07089 [gr-qc].
[13] A. Ghosh, N. K. Johnson-McDaniel, A. Ghosh, C. Kant Mishra, P. Ajith, W. Del Pozzo, C. P. L. Berry, A. B. Nielsen, and L. London, CQG **35**, 014002 (2018), arXiv:1704.06784 [gr-qc].
[14] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), PRL **116**, 221101 (2016), arXiv:1602.03841 [gr-qc].
[15] F. Pretorius, PRL **95**, 121101 (2005), gr-qc/0507014.
[16] M. Campanelli, C. O. Lousto, P. Marronetti, and Y. Zlochower, PRL **96**, 111101 (2006), gr-qc/0511048.
[17] M. A. Scheel, M. Boyle, T. Chu, L. E. Kidder, K. D. Matthews, and H. P. Pfeiffer, PRD **79**, 024003 (2009), arXiv:0810.1767 [gr-qc].
[18] F. Herrmann, I. Hinder, D. M. Shoemaker, P. Laguna, and R. A. Matzner, PRD **76**, 084032 (2007), arXiv:0706.2541 [gr-qc].

[19] M. Campanelli, C. O. Lousto, Y. Zlochower, and D. Merritt, PRL **98**, 231102 (2007), gr-qc/0702133.
[20] J. A. González, M. Hannam, U. Sperhake, B. Brügmann, and S. Husa, PRL **98**, 231101 (2007), gr-qc/0702052.
[21] J. A. González, U. Sperhake, B. Brügmann, M. Hannam, and S. Husa, PRL **98**, 091101 (2007), gr-qc/0610154.
[22] M. Campanelli, C. Lousto, Y. Zlochower, and D. Merritt, ApJ **659**, L5 (2007), gr-qc/0701164.
[23] L. Rezzolla, E. Barausse, E. N. Dorband, D. Pollney, C. Reisswig, J. Seiler, and S. Husa, PRD **78**, 044002 (2008), arXiv:0712.3541 [gr-qc].
[24] L. Rezzolla, P. Diener, E. N. Dorband, D. Pollney, C. Reisswig, E. Schnetter, and J. Seiler, ApJ **674**, L29 (2008), arXiv:0710.3345 [gr-qc].
[25] M. Kesden, PRD **78**, 084030 (2008), arXiv:0807.3043.
[26] W. Tichy and P. Marronetti, PRD **78**, 081501 (2008), arXiv:0807.2985 [gr-qc].
[27] C. O. Lousto and Y. Zlochower, PRD **77**, 044028 (2008), arXiv:0708.4048 [gr-qc].
[28] E. Barausse and L. Rezzolla, ApJ **704**, L40 (2009), arXiv:0904.2577 [gr-qc].
[29] Y. Pan, A. Buonanno, M. Boyle, L. T. Buchman, L. E. Kidder, H. P. Pfeiffer, and M. A. Scheel, PRD **84**, 124052 (2011), arXiv:1106.1021 [gr-qc].
[30] E. Barausse, V. Morozova, and L. Rezzolla, ApJ **758**, 63 (2012), [Erratum: ApJ, 2014, 786, 76], arXiv:1206.3803 [gr-qc].
[31] C. O. Lousto, Y. Zlochower, M. Dotti, and M. Volonteri, PRD **85**, 084015 (2012), arXiv:1201.1923 [gr-qc].
[32] C. O. Lousto and Y. Zlochower, PRD **87**, 084027 (2013), arXiv:1211.7099 [gr-qc].
[33] J. Healy, C. O. Lousto, and Y. Zlochower, PRD **90**, 104004 (2014), arXiv:1406.7295 [gr-qc].
[34] Y. Zlochower and C. O. Lousto, PRD **92**, 024022 (2015), arXiv:1503.07536 [gr-qc].
[35] F. Hofmann, E. Barausse, and L. Rezzolla, ApJ **825**, L19 (2016), arXiv:1605.01938 [gr-qc].
[36] D. Gerosa and M. Kesden, PRD **93**, 124066 (2016), arXiv:1605.01067 [astro-ph.HE].
[37] X. Jiménez-Forteza, D. Keitel, S. Husa, M. Hannam, S. Khan, and M. Pürrer, PRD **95**, 064024 (2017), arXiv:1611.00332 [gr-qc].
[38] J. Healy and C. O. Lousto, PRD **95**, 024037 (2017), arXiv:1610.09713 [gr-qc].
[39] J. Healy and C. O. Lousto, PRD **97**, 084002 (2018), arXiv:1801.08162 [gr-qc].
[40] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), PRX **6**, 041015 (2016), arXiv:1606.04856 [gr-qc].
[41] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), PRL **118**, 221101 (2017), [Erratum PRL, 2018, 21, 129901], arXiv:1706.01812 [gr-qc].
[42] N. K. Johnson-McDaniel, A. Gupta, P. Ajith, D. Keitel, O. Birnholtz, F. Ohme, and S. Husa, dcc.ligo.org/T1600168/public.
[43] M. Kesden, U. Sperhake, and E. Berti, PRD **81**, 084054 (2010), arXiv:1002.2643 [astro-ph.GA].
[44] L. E. Kidder, M. A. Scheel, S. A. Teukolsky, E. D. Carlson, and G. B. Cook, PRD **62**, 084032 (2000), gr-qc/0005056.
[45] V. Varma *et al.*, pypi.org/project/surfinBH, doi.org/10.5281/zenodo.1418525.
[46] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning, by C.E. Rasmussen and C.K.I. Williams. ISBN-13 978-0-262-18253-9* (2006).
[47] See Supplemental Material here, for details of the GPR fitting method, more detailed exploration of extrapolation errors, tests on the efficacy of GPR's error prediction, and details of the public Python implementation. This further includes Refs. [75–87].
[48] V. Varma, S. Field, M. A. Scheel, J. Blackman, L. E. Kidder, and H. P. Pfeiffer, (2018), arXiv:1812.07865 [gr-qc].
[49] G. Lovelace, R. Owen, H. P. Pfeiffer, and T. Chu, PRD **78**, 084017 (2008), arXiv:0805.4192 [gr-qc].
[50] M. Boyle *et al.*, (2018), in preparation.
[51] D. Gerosa, F. Hébert, and L. C. Stein, PRD **97**, 104049 (2018), arXiv:1802.04276 [gr-qc].
[52] M. Ruiz, M. Alcubierre, D. Núñez, and R. Takahashi, General Relativity and Gravitation **40**, 1705 (2008), arXiv:0707.4654 [gr-qc].
[53] M. Boyle and A. H. Mroué, PRD **80**, 124045 (2009), arXiv:0905.3177 [gr-qc].
[54] L. Boyle, M. Kesden, and S. Nissanke, PRL **100**, 151101 (2008), arXiv:0709.0299 [gr-qc].
[55] E. Berti, M. Kesden, and U. Sperhake, PRD **85**, 124049 (2012), arXiv:1203.2920 [astro-ph.HE].
[56] M. Campanelli, C. O. Lousto, and Y. Zlochower, PRD **74**, 041501 (2006), gr-qc/0604012.
[57] C. O. Lousto and Y. Zlochower, PRD **89**, 021501 (2014), arXiv:1307.6237 [gr-qc].
[58] M. A. Scheel, M. Giesler, D. A. Hemberger, G. Lovelace, K. Kuper, M. Boyle, B. Szilágyi, and L. E. Kidder, CQG **32**, 105009 (2015), arXiv:1412.1803 [gr-qc].
[59] A. Buonanno, Y. Chen, and M. Vallisneri, PRD **67**, 104025 (2003), gr-qc/0211087.
[60] M. Boyle, D. A. Brown, L. E. Kidder, A. H. Mroué, H. P. Pfeiffer, M. A. Scheel, G. B. Cook, and S. A. Teukolsky, PRD **76**, 124038 (2007), arXiv:0710.0158 [gr-qc].
[61] S. Ossokine, M. Boyle, L. E. Kidder, H. P. Pfeiffer, M. A. Scheel, and B. Szilágyi, PRD **92**, 104028 (2015), arXiv:1502.01747 [gr-qc].
[62] V. Varma, S. Field, M. A. Scheel, *et al.*, (2019), in preparation.
[63] M. Vallisneri, PRD **77**, 042001 (2008), gr-qc/0703086.
[64] B. P. Abbott *et al.* (VIRGO, KAGRA, LIGO Scientific), LRR **21**, 3 (2018), arXiv:1304.0670 [gr-qc].
[65] P. Amaro-Seoane *et al.* (LISA Core Team), (2017), arXiv:1702.00786 [astro-ph.IM].
[66] M. Punturo *et al.*, CQG **27**, 194002 (2010).
[67] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), CQG **34**, 044001 (2017), arXiv:1607.08697 [astro-ph.IM].
[68] C. Cahillane, J. Betzwieser, D. A. Brown, E. Goetz, E. D. Hall, K. Izumi, S. Kandhasamy, S. Karki, J. S. Kissel, G. Mendell, R. L. Savage, D. Tuyenbayev, A. Urban, A. Viets, M. Wade, and A. J. Weinstein, PRD **96**, 102001 (2017), arXiv:1708.03023 [astro-ph.IM].
[69] C. J. Moore and J. R. Gair, Physical Review Letters **113**, 251101 (2014), arXiv:1412.3657 [gr-qc].
[70] C. J. Moore, C. P. L. Berry, A. J. K. Chua, and J. R. Gair, PRD **93**, 064001 (2016), arXiv:1509.04066 [gr-qc].
[71] Z. Doctor, B. Farr, D. E. Holz, and M. Pürrer, PRD **96**, 123011 (2017), arXiv:1706.05408 [astro-ph.HE].
[72] E. A. Huerta, C. J. Moore, P. Kumar, D. George, A. J. K. Chua, R. Haas, E. Wessel, D. Johnson, D. Glennon, A. Rebei, A. M. Holgado, J. R. Gair, and H. P. Pfeiffer, PRD **97**, 024031 (2018), arXiv:1711.06276 [gr-qc].
[73] S. R. Taylor and D. Gerosa, PRD **98**, 083017 (2018),

arXiv:1806.08365 [astro-ph.HE].
[74] LIGO Scientific Collaboration and Virgo Collaboration, git.ligo.org/lscsoft/lalsuite.
[75] C. Cutler and É. E. Flanagan, PRD **49**, 2658 (1994), gr-qc/9402014.
[76] E. Poisson and C. M. Will, PRD **52**, 848 (1995), gr-qc/9502040.
[77] D. J. C. Mackay, *Information Theory, Inference and Learning Algorithms, by David J. C. MacKay, pp. 640. ISBN 0521642981. Cambridge, UK: Cambridge University Press, October 2003.* (2003) p. 640.
[78] P. Ajith, PRD **84**, 084037 (2011), arXiv:1107.1267 [gr-qc].
[79] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, A. Müller, J. Nothman, G. Louppe, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, Journal of Machine Learning Research **12**, 2825 (2012), 1201.0490.
[80] A. H. Mroué, M. A. Scheel, B. Szilágyi, H. P. Pfeiffer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, S. Ossokine, N. W. Taylor, A. Zenginoğlu, L. T. Buchman, T. Chu, E. Foley, M. Giesler, R. Owen, and S. A. Teukolsky, PRL **111**, 241104 (2013), arXiv:1304.6077 [gr-qc].
[81] P. Kumar, K. Barkett, S. Bhagwat, N. Afshari, D. A. Brown, G. Lovelace, M. A. Scheel, and B. Szilágyi, PRD **92**, 102001 (2015), arXiv:1507.00103 [gr-qc].
[82] J. Blackman, S. E. Field, C. R. Galley, B. Szilágyi, M. A. Scheel, M. Tiglio, and D. A. Hemberger, PRL **115**, 121102 (2015), arXiv:1502.07758 [gr-qc].
[83] T. Chu, H. Fong, P. Kumar, H. P. Pfeiffer, M. Boyle, D. A. Hemberger, L. E. Kidder, M. A. Scheel, and B. Szilagyi, CQG **33**, 165001 (2016), arXiv:1512.06800 [gr-qc].
[84] E. Jones, T. Oliphant, P. Peterson, *et al.*, "SciPy: Open source scientific tools for Python," `http://www.scipy.org/` (2001–).
[85] S. van der Walt, S. Colbert, and G. Varoquaux, Computing in Science Engineering **13**, 22 (2011).
[86] A. Collette, *Python and HDF5* (O'Reilly, 2013).
[87] Travis Continuous Integration, travis-ci.org.

Supplemental Material

Gaussian process regression– We construct fits in this work using Gaussian process regression (GPR) [S1, S2] as implemented in *scikit-learn* [S3].

The starting point is a training set of n observations, $\mathcal{TS} = \{(x^i, f(x^i))|i = 1, \ldots, n\}$, where x^i denotes an input vector of dimension D and $f(x^i)$ is the corresponding output. In our case, x is mass ratio and spins of the merging binary, and $f(x)$ is the remnant property we are fitting. Our goal is to use $\mathcal{TS}$ to make predictions for the underlying $f(x)$ at any point x_* that is not in $\mathcal{TS}$.

A Gaussian process (GP) can be thought of as a probability distribution of functions. More formally, a GP is a collection of random variables, any finite number of which have a joint Gaussian distribution [S1]. A GP is completely specified by its mean function $m(x)$ and covariance function $k(x, x')$, i.e. $f(x) \sim \mathcal{GP}(m(x), k(x, x'))$. Consider a prediction set of n_* test inputs and their corresponding outputs (which are unknown): $\mathcal{PS} = \{(x^i_*, f(x^i_*))|i = 1, \ldots, n_*\}$. By the definition of a GP, outputs of $\mathcal{TS}$ and $\mathcal{PS}$ (respectively $\mathbf{f} = \{f(x^i)\}$, $\mathbf{f}_* = \{f(x^i_*)\}$) are related by a joint Gaussian distribution

$$\begin{bmatrix} \mathbf{f} \\ \mathbf{f}_* \end{bmatrix} = \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} K_{xx} & K_{xx_*} \\ K_{x_*x} & K_{x_*x_*} \end{bmatrix}\right), \tag{S1}$$

where K_{xx_*} denotes the $n \times n_*$ matrix of the covariance $k(x, x_*)$ evaluated at all pairs of training and prediction points, and similarly for the other K matrices.

Eq. (S1) provides the Bayesian prior distribution for $\mathbf{f}_*$. The posterior distribution is obtained by restricting this joint prior to contain only those functions which agree with the observed data points, i.e. [S1]

$$p(\mathbf{f}_*|\mathcal{TS}) = \mathcal{N}\left(K_{x_*x}K_{xx}^{-1}\mathbf{f},\; K_{x_*x_*} - K_{x_*x}K_{xx}^{-1}K_{xx_*}\right). \tag{S2}$$

The mean of this posterior provides an estimator for $f(x)$ at x_*, while its width is the prediction error.

Finally, one needs to specify the covariance (or kernel) function $k(x, x')$. In this *Letter* we implement the following kernel

$$k(x, x') = \sigma_k^2 \exp\left[-\frac{1}{2}\sum_{j=1}^{D}\left(\frac{x^j - x'^j}{\sigma_j}\right)^2\right] + \sigma_n^2\, \delta_{x,x'}, \tag{S3}$$

where $\delta_{x,x'}$ is the Kronecker delta. In words, we use a product between a squared exponential kernel and a constant kernel, to which we add a white kernel term to account for additional noise in the training data [S1, S3].

GPR fit construction involves determining the $D+2$ hyperparameters (σ_k, σ_n and σ_j) which maximize the marginal likelihood of the training data under the GP prior [S1]. Local maxima are avoided by repeating the optimization with 10 different initial guesses, obtained by sampling uniformly in log in the hyperparameter space described below.

Before constructing the GPR fit, we pre-process the training data as follows. We first subtract a linear fit and the mean of the resulting values. Datapoints are then normalized by dividing by the standard deviation of the resulting values. The inverse of these transformations is applied at the time of the fit evaluation.

For each dimension of x, we define Δx^j to be the range of the values of x^j in $\mathcal{TS}$ and consider $\sigma_j \in [0.01 \times \Delta x^j, 10 \times \Delta x^j]$. Larger length scales are unlikely to be relevant and smaller length scales are unlikely to be resolvable. The remaining hyperparameters are sampled in $\sigma_k^2 \in [10^{-2}, 10^2]$ and $\sigma_n^2 \in [10^{-7}, 10^{-2}]$. These choices are meant to be conservative and are based on prior exploration of the typical magnitude and noise level in our pre-processed training data.

Input parameter space– Fits for *surfinBH3dq8* are parameterized using $x = [\log(q), \hat{\chi}, \chi_a]$, where $\hat{\chi}$ is the spin parameter entering the GW phase at leading order [S4–S7] in the PN expansion,

$$\chi_{\rm eff} = \frac{q\,\chi_{1z} + \chi_{2z}}{1+q}, \qquad \eta = \frac{q}{(1+q)^2}, \tag{S4}$$

$$\hat{\chi} = \frac{\chi_{\rm eff} - 38\eta(\chi_{1z} + \chi_{2z})/113}{1 - 76\eta/113}, \tag{S5}$$

and χ_a is the "anti-symmetric spin",

$$\chi_a = \tfrac{1}{2}(\chi_{1z} - \chi_{2z}). \tag{S6}$$

For *surfinBH7dq2* we use $x = [\log(q), \chi_{1x}, \chi_{1y}, \hat{\chi}, \chi_{2x}, \chi_{2y}, \chi_a]$. Subscripts x, y and z refer to components specified in the coorbital frame at $t = -100M$. We empirically found these parameterizations to perform more accurately than the more intuitive choice $x = [q, \chi_{1x}, \chi_{1y}, \chi_{1z}, \chi_{2x}, \chi_{2y}, \chi_{2z}]$.

In the main text we describe how we evolve spins given at earlier times to $t = -100M$, using PN and NRSur7qd2. Is it worth noting that the NR spins used to train NRSur7qd2 had some additional smoothening filters applied to them (see Eq. 6 in [S8]). This introduces additional systematics when evolving spins from times $t < -100M$. We verified that the resulting errors on our fits are subdominant.

Extrapolation erorrs– The right panel of Fig. 4 shows the errors in remnant quantities when extrapolating *surfinBH7dq2* to mass ratios beyond its training range ($q \leq 2$). These errors are computed using the spins at $t = -100M$. If the spins are given at earlier times, we expect larger extrapolation errors as this also involves extrapolation of the NRSur7dq2 waveform model (which was also trained in the $q \leq 2$ space). Figure S1 shows the extrapolation errors when the spins are specified at at orbital frequency $f_0 = 10\,\text{Hz}$ for a total mass $M = 70M_\odot$, computed by comparing against the same NR simulations as in Fig. 4. Errors are comparable to or lower than those

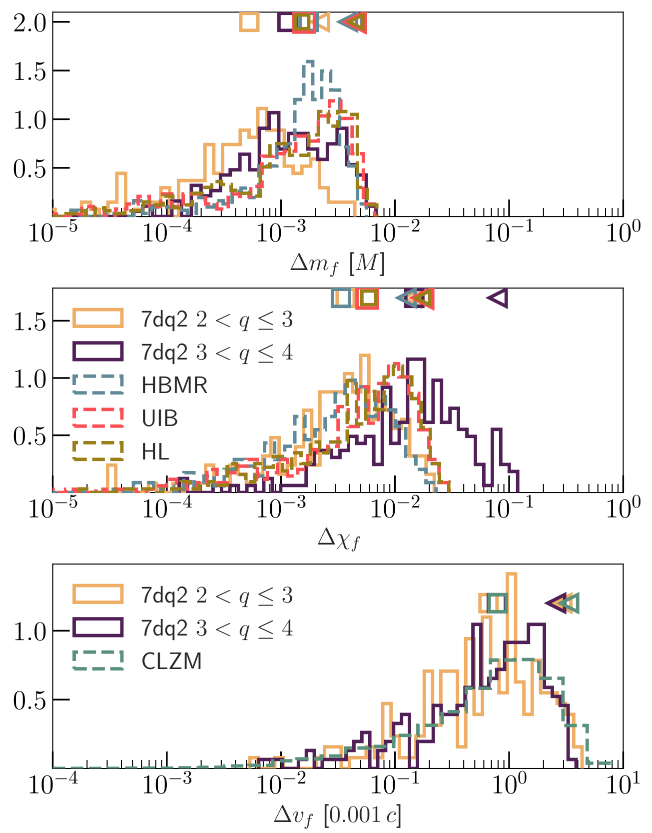

FIG. S1. Errors in *surfinBH7dq2* when extrapolating to higher mass ratios, and the spins are specified at an orbital frequency $f_0 = 10\,$Hz, for a total mass $M = 70 M_\odot$.

of existing fits for $q \leq 3$. For $3 < q \leq 4$, our errors for the remnant spin magnitude can become larger, but the remnant mass and kick magnitude remains as accurate as in other fits.

Figure S2 shows errors in *surfinBH3dq8* when extrapolated beyond its training space to higher mass ratios and/or spin magnitudes (this figure complements the results shown in Fig. 4 of the main text for *surfinBH7dq2*). Here we used some of the simulations of [S9–S13] with $q > 8$ and/or $\chi_1, \chi_2 > 0.8$. Accuracy in the remnant mass degrades noticeably only at high (~ 0.9) co-aligned spins. Errors in final spin become larger at both high spins and extreme mass ratios. For counter-aligned spins, our errors are always comparable to those found within the training region. Errors in kick magnitude and direction appear to be insensitive to extrapolation.

GPR error prediction– As stressed above and in the main body of our *Letter*, GPR naturally associates errors to the estimated quantities. In this Section we test the efficacy of this prediction by comparing the GPR errors against the out-of-sample errors. The GPR errors shown here are evaluated using the same cross-validation data sets used to generate the out-of-sample errors. Therefore, both error estimates are evaluated at points in parameter space where models were not trained.

Error comparisons for *surfinBH3dq8* and *surfinBH7dq2* are reported in Figs. S3 and S4, respectively. While GPR predictions miss some of the features captured by the “k-fold” cross validations, overall it provides faithful estimates of the fit errors.

Public python implementation– Our fits are made publicly available through the easy-to-use Python package, *surfinBH* [S14]. Our code is compatible with both `Python 2` and `Python 3`. The latest release can be installed from the Python Package Index using

```
pip install surfinBH
```

Python packages *numpy* [S15], *scipy* [S16], *h5py* [S17], *scikit-learn* [S3], *lalsuite* [S18], and *NRSur7dq2* [S8] are specified as dependencies and are automatically installed if missing. *surfinBH* is hosted on GitHub at github.com/vijayvarma392/surfinBH, from which development versions can be installed. Continuous integration is provided by *Travis* [S19]

The *surfinBH* module can be imported in Python using

```
import surfinBH
```

Documentation is provided for each submodule of surfinBH and can be accessed via Python's `help()` function. The fit class has to be initialized using, e.g.

```
fit = surfinBH.LoadFits("surfinBH7dq2")
```

Given mass ratio and component spins, the fits and 1σ GPR error estimates of the remnant mass, spin vector and kick vector can be evaluated as follows:

```
q = 1.2
chiA = [0.5, 0.05, 0.3]
chiB = [-0.5, -0.05, 0.1]
mf, mf_err = fit.mf(q, chiA, chiB)
chif, chif_err = fit.chif(q, chiA, chiB)
vf, vf_err = fit.vf(q, chiA, chiB)
```

Both the input spins as well as the remnant spin and kick vectors are assumed to be specified in the coorbital frame at $t = -100M$. Performance of *surfinBH* was tested on a 3.1 GHz Intel Core i5 processor by averaging over 10^3 evaluations at randomly chosen points in parameter space. For *surfinBH7dq2*, evaluation cost of final mass (spin) [kick] is 2.5 ms (7 ms) [7 ms]. For *surfinBH3dq8*, evaluation cost of final mass (spin) [kick] is 0.4 ms (0.4 ms) [0.6 ms].

We also allow specifying an orbital frequency (in units of rad/M), e.g.:

```
omega0 = 5e-3
mf, mf_err = fit.mf(q, chiA, chiB,
                    omega0 = omega0)
chif, chif_err = fit.chif(q, chiA, chiB,
                          omega0 = omega0)
vf, vf_err = fit.vf(q, chiA, chiB,
                    omega0 = omega0)
```

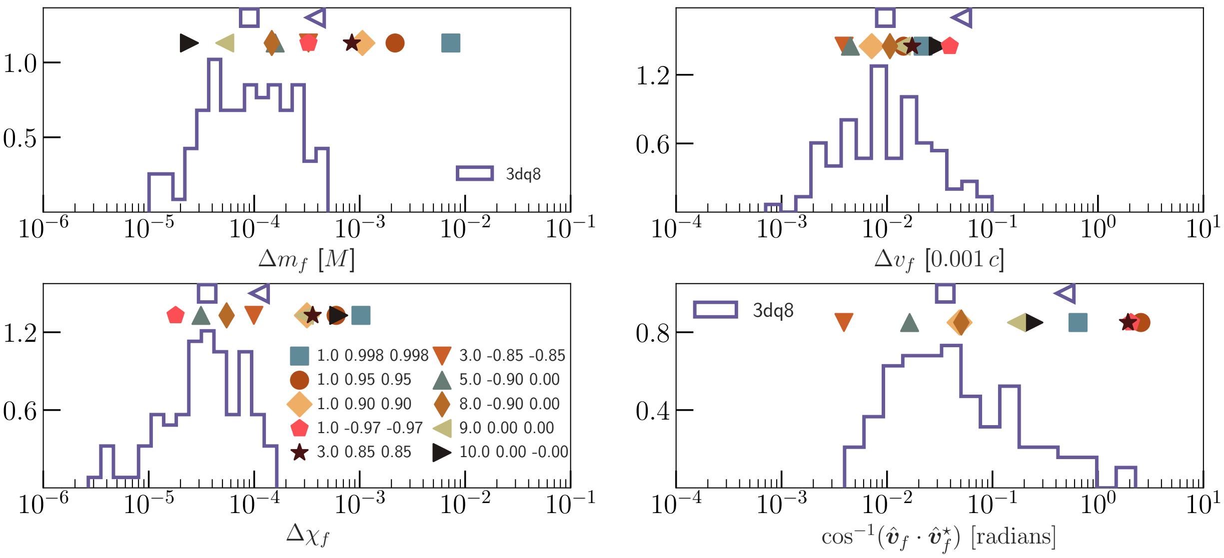

FIG. S2. Errors in predicting the remnant mass, spin, kick magnitude and kick direction for nonprecessing BBH when *surfinBH3dq8* is extrapolated outside of the training region (i.e. $q > 8$ and $\chi_1, \chi_2 > 0.8$). Each solid symbol marks the error of the extrapolated model against a single nonprecessing NR simulation. The legend in the bottom-left panel displays the mass ratio and spin components of the two BHs along the orbital angular momentum direction. Histograms of errors within the training region (from Fig. 2) are reproduced here for comparison. The hollow square (triangle) markers indicate the median (95^{th} percentile) values for those errors.

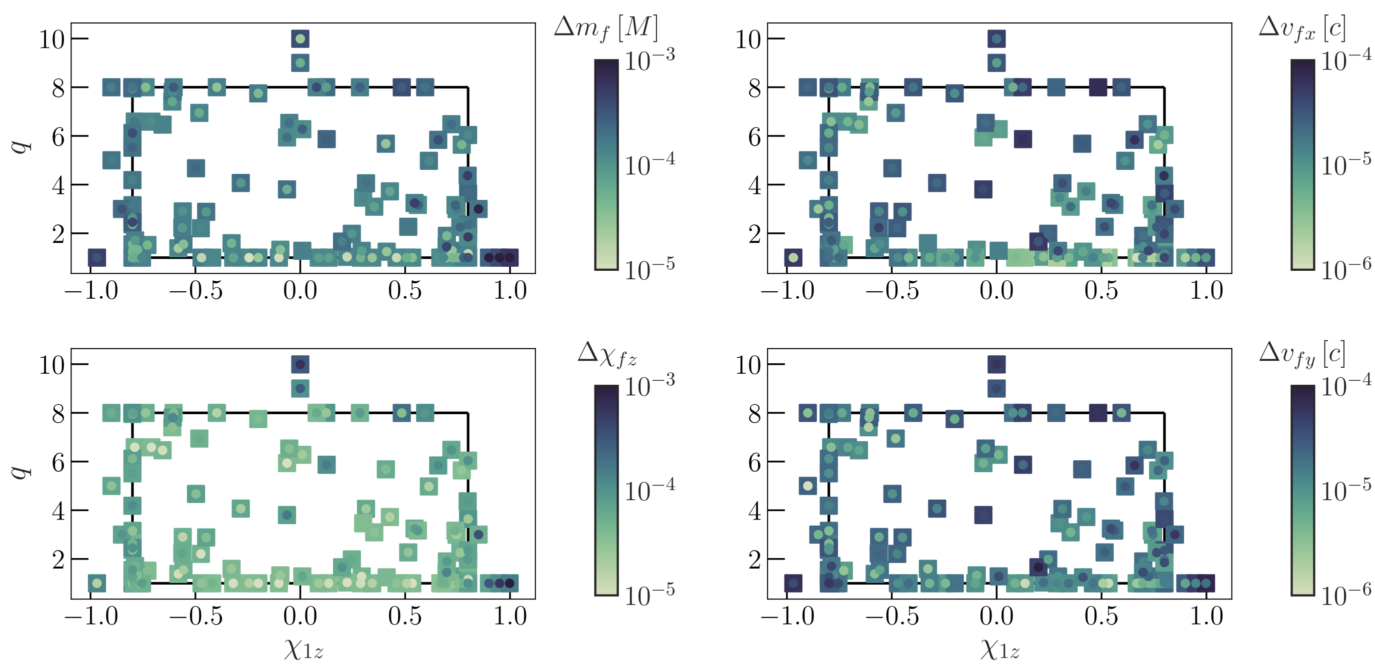

FIG. S3. Prediction errors for remnant mass, spin and kick for the model *surfinBH3dq8* against NR simulations. Two error estimates, as reported on the color scale, are compared: out-of-sample errors marked with circles, and 1σ GPR errors marked with squares. We include cases where *surfinBH3dq8* needs to be extrapolated to higher mass ratios and/or spin magnitudes. The bounds of the training parameter space are indicated as a black rectangle.

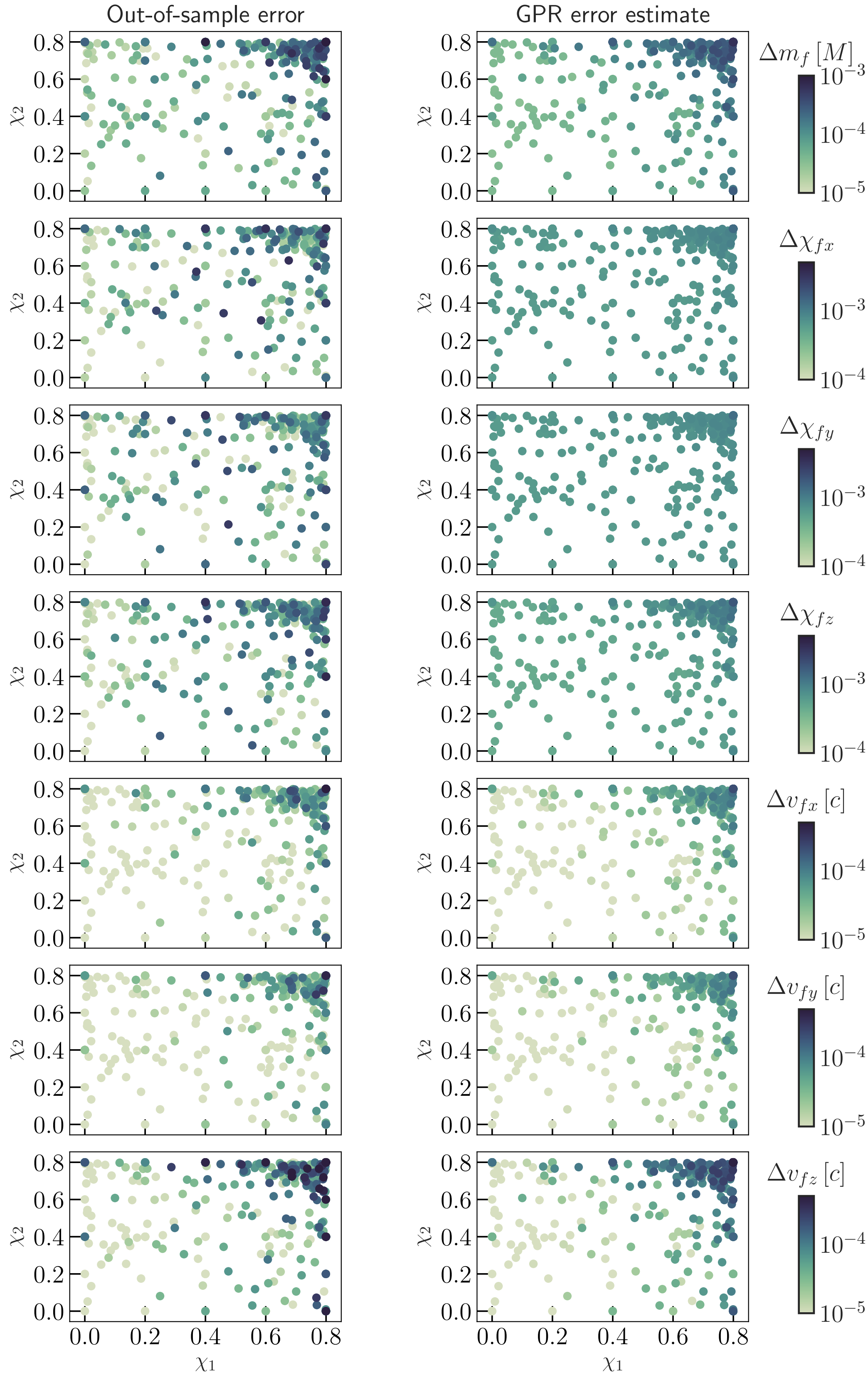

FIG. S4. Comparison between out-of-sample (left) and 1σ GPR (right) errors for *surfinBH7dq2*. The axes show the magnitudes of the component spins, and the color scale indicates the parameter error being plotted.

In this case, the component spins, as well as the final remnant spin/kick vectors are specified in the coorbital frame at this orbital frequency. The evaluation costs are larger when specifying an initial orbital frequency since this involves two additional stages of spin evolution. Execution times depend on the initial frequency, the specific PN approximant used and the time step size in the integration routine. For instance, with `omega0 = 5e-3`, SpinTaylorT4, and a step size of $0.1M$ the evaluation cost is ~ 0.5s for each of the remnant quantities.

Additional resources are provided in the package installation page [S14]. This includes example `jupyter` notebooks for both models presented in this *Letter*.

[S1] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning, by C.E. Rasmussen and C.K.I. Williams. ISBN-13 978-0-262-18253-9* (2006).

[S2] D. J. C. Mackay, *Information Theory, Inference and Learning Algorithms, by David J. C. MacKay, pp. 640. ISBN 0521642981. Cambridge, UK: Cambridge University Press, October 2003.* (2003) p. 640.

[S3] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, A. Müller, J. Nothman, G. Louppe, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, Journal of Machine Learning Research **12**, 2825 (2012), 1201.0490.

[S4] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, PRD **93**, 044007 (2016), arXiv:1508.07253 [gr-qc].

[S5] P. Ajith, PRD **84**, 084037 (2011), arXiv:1107.1267 [gr-qc].

[S6] C. Cutler and É. E. Flanagan, PRD **49**, 2658 (1994), gr-qc/9402014.

[S7] E. Poisson and C. M. Will, PRD **52**, 848 (1995), gr-qc/9502040.

[S8] 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, PRD **96**, 024058 (2017), arXiv:1705.07089 [gr-qc].

[S9] A. H. Mroué, M. A. Scheel, B. Szilágyi, H. P. Pfeiffer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, S. Ossokine, N. W. Taylor, A. Zenginoğlu, L. T. Buchman, T. Chu, E. Foley, M. Giesler, R. Owen, and S. A. Teukolsky, PRL **111**, 241104 (2013), arXiv:1304.6077 [gr-qc].

[S10] P. Kumar, K. Barkett, S. Bhagwat, N. Afshari, D. A. Brown, G. Lovelace, M. A. Scheel, and B. Szilágyi, PRD **92**, 102001 (2015), arXiv:1507.00103 [gr-qc].

[S11] J. Blackman, S. E. Field, C. R. Galley, B. Szilágyi, M. A. Scheel, M. Tiglio, and D. A. Hemberger, PRL **115**, 121102 (2015), arXiv:1502.07758 [gr-qc].

[S12] T. Chu, H. Fong, P. Kumar, H. P. Pfeiffer, M. Boyle, D. A. Hemberger, L. E. Kidder, M. A. Scheel, and B. Szilagyi, CQG **33**, 165001 (2016), arXiv:1512.06800 [gr-qc].

[S13] M. Boyle *et al.*, (2018), in preparation.

[S14] V. Varma *et al.*, pypi.org/project/surfinBH, doi.org/10.5281/zenodo.1418525.

[S15] S. van der Walt, S. Colbert, and G. Varoquaux, Computing in Science Engineering **13**, 22 (2011).

[S16] E. Jones, T. Oliphant, P. Peterson, *et al.*, "SciPy: Open source scientific tools for Python," `http://www.scipy.org/` (2001–).

[S17] A. Collette, *Python and HDF5* (O'Reilly, 2013).

[S18] LIGO Scientific Collaboration and Virgo Collaboration, git.ligo.org/lscsoft/lalsuite.

[S19] Travis Continuous Integration, travis-ci.org.