http://eprints.gla.ac.uk/148145/

Deposited on: 15 September 2017

Hierarchical data-driven approach to fitting numerical relativity data
for nonprecessing binary black holes with an application to final spin and radiated energy

Xisco Jiménez-Forteza,[1, *] David Keitel,[1, 2, †] Sascha Husa,[1, ‡] Mark Hannam,[3] Sebastian Khan,[3] and Michael Pürrer[3, 4]

[1]*Universitat de les Illes Balears, IAC3—IEEC, 07122 Palma de Mallorca, Spain*
[2]*School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, United Kingdom*
[3]*School of Physics and Astronomy, Cardiff University, The Parade, Cardiff, CF24 3AA, United Kingdom*
[4]*Albert-Einstein-Institut, Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany*

Numerical relativity is an essential tool in studying the coalescence of binary black holes (BBHs). It is still computationally prohibitive to cover the BBH parameter space exhaustively, making phenomenological fitting formulas for BBH waveforms and final-state properties important for practical applications. We describe a general hierarchical bottom-up fitting methodology to design and calibrate fits to numerical relativity simulations for the three-dimensional parameter space of quasicircular nonprecessing merging BBHs, spanned by mass ratio and by the individual spin components orthogonal to the orbital plane. Particular attention is paid to incorporating the extreme-mass-ratio limit and to the subdominant unequal-spin effects. As an illustration of the method, we provide two applications, to the final spin and final mass (or equivalently: radiated energy) of the remnant black hole. Fitting to 427 numerical relativity simulations, we obtain results broadly consistent with previously published fits, but improving in overall accuracy and particularly in the approach to extremal limits and for unequal-spin configurations. We also discuss the importance of data quality studies when combining simulations from diverse sources, how detailed error budgets will be necessary for further improvements of these already highly accurate fits, and how this first detailed study of unequal-spin effects helps in choosing the most informative parameters for future numerical relativity runs.

I. INTRODUCTION

According to general relativity, compact binary black holes (BBHs) coalesce through the emission of gravitational waves (GWs), as already observed by LIGO in at least two cases [1–4]. The merger remnant is a single Kerr BH [5] characterized only by its final spin and mass. An essential and robust tool for predicting BBH evolution, since the 2005 breakthrough [6–8], is numerical relativity (NR). Due to the large computational cost of each NR simulation, it is still computationally prohibitive to cover the BBH parameter space exhaustively. Hence it is natural to develop simple, yet accurate model fits to the existing set of NR simulations. One then has to study their quality of interpolation between NR simulations, and of extrapolation to regions of parameter space beyond the calibration range. For example, the phenomenological inspiral-merger-ringdown waveform models of [9–13], though used very successfully as one of two waveform families for LIGO O1 data analysis [1–4], still include only a limited set of physical effects and were calibrated to small sets of NR simulations with relatively simple fitting methods.

In this paper we develop a more general, hierarchical fitting approach for BBH properties and waveforms. The fit ansatz functions are developed through studying hierarchical structures present in the NR data set itself, and their complexity tailored to the actual predictive power of the data set by the use of information criteria.

We first apply this method to the spin and mass of the remnant black hole, which determine the frequencies of the quasinormal-mode ringdown [14–17] to a Kerr black hole. The ringdown is a crucial part in modeling full inspiral-merger-ringdown waveforms [9–13, 18–20]; and from param-

eter estimation with the full waveforms, the final mass and spin can be estimated with accuracy similar to other BBH parameters [2, 4]. Future observations of strong GW signals will allow to test general relativity through consistency tests between inspiral, merger and ringdown [21], significantly improving upon [22].

Apart from GW observations, the final state of a BBH merger is astrophysically interesting in itself, e.g. for the computation of merger trees [23–30]. The mass and spin of BHs surrounded by matter, e.g. accretion disks, may also be inferred from electromagnetic observations (see [31, 32] for stellar-mass BHs and [33–35] for supermassive BHs).

For this paper, we concentrate on nonprecessing quasicircular BBHs, where the black hole spins are parallel or antiparallel to the total orbital angular momentum of the binary. These configurations are fully described in a three-dimensional parameter space: given the masses $m_{1,2}$ and physical spins $S_{1,2}$, we use the two component spins $\chi_1 = S_1/m_1^2$ and $\chi_2 = S_2/m_2^2$ and the mass ratio, given either as $q = m_1/m_2$ with the convention $m_1 > m_2$, or as the symmetric mass ratio $\eta = (m_1 m_2)/(m_1 + m_2)^2 = q/(1 + q)^2$. The total mass is only a scaling factor, and here we work in units of $m_1 + m_2 = 1$.

As suggested by the post-Newtonian (PN) results [36–39] for radiated energy and angular momentum, and confirmed by many previous studies of NR-calibrated models [9, 18, 40, 41], the two dominant parameter dependencies of both final spin and final mass are on the mass ratio and on some appropriately chosen *effective* spin parameter of the binary, whereas the effects of any difference between the two individual spins are much smaller. However, only by accurately modeling these small unequal-spin effects can the full two-spin information be extracted from GW observations, disentangling any true physical degeneracies from systematic effects due to limitations of the waveform models.

A similar, but more complex, situation is encountered in the calibration of phenomenological models to precessing binaries, where so far a simple waveform model based on a sin-

_____
[*] francisco.forteza@ligo.org
[†] david.keitel@ligo.org
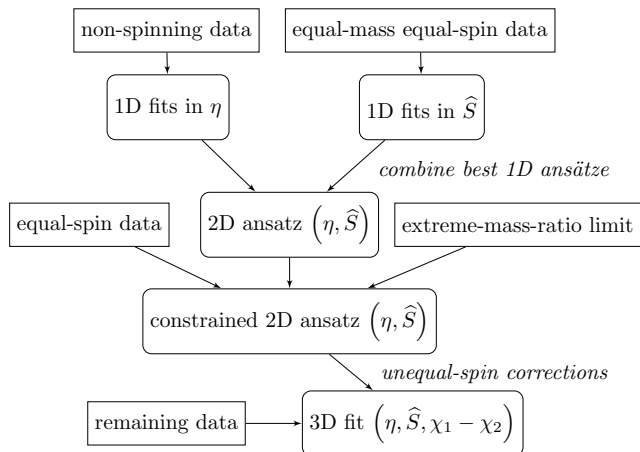[‡] sascha.husa@ligo.org

FIG. 1. Flowchart of the hierarchical step-by-step construction leading to a three-dimensional ansatz and fit for the quantity of interest over the $(\eta, \chi_1, \chi_2) \equiv \left(\eta, \widehat{S}, \Delta\chi\right)$ space.

FIG. 2. The NR data set used in this paper, over mass ratio $q = m_1/m_2$ and the two dimensionless spin components χ_1, χ_2, with color indicating the source catalog. Bright green points are cases removed from the analysis for data-quality reasons, as discussed in Appendix A.

gle effective precession parameter [11–13] has successfully been employed to analyze the first gravitational wave detections [2, 4]. (A complementary precessing analysis [42] based on the model of [20, 43] has also been published recently.) The current work, while mainly a first step to develop models for the full three-dimensional parameter space of nonprecessing waveforms, can also be considered as a toy model for the numerical calibration of subdominant effects in generic binaries, such as precession and higher modes.

For the current three-dimensional aligned-spin parameter space, in addition to the hierarchy of mass ratio, effective spin and spin difference, the sampling by available NR simulations still displays significant bias toward simple subsets: namely the one-dimensional subspaces of nonspinning cases (η dependence only) and equal-mass, equal-spin cases ($\eta = 0.25$, $q = 1$, $\chi_1 = \chi_2$, thus effective-spin dependence only) are covered particularly well by existing NR catalogs, while few simulations exist for unequal spins, high spins, and/or very unequal masses. Independently, the extreme-mass-ratio limit ($\eta \to 0$, $q \to \infty$, arbitrary spins) is known analytically [44].

Here we exploit and investigate this structure by parametrizing spin effects in terms of an effective spin and a spin-difference parameter. As the effective spin we choose

$$\widehat{S} = \frac{S}{m_1^2 + m_2^2}, \qquad \text{with } S = m_1^2 \chi_1 + m_2^2 \chi_2, \qquad (1)$$

which has already been found to work well for final-state quantities in [9]. We discuss other possible choices for the effective spin, for which our method also works robustly, in Appendix C. For spin difference, we use simply $\Delta\chi = \chi_1 - \chi_2$, which makes no assumptions on how spin-difference effects depend on mass ratio.

We develop our hierarchical approach, with the aim to ensure an accurate modeling of the subdominant spin-difference effects, along the lines illustrated as a flowchart in Fig. 1: First we consider the one-dimensional subspaces of nonspinning and of equal-mass-equal-spin black holes. We then combine and generalize these subspace fits, adding additional degrees of freedom to cover the entire two-dimensional space of equal-spin black holes, but constrain the generalized ansatz with information from the extreme-mass-ratio limit. In a third step, we investigate the leading subdominant terms, which are
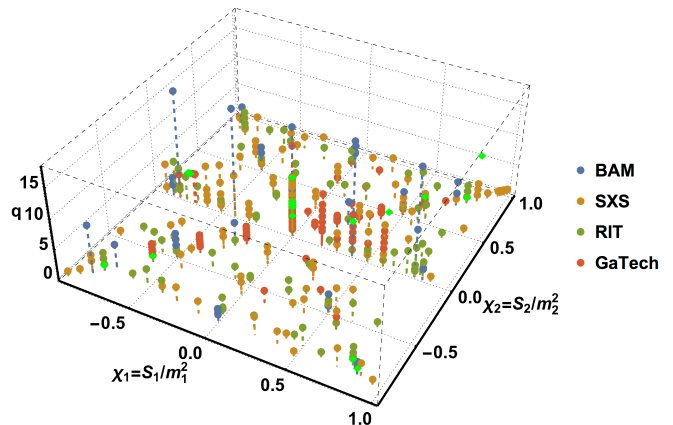
linear in the difference between spins, and also identify additional nonlinear spin-difference terms. We finally produce a three-dimensional fit to the complete data set with the hierarchically constructed ansatz. This way, we can construct a full ansatz with a relatively low number of free fit coefficients and avoid overfitting of spurious effects only due to small sample sizes, while still capturing the essential physical effects that are known from the well-constrained regions.

At each step, we evaluate the performance of different fit choices by several quantitative measures: by the overall residuals, by the Akaike and Bayesian information criteria (AICc, BIC, [45, 46]), and by how well determined the individual fit coefficients are. The information criteria are model selection tools to choose between fits with comparable goodness of fit but different degrees of complexity, i.e. they penalize high numbers of free coefficients. See Appendix B for details on these statistical methods.

Previous published fits for final spin and/or mass include [9, 27, 47–55], and we will compare our results to the most recent results in the literature, including both fits across the full three-dimensional nonprecessing parameter space, and the effective-spin-based fits from [9], which were used in the aligned-spin IMRPhenomD waveform model [9, 10] and (with in-plane-spin corrections) in the precessing IMRPhenomPv2 model [11–13]. These will be referred to as the "PhenomD fits" in the following. The plan of the paper is as follows: We first describe our data sets, built from several NR catalogs, in Sec. II, together with the available extreme-mass-ratio limit information. In Sec. III we develop the general fitting recipe for the example of the final spin. We then apply our method also for final mass – or equivalently, radiated energy – in Sec. IV, which illustrates some of the specific choices and adaptations required to apply the general method to each quantity. We summarize our method and results in Sec. V, and give additional details about NR data, fit construction, and fit uncertainty estimates in Appendices A–D.

Our fits for final spin and radiated energy are also provided in Mathematica and python formats as supplementary material [56], and implemented under the label "UIB2016" in the `nrutils.py` package of LALInference [57, 58]. This final version of the paper uses a larger NR calibration set than the initial arXiv submission (1611.00332v1), with fit results fully consistent but better constrained.
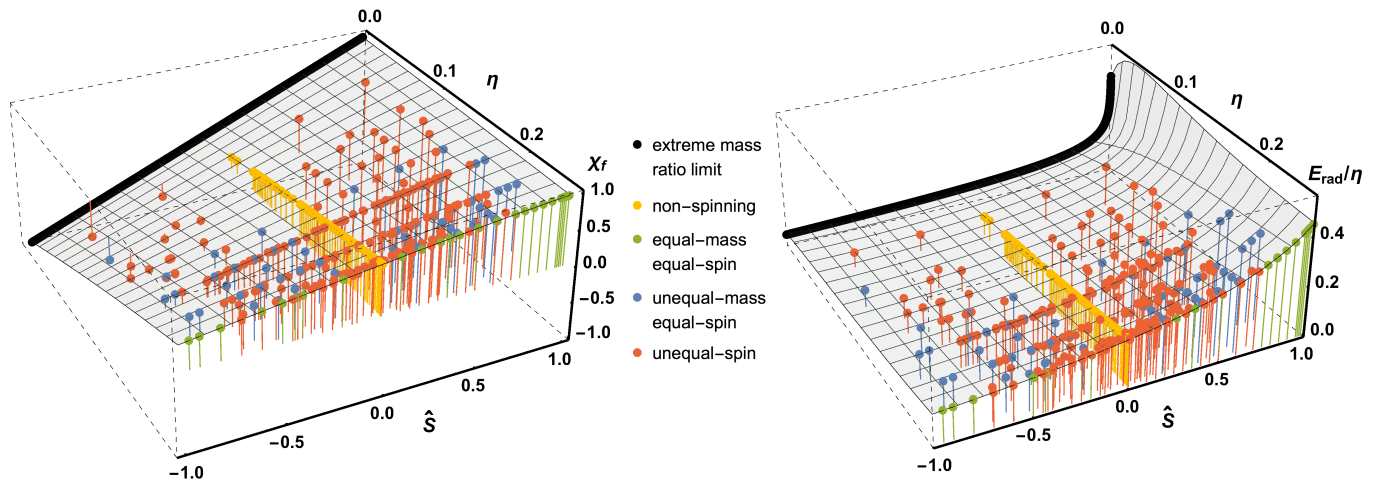
FIG. 3. Input data plotted against symmetric mass ratio η and effective spin $\widehat{S}$. Data consist of the combined set of NR simulations (colored points) and the analytically known [44] extreme-mass-ratio behavior (black line). Left panel: final spin; right panel: radiated energy rescaled by η. Both χ_f and E_{rad}/η follow a smooth surface in this space, and the well-constrained 1D subspaces together already give a good indication of its curvature.

II. INPUT DATA

A. Numerical Relativity data sets

We combine four data sets of aligned-spin numerical relativity BBH simulations from independent codes and sources: the public catalogs of SXS [59, 60], RIT [52, 61] and GaTech [62, 63]; as well as a set of our own simulations with the `BAM` code [9, 64, 65], including 27 new cases for which initial configurations and results are listed in Table XIV in Appendix A. After removing 16 cases from the combined data set due to data quality considerations as discussed in Appendix A, we have 161 cases from the SXS catalog, 107 from RIT, 114 from GaTech and 45 from `BAM`; for a total of 427 cases. The sampling of our three-dimensional parameter space by the four data sets is shown in Fig. 2. The initial arXiv version of this paper (1611.00332v1) used a smaller data set of 256 NR cases, with the increase coming from an update of the public SXS catalog and new RIT results from [61].

To obtain a qualitative understanding of the hierarchical structure in the two-dimensional parameter space of mass ratio and effective spin, in Fig. 3 we show the NR data set over the $(\eta, \widehat{S})$ plane together with the analytical extreme-mass-ratio results, discussed below in Sec. II B. For both final spin and radiated energy, we find a reasonably smooth surface spanned by the NR data points. These plots already suggests that – together with the known extreme-mass-ratio results to compensate the sparsity of NR simulations at increasingly unequal masses – good one-dimensional fits in the two best-sampled one-dimensional subsets (equal-mass-equal-spin and nonspinning BHs) will significantly constrain any two-dimensional fits.

Furthermore, as a first quantitative check that the assumption of a hierarchical structure in three-dimensional BBH parameter space holds, with unequal-spin effects subdominant to the dependence on η and $\widehat{S}$, we can study the residuals of this data set under the two-dimensional PhenomD fits [9]. For final spin, we find that 90% of relative errors are below 3%. (The only cases over 10% are those with absolute values close to zero, where relative error is not a good measure, and absolute errors (residuals) are limited to $\Delta\chi_f \leq 0.025$.) Still, this com-
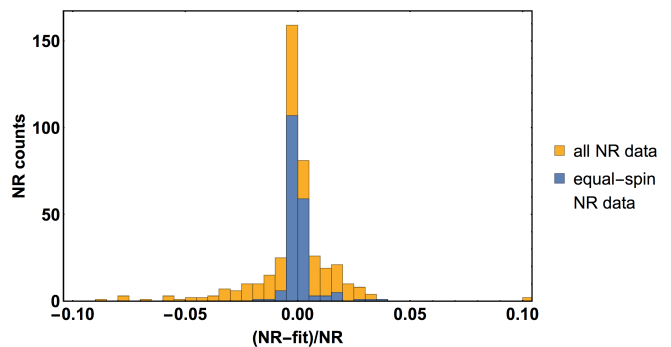

FIG. 4. Relative errors in final spin of the combined NR data set for this paper under the two-dimensional PhenomD fit [9].

parison suggests that unequal-spin effects make a large contribution to these small errors, as shown by four times smaller 90% quantiles when restricting to equal-spin cases only. See also Fig. 4 for histograms of these distributions. For radiated energy, 90% of relative errors are below 2%, with a reduction of that quantile by 1.4 for equal-spin cases only, indicating that spin-difference effects are even smaller for this quantity, which we will also see confirmed in our final results.

For details about extraction of final-state quantities, NR data quality and weight assignment, see Appendix A. As explained there, we do not have a full set of NR error estimates available, so we assign heuristic fit weights to each case based on the expected accuracy of the respective NR code in that particular parameter space region. For example, high-mass-ratio cases are down-weighted more for puncture codes.

B. Extreme-mass-ratio limit

The computational cost of numerical simulations of BH binaries in full general relativity diverges in the extreme-mass-ratio limit $\eta \to 0$. However, this limit is also equivalent to the much simpler case of a test particle orbiting a Kerr black hole. The energy and orbital angular momentum for that configuration have long been known analytically [44]: inserting the radius of the innermost stable circular orbit (ISCO) from Eq. (2.21) of [44] into Eqs. (2.12) and (2.13) of the same reference yields the test-particle energy (equivalent to the radiated

energy) and orbital angular momentum at ISCO:

$$E_{\text{ISCO}}(\eta, \chi) = \eta \left(1 - \sqrt{1 - \frac{2}{3\,\rho_{\text{ISCO}}(\chi_{\text{f}})}} \right), \quad \text{(2)a}$$

$$L_{\text{orb,ISCO}}(\eta, \chi) = \frac{2\eta \left(3\,\sqrt{\rho_{\text{ISCO}}(\chi)} - 2\chi \right)}{\sqrt{3\,\rho_{\text{ISCO}}(\chi)}}, \quad \text{(2)b}$$

with

$$\rho_{\text{ISCO}}(\chi) = 3 + Z_2 - \text{sgn}(\chi)\,\sqrt{(3 - Z_1)(3 + Z_1 + 2\,Z_2)}, \quad \text{(3)a}$$

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

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

Note that both E_{ISCO} and $L_{\text{orb,ISCO}}$ depend linearly on η.

In the test-particle limit, the small BH plunges after reaching the ISCO, and further mass loss scales with η^2 [66]. Similar to previous work [47, 52, 55, 67], we will exploit this fact to compute the final spin and radiated energy to linear order in η from the analytical expressions, Eq. (2), holding at the ISCO. To linear order in η, we thus simply have $E_{\text{rad}} = E_{\text{ISCO}}$ or $M_{\text{f}} = 1 - E_{\text{ISCO}}$ for the final mass, and for the final spin χ_{f} we obtain the implicit equation

$$\chi_{\text{f}}\,M_{\text{f}}(\eta, \chi_{\text{f}})^2 = L_{\text{orb,ISCO}}(\eta, \chi_{\text{f}}) + S_1 + S_2, \quad \text{(4)}$$

where the individual BH spins can be written in terms of our effective spin as

$$S_1 + S_2 = (1 - 2\eta)\,\widehat{S}. \quad \text{(5)}$$

Equation (4) can then be solved numerically for the final spin χ_{f} as a function of η and of the effective spin $\widehat{S}$. Since this result holds to linear order in η, and assuming that the final spin and mass are regular functions of η, we have thus essentially computed the derivatives $\partial E_{\text{rad}}/\partial\eta$ and $\partial\chi_{\text{f}}/\partial\eta$ at $\eta = 0$, in addition to the values at $\eta = 0$, which are $E_{\text{rad}}(0) = 0$ and $\chi_{\text{f}}(0) = S_1/M^2$.

Additionally, assuming that the final state is indeed a Kerr BH, its final spin has to satisfy $\chi_{\text{f}} \leq 1$. One would also expect the final spin for maximal effective spin, $\widehat{S} = 1$, to decrease monotonically with increasing η. To construct an accurate fit in a neighborhood of $\widehat{S} \to 1$ that satisfies these expectations – in particular the Kerr limit – we will constrain our ansatz with the analytically computed value of $\chi_{\text{f}}' = \partial\chi_{\text{f}}/\partial\eta$ at $(\eta = 0, \widehat{S} = 1)$. By perturbing Eq. (4) around $\{\eta \to 0, \chi_{\text{f}} \to 1\}$ to linear order before taking the derivative in η at the same point, we find

$$\chi_{\text{f}}'\left(\eta \to 0, \widehat{S} \to 1\right) = 0. \quad \text{(6)}$$

Several variations of this procedure have been used for previous final-spin fits, and differences are due to previous works neglecting the radiated energy in Eq. (4) [47, 52], or not enforcing the derivative for satisfying the Kerr limit [55].

III. FINAL SPIN

We will now first develop the details of our hierarchical fitting procedure for the example of the final spin of BBH merger remnants, giving more detail here than we will do for the radiated energy in Sec. IV.

A. Choice of fit quantity

We first need to decide which quantity exactly we want to fit. It appears natural to fit a quantity related to the "final" orbital angular momentum L_{orb} near merger, i.e. separating out the known initial spins S_i. This is particularly useful in connection with the extreme-mass-ratio limit, since with Eq. (2)b, L_{orb} is linear in η to leading order. We can use the relation from Eq. (4) between L_{orb} and the dimensionless Kerr parameter χ_{f} of the remnant BH, $M_{\text{f}}^2 \chi_{\text{f}} = L_{\text{orb}} + S_1 + S_2 = L_{\text{orb}} + S$, also outside the extreme-mass-ratio limit. Here M_{f} is the final mass of the remnant BH.

Instead of the actual angular momentum L_{orb}, we take the liberty of fitting the quantity $L_{\text{orb}}' = M^2 \chi_{\text{f}} - S$, where (as throughout the paper) M is set to unity. This way, all fit results are easily converted to the final Kerr parameter χ_{f} by adding the total initial spin S, and no correction for radiated energy has to be applied.

B. One-dimensional subspace fits

Motivated by the the unequal sampling of the parameter space by NR simulations, as visualized in Fig. 3, we start our hierarchical fit development with the simplest and best-sampled subspaces of the NR data set, constructing one-dimensional fits $L_{\text{orb}}'\left(\eta, \widehat{S} = 0\right)$ and $L_{\text{orb}}'\left(\eta = 0.25, \widehat{S}\right)$ over 92 nonspinning and 37 equal-mass-equal-spin cases. We do not restrict ourselves to polynomial fits, and also include ansätze in the form of rational functions. We have also found good fits for more general functions, but omit these here since we have not explored that option systematically.

All fits are performed with Mathematica's `NonlinearModelFit` function, but also partially cross-checked with the `nls` package of R. Since rational functions can have singularities, codes such as `NonlinearModelFit` may not converge to a valid solution without good starting values for the coefficients. We solve this problem by first performing a sufficiently high-order polynomial fit, from which we compute a Padé approximant at the desired order, and use the coefficients of this approximant as starting values for the rational-function fit. We denote rational functions with a numerator of polynomial order m and denominator of polynomial order k as an ansatz of order (m, k). Before fitting to the NR data, all ansätze are constrained by two facts: For nonspinning BBHs, both χ_{f} and L_{orb}' have to vanish for $\eta \to 0$, so that there can be no constant term in the ansatz. Furthermore, from the extreme-mass-ratio prediction Eq. (2), it also follows that the spin-independent coefficient linear in η is $2\sqrt{3}$ (see also [68]). We will include spin-dependent information linear in η in Sec. III C.

Thus, we obtain $L_{\text{orb}}'\left(\eta, \widehat{S} = 0\right)$ fits for a large set of polynomial and rational functions. Several of them produce competitive goodness of fit, as measured by the root-mean-square-error (RMSE) or the full distribution of residuals. However, we do not want to overfit the data, which could induce spurious oscillations in the region of very unequal BH masses that is not covered by NR data. Hence, we rank the fits by information criteria penalizing superfluous free coefficients.

Figure 5 shows the top-ranked fit in terms of Schwarz's Bayesian information criterion (BIC), which is a rational func-
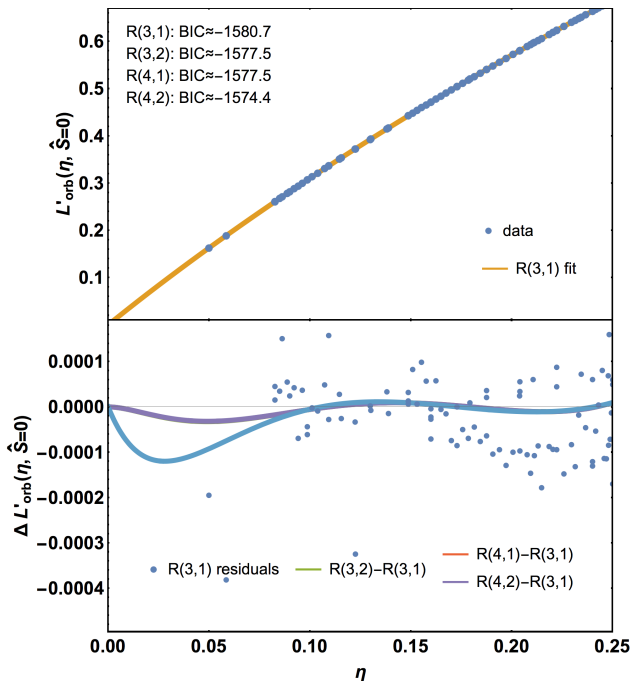
FIG. 5. Nonspinning NR data and one-dimensional $L'_{\rm orb}\left(\eta, \widehat{S}=0\right)$ fit as a function of mass ratio η. Top panel: best fit in terms of the Bayesian information criterion (BIC), a rational function R(3,1), see Eq. (7). Lower panel: residuals ($\Delta L'_{\rm orb}=$ data$-$fit) of this fit (points) and differences from the three next-best-ranking fits in terms of BIC (lines). See also Fig. 29 in Appendix B for an illustration of BIC ranking for this example.

	Estimate	Standard error	Relative error [%]
a_2	3.833	0.085	2.2
a_3	-9.49	0.24	2.5
a_5	2.513	0.046	1.8

TABLE I. Fit coefficients for the one-dimensional nonspinning $L'_{\rm orb}\left(\eta, \widehat{S}=0\right)$ fit over 92 NR cases, along with their uncertainties (standard errors) and relative errors (Std.err./estimate).

tion of order (3,1):

$$L'_{\rm orb}\left(\eta, \widehat{S}=0\right) = \frac{1.3a_3\eta^3 + 5.24a_2\eta^2 + 2\sqrt{3}\eta}{2.88a_5\eta + 1}. \quad (7)$$

The fit coefficients a_i along with their uncertainties are given in Table I; all are well determined. The exact ranking of fits can depend on the choice of fit weights (see Appendix A) and on the ranking criterion, but we find that Eq. (7) is top-ranked by both BIC and AICc. While only ranked 6th by RMSE, none of the considered fits is better than Eq. (7) by more than 6% in that metric either. Additionally, under variations of the weighting scheme, this is robustly the fit among the top-ranked group – by all three criteria – with the lowest number of fitting coefficients, indicating it is a robust choice. For comparison, a simple third-order polynomial (two free coefficients) is disfavored clearly, by more than a factor of 8 in RMSE and an offset of +452 in BIC, and a fourth-order polynomial (three free coefficients, just as Eq. (7)) by almost a factor of 2 and by +332 respectively.

The lower panel of Fig. 5 also compares the preferred fit both to the NR data and to the three next-best ranking fits by BIC. We find that the residuals are centered around zero with
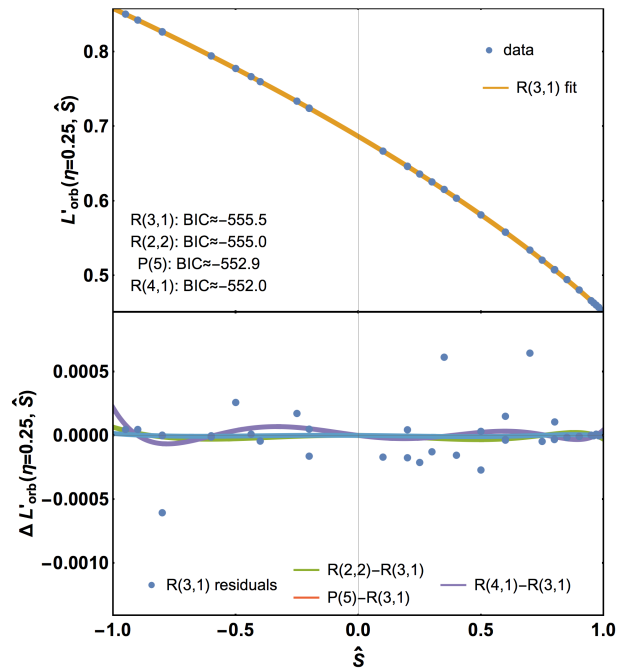
FIG. 6. Equal-mass-equal-spin NR data and one-dimensional fits of $L'_{\rm orb}\left(\eta = 0.25, \widehat{S}\right)$ as a function of effective spin $\widehat{S}$. Top panel: best fit in terms of BIC, a rational function R(3,1), see Eq. (8). Lower panel: residuals of this fit (points) and differences from the three next-best-ranking fits in terms of BIC (lines).

	Estimate	Standard error	Relative error [%]
b_1	1.00096	0.00068	0.1
b_2	0.788	0.042	5.3
b_3	0.654	0.074	11.4
b_5	0.840	0.030	3.6

TABLE II. Fit coefficients for the one-dimensional equal-mass-equal-spin $L'_{\rm orb}\left(\eta = 0.25, \widehat{S}\right)$ fit over 37 NR cases.

no major trends, while the differences among high-ranked fits are much smaller than the scatter of residuals for the well-covered high-η range, and that the "systematic uncertainty", as indicated by the difference of high-ranked fits, is still at the same level even in the extrapolatory low-η region. The BIC ranking for this example is also illustrated in Fig. 29 in Appendix B.

Before the second 1D fit in the effective spin parameter $\widehat{S}$, we need to decide how we are later going to construct a 2D ansatz combining both 1D fits. We can either subtract or divide the NR data by the nonspinning fit, and find that subtraction exhibits a simpler functional form. Thus we decide to construct a 2D ansatz as the sum of nonspinning and spinning contributions. (We will choose a product ansatz for the radiated energy discussed in the next section.) We thus constrain the constant term of the 1D ansatz in $\widehat{S}$ to reproduce the $\eta = 0.25$ nonspinning result, i.e. $L'_{\rm orb}(\eta = 0.25, \widehat{S} = 0)$ must be identical for both 1D fits.

The top-ranked $L'_{\rm orb}\left(\eta = 0.25, \widehat{S}\right)$ fits by BIC are shown in Fig. 6, and again we find a unique top-ranked fit by both AICc and BIC, a rational function of order (3,1):

$$L'_{\rm orb}\left(\eta = 0.25, \widehat{S}\right) = \frac{0.00954b_3\widehat{S}^3 + 0.0851b_2\widehat{S}^2 - 0.194b_1\widehat{S}}{1 - 0.579b_5\widehat{S}} + 0.68637, \quad (8)$$
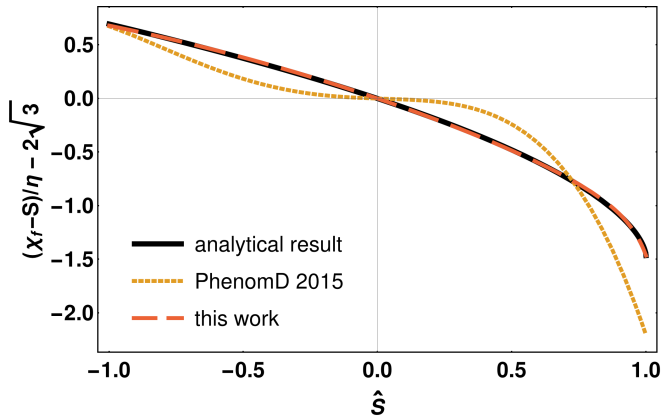
FIG. 7. Extreme-mass-ratio comparison of the rescaled final spin: analytical results from solving Eq. (4), the previous PhenomD final-spin fit of [9], and this work.

	Estimate	Standard error	Relative error [%]
f_{21}	8.774	0.019	0.2
f_{31}	22.83	0.27	1.2
f_{50}	1.8805	0.0025	0.1
f_{11}	4.4092	0.0047	0.1

TABLE III. Fit coefficients for the extreme-mass-ratio limit of the final spin, fitted to discretized analytical results. The fourth coefficient, f_{11}, is fixed by the derivative constraint in Eq. (13) and its estimate and error computed from the others.

with four fit coefficients b_i as listed in Table II. This ansatz is ranked 8th by RMSE, but with only 3% difference from the lowest RMSE, which is attained by a P(5) fit with one more coefficients, marginally disfavored by about +1.7 AICc and +2.6 in BIC. The best three-coefficient fit R(2,1) is significantly worse, with differences of over +40 in BIC/AICc and 40% in RMSE. Again, the distribution of residuals is well behaved, and differences between the four top-ranked fits by BIC are smaller than the scatter of residuals.

C. Two-dimensional fits

Next, we want to construct a two-dimensional fit covering the $(\eta, \widehat{S})$ space, as it was illustrated in Fig. 3, by combining both the 1D subspace fits and the extreme-mass-ratio limit. As discussed above, we take the sum of Eq. (7) and the spin-dependent terms of Eq. (8). We introduce the necessary flexibility to describe 2D curvature and the extreme-mass-ratio limit by generalizing the $\widehat{S}$-dependent terms, inserting a polynomial of order J in η for each b_i through the substitution

$$b_i \to b_i \sum_{j=0}^{j=J} f_{ij} \eta^j . \quad (9)$$

The general 2D ansatz is thus

$$L'_{\text{orb}}\left(\eta, \widehat{S}\right) = L'_{\text{orb}}\left(\eta, 0\right) + L'_{\text{orb}}\left(0.25, \widehat{S}, f_{ij}\right) - L'_{\text{orb}}\left(0.25, 0\right) . \quad (10)$$

Here we choose to expand to third order in η ($J = 3$), which is the lowest order leaving enough freedom to incorporate all available constraints from the 1D fits and the extreme-mass-ratio limit, and, as evidenced by the residuals we find below,
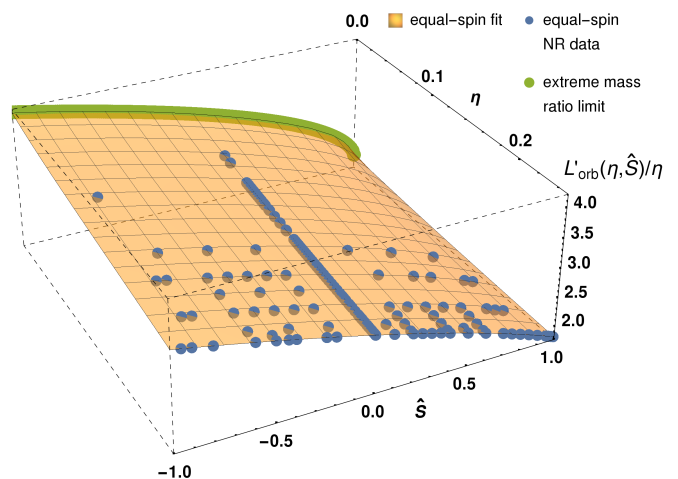

FIG. 8. Two-dimensional $L'_{\text{orb}}\left(\eta, \widehat{S}\right)$ fit, visualized as $L'_{\text{orb}}\left(\eta, \widehat{S}\right)/\eta$. Application of the extreme-mass-ratio limit helps in avoiding extrapolation artifacts which would otherwise appear in low-η, high-$|\widehat{S}|$ regions that are uncovered by NR simulations.

also high enough to adequately model this data set. Of the resulting 16 coefficients, the three f_{i0} in the numerator must vanish to preserve the $L'_{\text{orb}}\left(\eta = 0, \widehat{S}\right) = 0$ limit, while consistency with the equal-mass fit from Eq. (8) provides four constraints which we use to fix the f_{i3} terms:

$$f_{i3} = 64 - 64f_{i0} - 16f_{i1} - 4f_{i2} . \quad (11)$$

Four more coefficients are fixed by the extreme-mass-ratio information discussed in Sec. II B: we re-express Eq. (4) in terms of L'_{orb}/η and fit the discretized quantity

$$\lim_{\eta \to 0} \frac{L'_{\text{orb}}\left(\eta, \widehat{S}\right)}{\eta} - 2\sqrt{3} = \lim_{\eta \to 0} \frac{\chi_{\text{f}}\left(\eta, S\right) - S}{\eta} - 2\sqrt{3} \quad (12)$$

where $2\sqrt{3}$ is the linear contribution from the nonspinning part (cf. Eq. (7)) and the $\chi_{\text{f}}\left(\eta \to 0, S\right)$ values are obtained by solving Eq. (4) numerically for small η. Before fitting, we apply the derivative constraint from Eq. (6), which for the sum ansatz Eq. (10) implies a coefficient constraint

$$f_{11} \to 0.345225f_{21} + 0.0321306f_{31} - 3.66556f_{50} + 7.5397. \quad (13)$$

We find this extra physical constraint to be essential in avoiding superextremal χ_{f} results due to fitting artifacts. The extreme-mass-ratio limit fit coefficients are listed in Table III, and the improved agreement between analytical results and this new fit, as compared with the previous fit of [9], is illustrated in Fig. 7.

In summary, after constraining to the well-covered one-dimensional NR data subsets and the analytically known extreme-mass-ratio limit, the 2D ansatz from Eq. (10) has reduced from 16 to 5 free coefficients. We fit it to the 60 remaining equal-spin NR cases that were not yet included in the 1D subsets. To remove possible singularities in the $(\eta, \widehat{S})$ plane for this rather general rational ansatz, we set the least-constrained denominator coefficient f_{52} to zero. Thus we obtain a smooth four-coefficient fit as shown in Fig. 8. It has RMSE four times larger than the 1D η fit and twice as
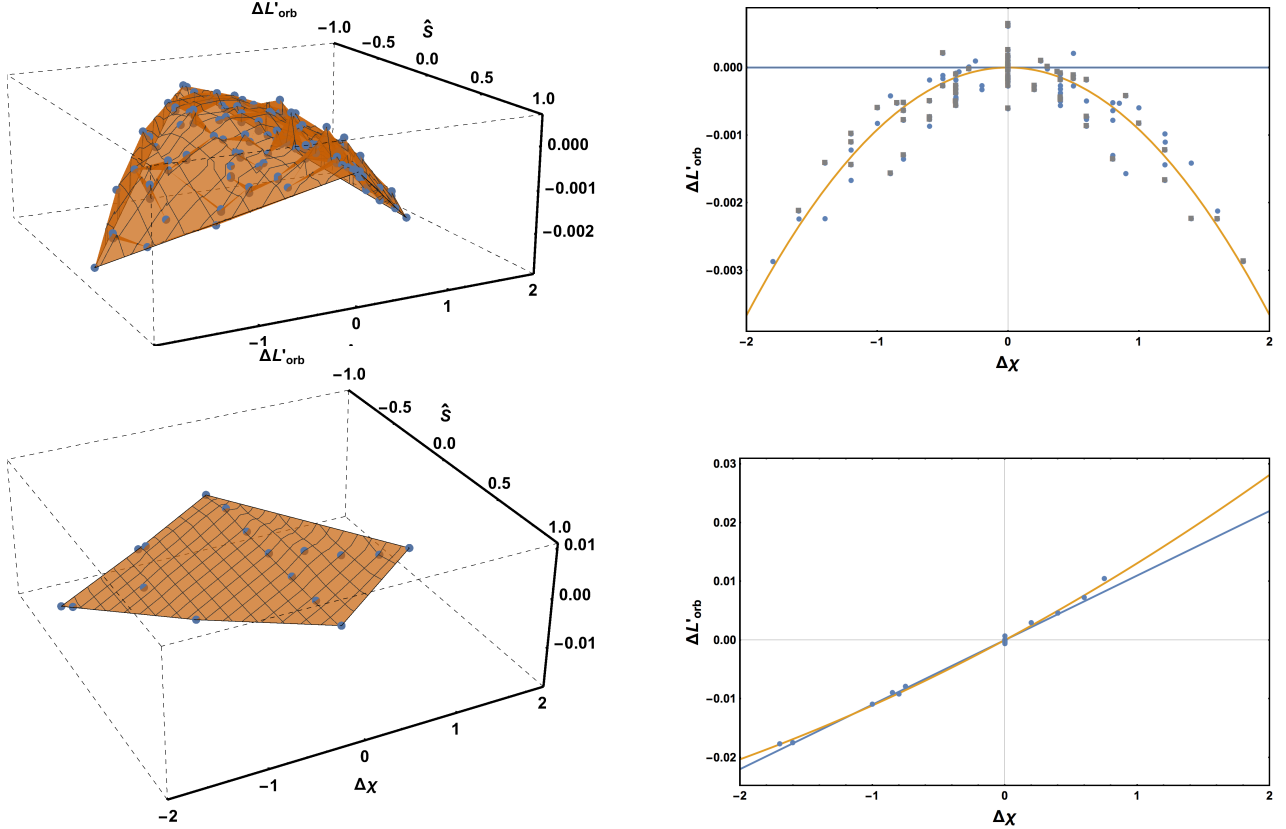
FIG. 9. Examples of spin-difference behavior at fixed mass ratios, for residuals $\Delta L'_{\text{orb}}$ after subtracting the two-dimensional $L'_{\text{orb}}\left(\eta, \widehat{S}\right)$ fit, as defined in Eq. (14). Top row: $q = 1$; lower row: $q = 4$; left column: surfaces in $\left(\widehat{S}, \Delta\chi, \Delta L'_{\text{orb}}\right)$ space; right column: projections onto the $\Delta\chi$ axis with linear and quadratic fits. At equal mass, the surface is parabolic, with the linear term (blue line) and mixture term (not shown) vanishing, but a clear quadratic dependence (orange line). At $q = 4$ and other intermediate mass ratios, the surface is very close to flat and the linear term dominates.

high as the $\widehat{S}$ fit, but these residuals are still smaller than expected unequal-spin effects and rather noisily distributed without any clear parameter-dependent trends, thus indicating that the 2D fit sufficiently captures the dominant $\left(\eta, \widehat{S}\right)$ dependence to form the basis of studying subdominant spin-difference effects. The least-constrained coefficient at this point has a relative error of about 25%, which is good enough to keep it in the ansatz for the next, 3D step where we will refit to a larger data set.

D. Unequal-spin contributions and 3D fit

Now the final step in the hierarchical procedure is to explore the subdominant effects of unequal spins, parametrized by the spin difference $\Delta\chi = \chi_1 - \chi_2$. We first study the residuals of the 238 unequal-spin NR cases under the equal-spin 2D fit:

$$\Delta L'_{\text{orb}}\left(\eta, \widehat{S}, \Delta\chi\right) := L'_{\text{orb,NR}}\left(\eta, \widehat{S}, \Delta\chi\right) - L'_{\text{orb}}|_{\text{eqSpinFit}}\left(\eta, \widehat{S}\right).$$
(14)

We do this at fixed steps in mass ratio, having sufficient numbers of NR cases for this analysis at mass ratios $q = \{1, 1.33, 1.5, 1.75, 2, 3, 4, 5, 6, 7, 8\}$. This per-mass-ratio analysis is only used to guide the construction of the full 3D ansatz and as a consistency check, while the final full 3D fit will consist of fitting the constrained 2D ansatz plus spin-difference terms directly to the full data set.

At each mass ratio, we visually inspect the residuals, which span 2D surfaces in $\left(\chi_1, \chi_2, L'_{\text{orb}}\right)$ or, equivalently, $\left(\widehat{S}, \Delta\chi, L'_{\text{orb}}\right)$

space. As illustrated in Fig. 9, we find surfaces close to a plane, indicating a dominant linear dependence on $\Delta\chi$ and possibly a mixture term $\widehat{S}\,\Delta\chi$. The exception is at equal masses, where quadratic curvature in the $\Delta\chi$ dimension dominates. In this case, exchange of χ_1 and χ_2 yields an identical binary configuration, so that terms linear in $\Delta\chi$ indeed have to vanish. We have also exploited this fact in the $q = 1$ analysis by adding mirror duplicates of each NR data point. Motivated by these empirical findings and symmetry argument, we introduce up to three spin-difference terms,

$$\Delta L'_{\text{orb}}\left(\eta, \widehat{S}, \Delta\chi\right) = A_1(\eta)\,\Delta\chi + A_2(\eta)\,\Delta\chi^2 + A_3(\eta)\,\widehat{S}\,\Delta\chi.$$
(15)

The full 3D ansatz is then simply the sum of Eqs. (10) and (15):

$$L'_{\text{orb}}\left(\eta, \widehat{S}, \Delta\chi\right) = L'_{\text{orb}}\left(\eta, \widehat{S}\right) + \Delta L'_{\text{orb}}\left(\eta, \widehat{S}, \Delta\chi\right).$$
(16)

Adding higher orders in the effective spin or spin difference is not supported by visual inspection. At each mass ratio, we now perform four fits in $\Delta\chi$ for the values of the A_i: linear, linear+quadratic, linear+mixed, or the sum of all three terms. Examples are also shown in Fig. 9.

We then collect the coefficients of each of these fits and use them as data $A_i(\eta)$ to be fitted as functions of mass ratio (see the 'per-mass-ratio data' in Fig. 10), using as weights the fit uncertainty from each mass ratio rescaled by the average data weight for that mass ratio. We also apply what we know about the extreme-mass-ratio and equal-mass limits: all three $A_i(\eta)$
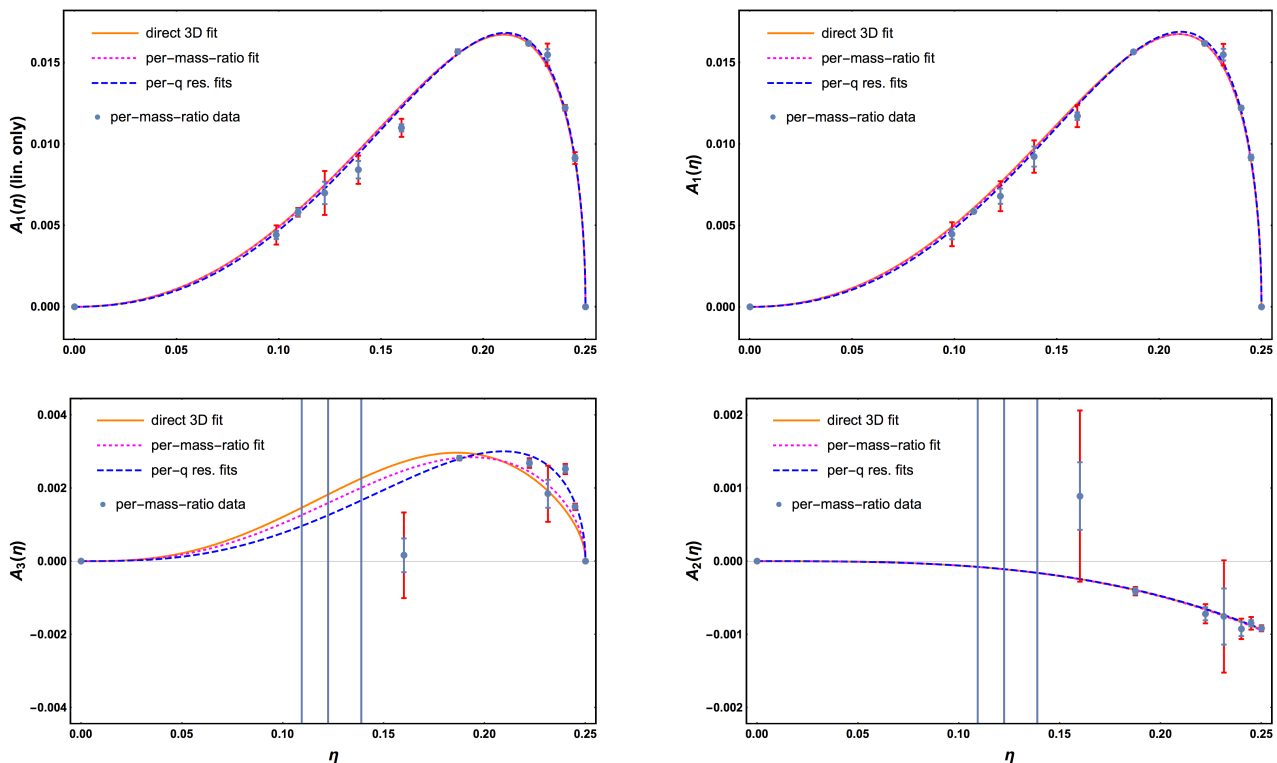
FIG. 10. Spin-difference behavior of final-spin data after subtraction of the two-dimensional $L'_{orb}(\eta, \widehat{S})$ fit, showing the results of fits as in Fig. 9 at η steps corresponding to $q = \{1, 1.33, 1.5, 1.75, 2, 3, 4, 5, 6, 7, 8\}$ and three estimates for the three ansatz functions $A_i(\eta)$ from Eqs. (15) and (19): (i) unequal-spin part of the final 3D fit from Eq. (16) ("direct 3D fit"), (ii) fit of the unequal-spin terms from Eq. (19) ("fit to residuals") to the residuals of the 2D fit from Eq. (10) over all mass ratios, (iii) fits of Eq. (19) to the per-mass-ratio results. Top-left panel: linear term A_1 only. The remaining panels are for the combined linear+quadratic+mixture fit, in clockwise order: linear term A_1, quadratic term A_2 and mixture term A_3. The A_1 results from the combined fit are very similar to those from the linear-only fit, demonstrating the robustness of extracting leading-order spin-difference effects. For the two lower panels, data points for low η are outside the displayed range, but the error bars are huge and hence this region does not contribute significantly to the weighted per-mass-ratio fits. In the direct 3D fit to the full data set, however, low-η information can be better incorporated, leading to the somewhat different shape of the mixture-term fit. See Sec. III E for more discussion of how well constrained these shapes actually are with the current data set.

have to vanish in the limit $\eta = 0$, and the A_1, A_3 linear in $\Delta\chi$ have to vanish for $\eta = 0.25$. We thus choose ansätze of the form

$$A_i = d_{i0}\, \eta^{p_i} \left(\sqrt{1-4\eta}\right)^{q_i} (1 + d_{i1}\eta) \qquad (17)$$

for $A_{i=1,3}$ linear in $\Delta\chi$, where the factor $\left(\sqrt{1-4\eta}\right)^{q_i}$ is motivated from post-Newtonian (PN) results [38, 39], and

$$A_2 = d_{20}\, \eta^{p_2} \left(1 + d_{21}\left(\sqrt{1-4\eta}\right)^{q_2}\right) \qquad (18)$$

for the term quadratic in $\Delta\chi$. We find that the data can be well fit without any higher-order terms and by reducing some of the freedom of these three terms exploratory fits keeping all coefficients free give results close to integer numbers for the p_i, $q_i = 1$ and $d_{21} = 0$. Hence we choose the three parsimonious ansätze

$$A_1(\eta) = d_{10}(1 - 4\eta)^{0.5}\eta^2 (d_{11}\eta + 1) \qquad (19)a$$

$$A_2(\eta) = d_{20}\eta^3 \qquad (19)b$$

$$A_3(\eta) = d_{30}(1 - 4\eta)^{0.5}\eta^3 (d_{31}\eta + 1) . \qquad (19)c$$

The blue points and lines in Fig. 10 show these per-mass-ratio results. The shape and numerical results of the dominant linear term A_1 are quite stable under adding one or two of the

other terms. Fitting two terms, either linear+quadratic or linear+mixture, yields quadratic/mixture effects of very similar magnitude, with the quadratic term following the same basic shape (an intermediate-mass-ratio bulge) as the other two. However, combining all three terms, the results match better with the expectations from symmetry detailed before, with the bulge shape limited to the linear and mixture terms while the quadratic term provides a correction mostly at similar masses.

Using again the $q = 1$, $\widehat{S} = 0$ and $\eta \to 0$ constraints on the general ansatz from Eq. (16), we end up with a total of nine free coefficients in this final step. We now fit to 298 cases with arbitrary spins not yet used in the 1D fits, with results given in Table IV. Together with the coefficients from Tables I–III, these fully determine the fit. To convert back from our fit quantity L'_{orb} to the actual dimensionless final spin χ_f, just add the total initial spin $S = m_1^2 \chi_1 + m_2^2 \chi_2$.

We find that the data set is sufficiently large and clean, and the equal-spin part modeled well enough from the 2D step, to confidently extract the linear spin-difference term and its η-dependence, which is stable when adding the other terms; and to find some evidence for the combined mixture and quadratic terms, whose shape however is not fully constrained yet.

	Estimate	Standard error	Relative error [%]
d_{10}	0.322	0.020	6.2
d_{11}	9.33	0.87	9.3
d_{20}	−0.0598	0.0021	3.5
d_{30}	2.32	0.28	12.1
d_{31}	−3.26	0.20	6.1
f_{12}	0.512	0.085	16.7
f_{22}	−32.1	3.6	11.3
f_{32}	−154	10	6.5
f_{51}	−4.77	0.34	7.1

TABLE IV. Fit coefficients for the final 3D step of the L'_{orb} fit to 298 cases not yet used in the 1D fits of Sec. III B.

E. Fit assessment

In Fig. 10 we also compare the spin-difference terms from this final "direct 3D" fit to those obtained from the per-mass-ratio residuals analysis. The linear term is fully consistent, confirming that it is well determined by the data, while for the quadratic and mixture terms both approaches agree on the qualitative shape, but do not match as closely. Under the chosen ansätze, the 3D fit coefficients even for those terms are tightly determined (see Table IV). However, we have explicitly chosen the spin-difference terms in Eq. (19) to achieve this goal, while several other ansatz choices (changing the fixed exponents of the multiplicative η or $\sqrt{1-4\eta}$ terms, or adding more terms with free coefficients in the η polynomials) can produce fits that are indistinguishable by summary statistics (AICc, BIC, RMSE). Still, most of these have some strongly degenerate and underconstrained coefficients, while the reported fit has the desirable property of sufficient complexity to be within the plateau region of summary statistics while not having any degenerate coefficients.

Yet, the shape of the functions $A_2(\eta)$ and $A_3(\eta)$ for the mixture and quadratic terms is not actually as closely constrained from the current data set as the coefficient uncertainties alone seem to imply, due to this ambiguity in ansatz selection. This becomes clear from the comparison of direct 3D fit and per-mass-ratio analysis in Fig. 10. The per-mass-ratio analysis also demonstrates that the data at mass ratios $\eta < 0.16$ are not yet constraining enough to help characterize these terms. (The error bars are so large, and hence the weights so low, that they effectively do not contribute to the fit.) It also becomes clear that additional unequal-spin data at intermediate mass ratios would be very useful in constraining the $A_{2,3}(\eta)$ functions. Meanwhile, it is important to note again that the leading linear spin-difference term is already determined much more narrowly and robustly with the current data set.

We can further assess the success of the hierarchical 3D fitting procedure by comparing

- a 2D fit (equal-spin physics only) to equal-spin NR cases only (same as in Fig. 8),
- a 2D fit (equal-spin physics only) to all NR data,
- and the 2D part of the full 3D fit.

As shown in Fig. 11, fitting the 2D equal-spin ansatz to the full data set induces strong curvature in the $(\eta, \widehat{S})$ plane, which the full 3D fit is able to correct by the additional degrees of freedom in the spin-difference dimension. This is how it was possible to pull out the subdominant spin-difference effects
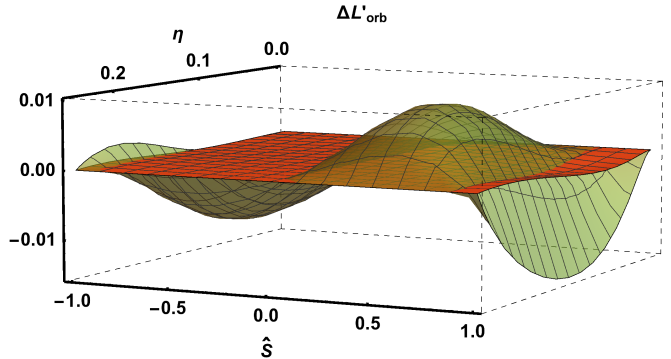

FIG. 11. Green: Difference $\Delta L'_{orb}$ of a 2D fit (equal-spin physics only) to the full data set minus the 2D fit to equal-spin cases only, both including extreme-mass-ratio constraints. The strong curvature at intermediate mass ratios and nonzero spins is due to the equal-spin-physics-only fit trying to compensate for the addition of unequal-spin NR cases.
Orange: Difference $\Delta L'_{orb}$ of the 2D part of the 3D fit to the full data set minus the 2D-only fit to equal-spin data. The bulk of the parameter space is no longer distorted, and only at high effective-spin magnitudes a small opposite effect to the η-dependent behavior of the spin-difference terms (cf. Fig. 10) can be seen.

	N_{data}	N_{coeff}	RMSE	AICc	BIC
1D η	92	3	9.41×10^{-5}	−1590.8	−1580.7
1D $\hat{S}$	37	4	2.05×10^{-4}	− 563.6	− 555.5
2D ($\chi_1 = \chi_2$)	60	4	3.90×10^{-4}	− 880.5	− 870.8
2D all	298	4	8.05×10^{-3}	−2247.4	−2229.0
3D lin	298	6	9.20×10^{-4}	−3628.4	−3602.9
3D lin+quad	298	7	8.28×10^{-4}	−3765.0	−3735.8
3D lin+mix	298	8	8.11×10^{-4}	−3693.4	−3660.6
3D lin+quad+mix	298	9	6.10×10^{-4}	−4087.3	−4050.9

TABLE V. Summary statistics for the various steps of the hierarchical final-spin fit. Note that it is not meaningful to compare AICc and BIC between data subsets of different sizes. There is statistical preference for the 3D fit including all three linear+mixture+quadratic terms, although many different choices of the $A_i(\eta)$ ansatz functions yield similar results with just ± a few percent in RMSE and ± a few in AICc/BIC, so that the shape of the mixture and quadratic terms is not yet fully constrained.

with this enlarged data set. The same conclusion is supported by the comparison of summary statistics between the various steps and 2D/3D fit variants in Table V, showing that the RMSE only increases by 50% from the 2D equal-spin case to the full 3D fit using all data.

The distribution of fit residuals for the full data set, projected onto the $(\eta, \widehat{S})$ plane, is shown in Fig. 12, and a comparison of fit residuals with other previously published fits, over the calibration data set of the current work, is shown as histograms in Fig. 13 and summarized in Table VI along with AICc and BIC metrics. The shape of the distributions is consistent, and for all fits the means are much smaller than the standard deviations, showing no evidence for any systematic bias. Our new fit improves significantly over the previous fit [9] used in the calibration of the IMRPhenomD waveform model [10], and also yields some improvement over recent fits from other groups [55, 61], even when those ansätze are refit to our present NR data set.

Refitting our final hierarchically obtained ansatz directly to the full data set produces slightly better summary statis-
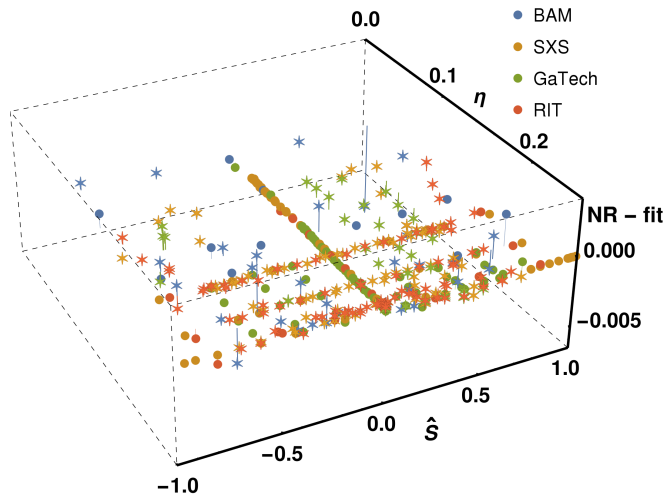
FIG. 12. Residuals of the new 3D final-spin fit, projected to the 2D parameter space of η and $\widehat{S}$. The four NR data sets are distinguished by colors, and unequal-spin points highlighted with stars.
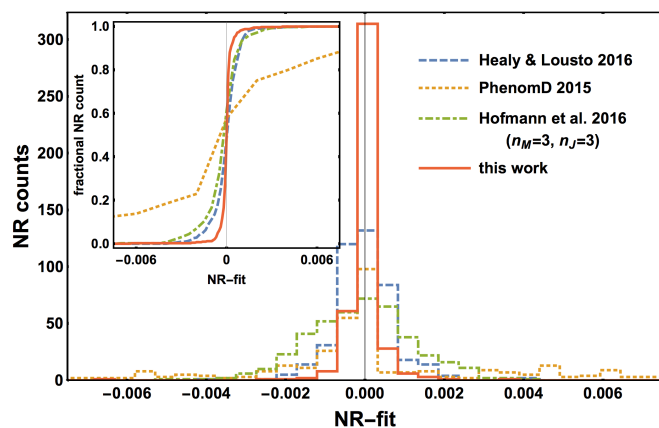

FIG. 13. Fit residuals of the final spin χ_f, for this work and for previously published fits [9, 55, 61], evaluated over the set of 427 NR simulations shown in Fig. 2. Main panel: histograms, with 102 outliers for PhenomD with |NR − fit| > 0.0075 outside of the plot range. Inset: cumulative distributions over the same range.

	N_{coef}	mean	stdev	AICc	BIC
HLZ2014 [52]	19	-4.8×10^{-5}	8.9×10^{-4}	−5141.0	−5061.7
HL2016 [61]	19	8.1×10^{-7}	7.9×10^{-4}	−5358.1	−5278.9
PhenomD [9]	11	-4.7×10^{-5}	7.2×10^{-3}	−3309.0	−3260.9
(refit)	11	-1.7×10^{-4}	7.0×10^{-3}	−3334.5	−3286.5
HBR2016 [55]	6	-1.2×10^{-4}	1.4×10^{-3}	−4717.2	−4689.0
(refit)	6	-1.4×10^{-4}	1.3×10^{-3}	−4791.4	−4763.2
HBR2016 [55]	16	-2.8×10^{-4}	1.2×10^{-3}	−4877.3	−4809.7
(refit)	16	-1.4×10^{-5}	1.0×10^{-3}	−4975.8	−4908.2
This work	16	-2.3×10^{-5}	5.2×10^{-4}	−5991.5	−5923.9
(refit)	16	-2.1×10^{-5}	5.1×10^{-4}	−6011.3	−5943.6
(uniform)	16	-1.2×10^{-5}	5.0×10^{-4}	−5240.1	−5172.5
(uniform refit)	16	-6.9×10^{-6}	4.9×10^{-4}	−5256.8	−5189.2

TABLE VI. Summary statistics for the new final-spin fit compared with previous fits [9, 52, 55, 61], evaluated over the 427 NR simulations shown in Fig. 2. For Hofmann *et al.* [55], both the $(n_M = 1, n_J = 2)$ fit (6 coefficients) and the $(n_M = 3, n_J = 3)$ version (16 coefficients) are listed. The new fit has a total of 16 coefficients calibrated to NR, corresponding to Tables I, II and IV, not counting those constrained from the extreme-mass-ratio limit. We also show results for refitting previous ansätze to the present NR data set, for a refit of our hierarchically obtained ansatz directly using the full data set, and for the same fitting procedure, but using uniform weights.
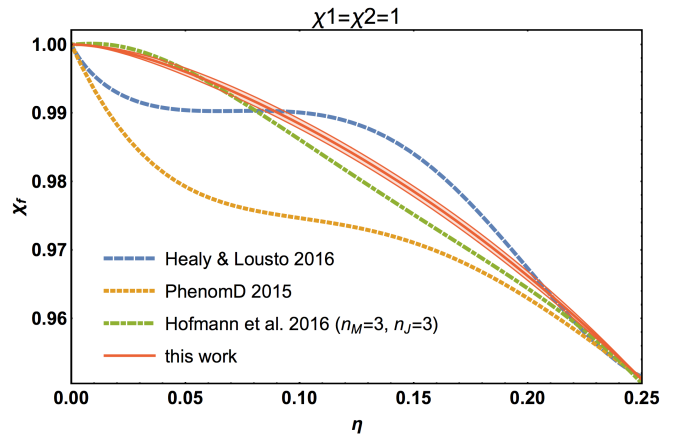

FIG. 14. Comparison of this work with previously published fits [9, 55, 61] in the limit of extremal aligned spins, $\chi_1 = \chi_2 = 1$. The shaded region shows our fit's 90% confidence interval, which is narrow enough to indicate that discrepancies with the referenced fits are significant and due to the different ansatz constructions, especially in the extreme-mass-ratio limit (cf. Sec. II B), and not just a consequence of insufficient data.

tics, but also allows uncertainties from the less well-controlled unequal-spin set to influence the other parts of the fit, while the stepwise fit gives better control over the extreme-mass-ratio behavior and better-determined coefficients for the well-constrained subspaces.

As a further test of robustness, we have repeated the hierarchical fitting procedure with uniform weights instead of the weights used so far and discussed in Appendix A. This yields a fit consistent with our main result, though slightly less well constrained, but still improving over previous fits, thus demonstrating the robustness of the hierarchical fit construction under weighting choice.

We have also verified that our new fit does not violate the $\chi_f \leq 1$ Kerr bound, particularly in the extreme-spin limit ($\widehat{S} = 1$) and at low η, see Fig. 14.

F. Precessing binaries

While some existing final-spin fits [53–55] also include a calibration to precessing cases, it is also possible to use a simple "augmentation" procedure [49] (see also [67]) for aligned-spin-only calibrations by adding the contribution of in-plane spins in quadrature to the aligned-spin fit result:

$$\chi_f^{aug} = \sqrt{\left(\chi_f^{aligned}\right)^2 + \left(S^{in\text{-}plane}/M^2\right)^2} \,. \tag{20}$$

This procedure is known to significantly improve accuracy and reduce bias for precessing binaries. For example, it has been applied to the aligned-spin PhenomD fit [9] for the precessing PhenomPv2 model [12, 13], and to the RIT fit [52] in recent parameter estimation work of the LIGO-Virgo collaboration [4, 42, 69] (including spin evolution according to [41]).

Applying Eq. (20) to our aligned-spin fit, we find a small overshooting of the $|\chi_f| \leq 1$ Kerr bound for mass ratios $q \gtrsim 24$, when the spin magnitude of the heavier BH is very close to extremal, and for certain orientation angles θ_i of the black holes' spins to the angular momentum. The worst cases give an excess in χ_f of about 0.12% at $q \sim 60$ and intermediate opening

angles, comparable to the aligned-spin fit residuals. No overshooting occurs if only the linear-in-η term in the final spin is used. Such a small inaccuracy when extending the aligned-spin fit to precessing cases is in principle not surprising, as this parameter-space region is not covered with NR simulations and hence the fit slope in this region is purely determined by extrapolation between the NR data and the extreme-mass-ratio limit, which we have ensured to be smooth with a flat approach to $\chi_f = 1$ at ($\eta = 0$, $\chi_1 = 1$) (see Sec. II B and Fig. 14). Very small inaccuracies in the intermediate-η extrapolation region can thus lead to a minimal Kerr violation when adding the in-plane spins according to Eq. (20). A clean solution to this issue would require more calibration NR simulations in the critical region and a study of precessing spin contributions in the extreme-mass-ratio limit.

However, as the overshooting is very small, we have investigated two easy *ad hoc* solutions: We could take the worst-case point and enforce our 3D fit to be at or below the corresponding L'_{orb} value for the aligned-spin projection $\chi_1 = \cos(\theta_1)$ by putting a constraint on one of the f_{i2} coefficients. This can remove the overshooting at the worst-case point and nearby, but not over the whole problematic region, as the fit still has enough freedom in other parameters. But the accuracy of the aligned-spin fit already suffers from this one extra constraint, and using constraints on more than one coefficient to pull down the augmented χ_f over a wider parameter region is fully prohibitive because insufficient freedom will remain in the fit to properly calibrate to the actual NR data.

Hence, we opt for an even simpler solution, truncating the augmentation from Eq. (20) at unity: $\chi_f = \min\left(\chi_f^{aug}, 1.0\right)$. This is justified as the overshooting is very small, on the order of the fit residuals, and limited to an extremal parameter-space region. The need for this *ad hoc* truncation will reduce or become obsolete when low-η-high-spin NR simulations and/or precessing extreme-mass-ratio information become available. A detailed comparison of fit accuracies over a representative set of precessing NR runs is left to future work.

IV. FINAL MASS AND RADIATED ENERGY

To fit the final mass of remnant BHs from BBH mergers, we use the same hierarchical approach as for the final spin, summarized in Fig. 1, with only minor modifications. The quantity we are going to fit here is the dimensionless radiated energy, $E_{rad} = M - M_f = 1 - M_f$. For $\eta = 0$, it has to vanish even in spinning cases, while the analytical expectation for the leading order in η, as $\eta \to 0$, is $E_{rad}(\eta)/\eta \sim 1 - \left(2\sqrt{2}\right)/3$. We construct the two-dimensional $E_{rad}\left(\eta, \widehat{S}\right)$ ansatz as a product of the 1D ansätze, instead of a sum as in the final-spin case.

In principle, the fitting procedure is robust enough to use either a sum or product ansatz for either final-state quantity. Actually, by carrying out the full procedure for E_{rad} as described in the following subsections, but using a sum of the 1D contributions, we found a fit statistically at least competitive with that obtained from a product. However, in this case the sum ansatz tends to produce suspicious curvature in the $\widehat{S} = 1$, low-η region, which cannot be suppressed by the extreme-mass-ratio information. With additional NR data in this region, that problem might be alleviated, but with the current data set we find the product ansatz to be more robust and able to yield a final fit that is both accurate and well determined
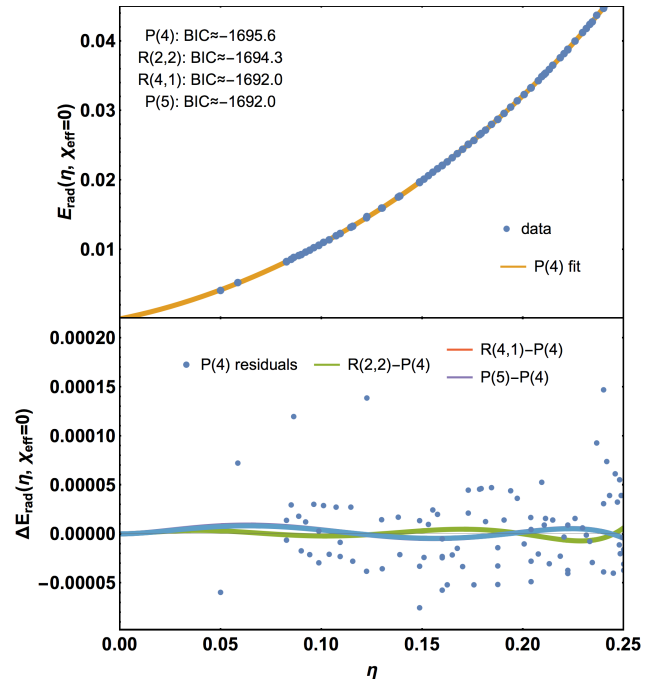

FIG. 15. Nonspinning NR data and one-dimensional $E_{rad}\left(\eta, \widehat{S} = 0\right)$ fit as a function of mass ratio η, Top panel: selected fit, a fourth-order polynomial P(4), see Eq. (21). Lower panel: residuals of this fit (points) and differences from the three next-best-ranking in terms of BIC (lines).

	Estimate	Standard error	Relative error [%]
a_2	0.5610	0.0026	0.5
a_3	−0.847	0.027	3.2
a_4	3.145	0.069	2.2

TABLE VII. Fit coefficients for the one-dimensional nonspinning $E_{rad}\left(\eta, \widehat{S} = 0\right)$ fit over 92 NR cases.

over the calibration region and without obvious bad behavior in extrapolation.

A. One-dimensional fits

For the nonspinning 1D fit in symmetric mass ratio η, a simple fourth-order polynomial

$$E_{rad}\left(\eta, \widehat{S} = 0\right) = a_4\eta^4 + a_3\eta^3 + a_2\eta^2 + \left(1 - \frac{2\sqrt{2}}{3}\right)\eta \quad (21)$$

with three free coefficients, listed in Table VII, is marginally preferred by both AICc and BIC. More complicated rational functions are not able to yield any significant change in residuals (only up to 1% in RMSE), while the differences between Eq. (21) and the next-ranked fits are again much smaller than the remaining residuals, as shown in Fig. 15.

For the effective-spin dependence, again the value at $\left(\eta = 0.25, \widehat{S} = 0\right)$ is fixed from the η fit. A rational function of order (3,1) is top-ranked by AICc, BIC and RMSE and thus unambiguously selected as the preferred ansatz:

$$E_{rad}\left(\eta = 0.25, \widehat{S}\right) = \frac{0.0484161\left(0.128b_3\widehat{S}^3 + 0.211b_2\widehat{S}^2 + 0.346b_1\widehat{S} + 1\right)}{1 - 0.212b_5\widehat{S}}$$
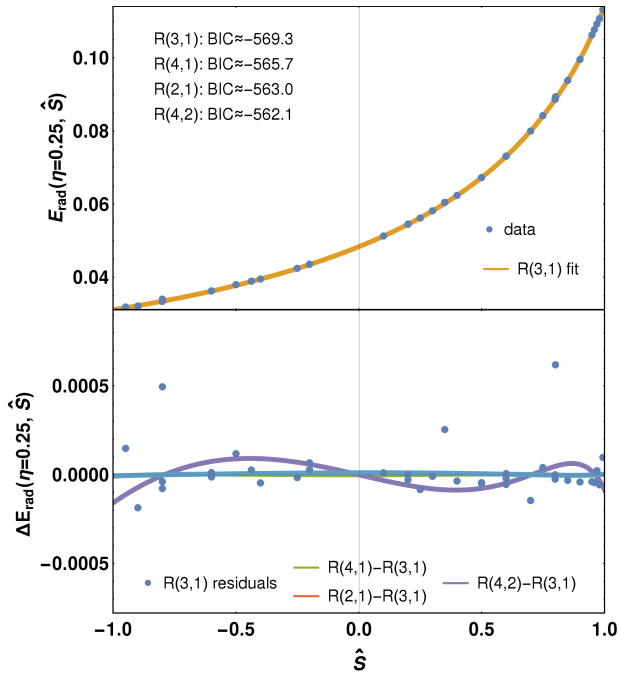
$$(22)$$

FIG. 16. Equal-mass-equal-spin NR data and one-dimensional fits of $E_{\rm rad}\left(\eta = 0.25, \widehat{S}\right)$ as a function of effective spin $\widehat{S}$. Top panel: selected fit, a rational function R(3,1), see Eq. (22). Lower panel: residuals of this fit (points) and differences from the three other top-ranking fits in terms of BIC (lines).

	Estimate	Standard error	Relative error [%]
b_1	-0.209	0.016	7.6
b_2	-0.197	0.026	13.2
b_3	-0.159	0.049	31.1
b_5	2.985	0.034	1.1

TABLE VIII. Fit coefficients for the one-dimensional equal-mass-equal-spin $E_{\rm rad}\left(\eta = 0.25, \widehat{S}\right)$ fit over 37 NR cases.

with four free coefficients listed in Table VIII, and well-behaved residuals as seen in Fig. 16.

B. Two-dimensional fits

For the 2D ansatz, we combine the two 1D fits from Eqs. (21) and (22), expanding each $\widehat{S}$-dependent term with a polynomial in η, according to Eq. (9), and removing the $\left(\eta = 0.25, \widehat{S} = 0\right)$ value from the spin ansatz before multiplying with the η terms:

$$E_{\rm rad}\left(\eta, \widehat{S}\right) = E_{\rm rad}\left(\eta, 0\right) \frac{E_{\rm rad}\left(0.25, \widehat{S}, f_{ij}\right)}{E_{\rm rad}\left(0.25, 0\right)} . \quad (23)$$

Contrary to the sum ansatz for $\chi_{\rm f}$ in Eq. (10), we do not need to set the η-independent coefficients f_{i0} of the $\widehat{S}$ terms to zero, as the $E_{\rm rad}\left(\eta, \widehat{S}\right) = E_{\rm rad}\left(\eta, 0\right)\left(1 + \ldots\right)$ form of Eq. (23) already guarantees the correct $\eta = 0$ limit. Hence an expansion up to third order in η of each $\widehat{S}$ term, as we chose for the $\chi_{\rm f}$ fit, would yield too many free coefficients, and instead we only expand up to second order. The four f_{i2} coefficients are again fixed by the equal-mass boundary conditions:

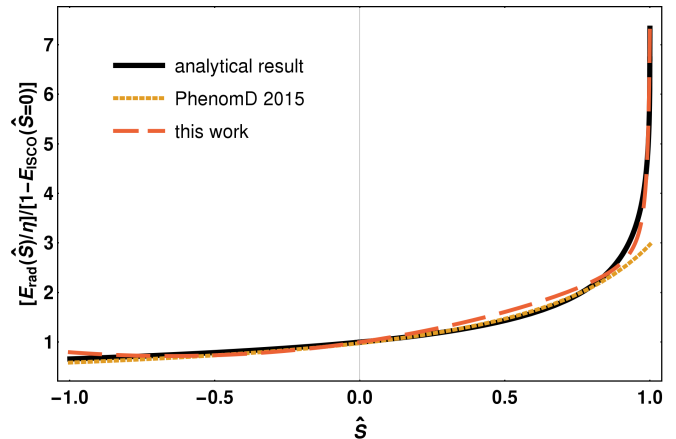$$f_{i2} = 16 - 16f_{i0} - 4f_{i1} . \quad (24)$$

FIG. 17. Radiated energy: extreme-mass-ratio comparison of analytical results, the previous PhenomD radiated-energy fit of [9], and this work. Note that, constrained to capture the steep rise at $\widehat{S} \to +1$, the new fit actually deviates slightly more from the analytical result at low positive $\widehat{S}$. This could be avoided with a more complex 1D $\widehat{S}$ ansatz, which is however disfavored by the current NR data set.

	Estimate	Standard error	Relative error [%]
f_{20}	4.27	0.38	8.9
f_{30}	31.09	0.71	2.3
f_{50}	1.56735	0.00032	0.02
f_{10}	1.81	0.15	8.2

TABLE IX. Fit coefficients for the extreme-mass-ratio limit of the radiated energy, fitted to discretized analytical results. The fourth coefficient, f_{10}, is fixed by the constraint at $\widehat{S} = 1$, cf. Eq. (26), and its estimate and error are computed from the others.
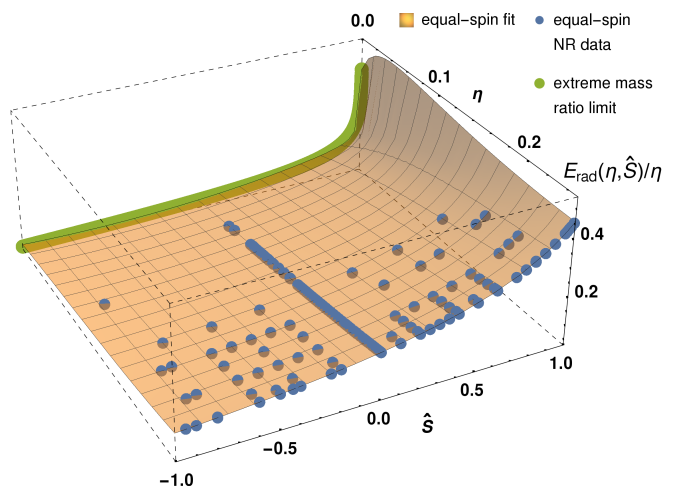

FIG. 18. Two-dimensional $E_{\rm rad}\left(\eta, \widehat{S}\right)$ fit, visualized as $E_{\rm rad}\left(\eta, \widehat{S}\right)/\eta$. Application of the extreme-mass-ratio limit helps in avoiding extrapolation artifacts which would otherwise appear at low-η, high-$|\widehat{S}|$ regions that are uncovered by NR simulations.

Similar to the procedure for $\chi_{\rm f}$, we can use the extreme-mass-ratio limit to fix the four coefficients f_{i0} of the linear-in-η terms. Using the analytic result from Eq. (2)a, we force the fit to satisfy the equality

$$E_{\rm rad}(\eta \to 0, \widehat{S}) = 1 - E_{\rm ISCO}(\widehat{S}) \quad (25)$$

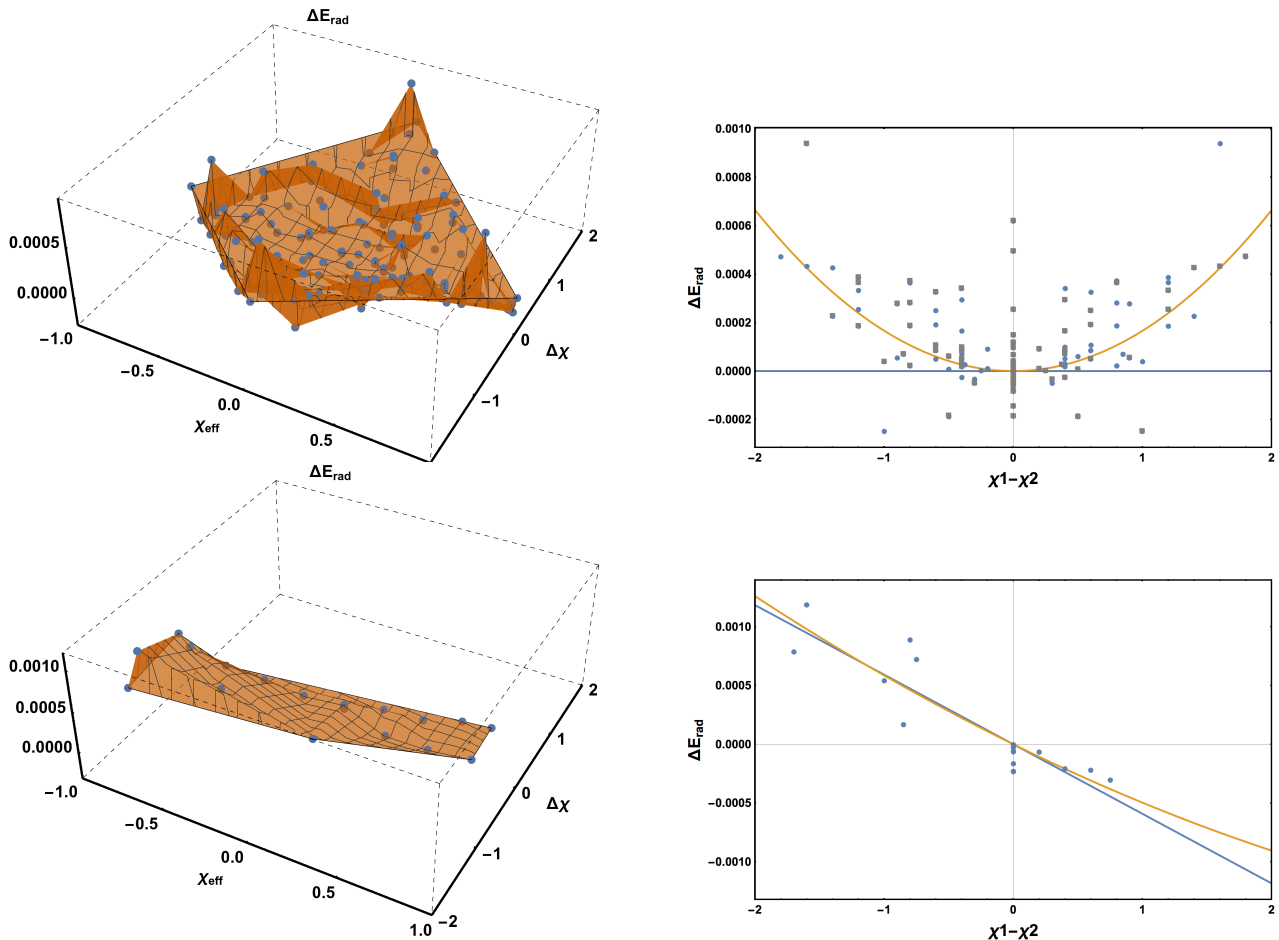and fit the corresponding leading-order η dependence of our

FIG. 19. Examples of spin-difference behavior of the radiated energy at fixed mass ratios, for residuals $\Delta E_{\rm rad}$ after subtracting the two-dimensional $E_{\rm rad}\left(\eta, \widehat{S}\right)$ fit. Top row: $q = 1$ (mirror-duplicated data points shown in gray); lower row: $q = 4$; left column: surfaces in $\left(\widehat{S}, \Delta\chi, \Delta E_{\rm rad}\right)$ space; right column: projections unto the $\Delta\chi$ axis with linear and quadratic fits. At equal mass, the linear term and mixture term vanish, but the expected quadratic dependence (parabolic surface) is less clearly pulled out from rather noisy residuals than for the final spin (cf. Fig. 9). At $q = 4$ and other intermediate mass ratios, the surface is not as close to flat as in the final-spin case, and the noisy data still shows some quadratic dependence.

2D ansatz to discretized values of this quantity. Again we fix one of the four free coefficients of $E_{\rm rad}\left(\eta \to 0, \widehat{S}\right)$ by a constraint fixing the value at $\widehat{S} = 1$, which is necessary to capture the very steep rise of Eq. (25) as $\widehat{S} \to +1$:

$$f_{10} \to -0.574752 f_{20} - 0.280958 f_{30} + 64.6408 f_{50} - 88.3165 . \tag{26}$$

The agreement between discretized analytical result and fit is shown in Fig. 17, and fit coefficients are listed in Table IX.

We thus have $12 - 4 - 4 = 4$ free coefficients f_{i1}, of which f_{21} turns out to be extremely poorly constrained, so that we set it to zero before refitting. Results of the 2D fit, calibrated to equal-spin simulations only, are shown in Fig. 18, which shows that the steep shape of the extreme-mass-ratio limit at high $\widehat{S}$ is smoothly attained by the extrapolated fit. For the curvature at low η and extremal $\widehat{S} = 1$, where there is no NR data, there might be also a contribution from the small remaining fit issues in the extreme-mass-ratio limit (cf. Fig. 17). The residuals again have larger RMSE than the 1D fits in η and $\widehat{S}$, by factors of 6.5 and 1.8 respectively, but show no clear apparent trends, allowing us to use this 2D fit as the basis for an unequal-spin residuals study in the next step.

C. Unequal-spin contributions and 3D fit

The spin-difference dependence of unequal-spin residuals is less clear here than for the final spin: As seen in the examples of Fig. 19, the general trend is the same with a quadratic dependence on $\Delta\chi$ at equal masses and more dominant linear effects as η decreases, but the distributions are generally noisier and the second-order terms (quadratic and mixture $\propto \widehat{S}\Delta\chi$) cannot be as cleanly separated.

For both the per-mass-ratio-step analysis and the direct 3D fit, we use the same general functional forms for possible linear, quadratic and mixture terms as in Eqs. (15), (17) and (18). After fixing ill-constrained coefficients to integer values, these reduce to

$$A_1(\eta) = d_{10}(1 - 4\eta)^{0.5}\eta^2 (d_{11}\eta + 1) \tag{27a}$$

$$A_2(\eta) = d_{20}\eta^3 \tag{27b}$$

$$A_3(\eta) = d_{30}(1 - 4\eta)^{0.5}\eta (d_{31}\eta + 1) . \tag{27c}$$

Figure 20 shows that the linear term is again robustly determined and does not change shape much when adding the two additional terms, but already the per-mass-ratio and direct-3D fits for this term do not agree quite as closely as in the $\chi_{\rm f}$ fit. The quadratic term is more noisy, and for the mixture term the
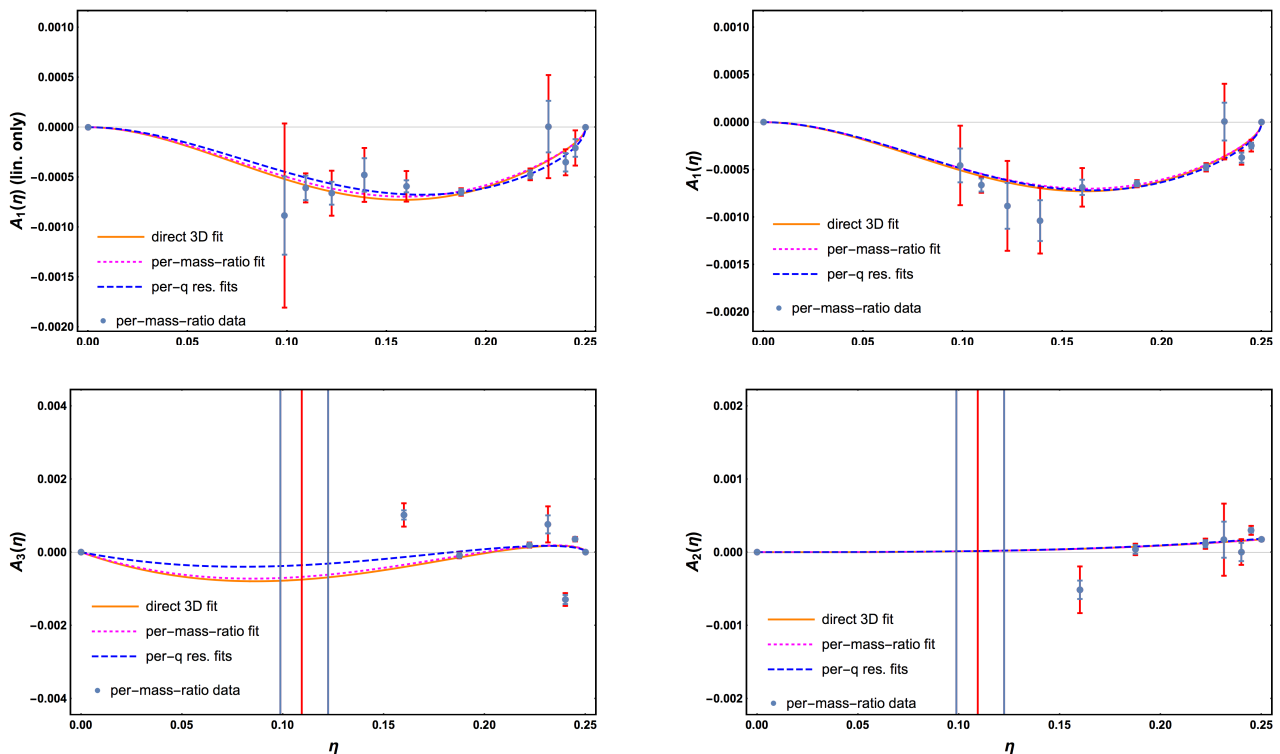
FIG. 20. Spin-difference behavior of radiated-energy data after subtraction of the two-dimensional $E_{\rm rad}\left(\eta,\widehat{S}\right)$ fit, for the three ansatz functions $A_i(\eta)$ from Eq. (27), with the same mass-ratio steps and fits as in Fig. 10. Top-left panel: linear term A_1 only. The remaining panels are for the combined linear+quadratic+mixture fit, in clockwise order: linear term A_1, quadratic term A_2 and mixture term A_3. The A_1 results from the combined fit are very similar to those from the linear-only fit, demonstrating the robustness of extracting leading-order spin-difference effects. For the two lower panels, results are much more uncertain, and the error bars for low η go far outside the displayed range, so that this region does not contribute significantly to the weighted per-mass-ratio fits.

results are rather uncertain, with an apparent sign change in the effect over η, but the stepwise cross-checks at least agreeing on the overall shape. Still, we will see below that inclusion of both these effects is statistically justified.

The full 3D ansatz for $E_{\rm rad}\left(\eta,\widehat{S},\Delta\chi\right)$ is then built up as

$$E_{\rm rad}\left(\eta,\widehat{S},\Delta\chi\right) = E_{\rm rad}\left(\eta,\widehat{S}\right) + \Delta E_{\rm rad}\left(\eta,\widehat{S},\Delta\chi\right),\qquad(28)$$

and this time has eight free coefficients (three from the 2D ansatz and five from the spin-difference terms). Results for the fit to 298 NR cases not previously used in the 1D fits are listed in Table X.

	Estimate	Standard error	Relative error [%]
d_{10}	-0.098	0.011	11.3
d_{11}	-3.23	0.18	5.6
d_{20}	0.0112	0.0012	10.5
d_{30}	-0.0198	0.0036	18.4
d_{31}	-4.92	0.19	3.9
f_{11}	15.7	1.2	7.9
f_{31}	-243.6	8.0	3.3
f_{51}	-0.58	0.13	21.6

TABLE X. $E_{\rm rad}$ fit coefficients for the final 3D step, using 298 cases.

D. Fit assessment

Tbl. XI gives a statistical summary of the various 1D, 2D and 3D fits for $E_{\rm rad}$ as discussed in this section. We find that the full linear+quadratic+mixture fit to all data has better RMSE than the simpler versions, and actually on the same level as the 2D equal-spin fit to equal-spin data only. The additional coefficients are also justified by yielding the best AICc and BIC, though not as clearly as for the final-spin case in Table V. Thus, we choose this three-term ansatz, matching the final-spin choice. Also, from Table X we see that all coefficients of this ansatz are sufficiently well constrained, with a worst standard error of 21.6%, to be useful for applications, at least under the assumptions made for the ansatz terms in Eq. (27). Just as was the case for the final-spin fit, and even more so because of the noisier data, we caution that ambiguity over the exact shape of the spin-difference terms remains, because different choices of the exponents and expansion orders in Eq. (27) can yield statistically indistinguishable fits.

As an additional check on the overall functional behavior of the fit, in Fig. 21 we show again comparisons between the intermediate 2D fit calibrated to equal-spin data only, a 2D fit to all data (not working well, as expected) and the 2D part of the final 3D fit. The equal-spin 2D fit and 3D fit agree well, with the strong excess curvature of the all-data 2D fit in the high-$\widehat{S}$ region reduced significantly thanks to the unequal-spin terms.

	N_{data}	N_{coeff}	RMSE	AICc	BIC
1D η	92	3	4.14×10^{-5}	-1705.7	-1695.6
1D $\hat{S}$	37	4	1.51×10^{-4}	-577.3	-569.3
2D ($\chi_1 = \chi_2$)	60	3	2.67×10^{-4}	-875.3	-867.4
2D all	298	3	4.26×10^{-4}	-4070.9	-4056.2
3D lin	298	5	3.24×10^{-4}	-4282.9	-4260.9
3D lin+quad	298	6	2.72×10^{-4}	-4391.9	-4366.3
3D lin+mix	298	7	2.91×10^{-4}	-4339.3	-4310.1
3D lin+quad+mix	298	8	2.62×10^{-4}	-4417.8	-4385.0

TABLE XI. Summary statistics for the various steps of the hierarchical radiated-energy fit. Evidence for spin-difference terms beyond linear order is weaker than for the final-spin fits in Table V.
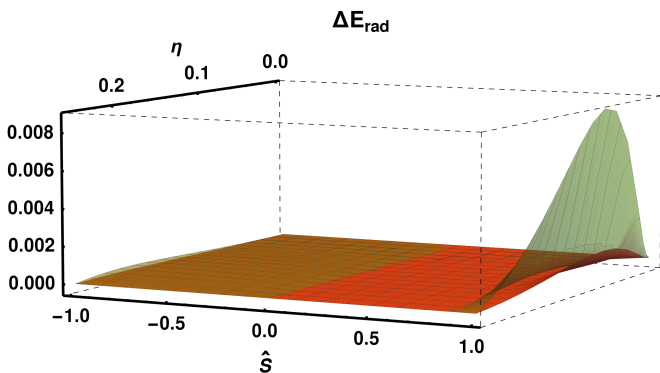

FIG. 21. Green: Difference ΔE_{rad} of a 2D fit to the full radiated energy data set minus the 2D fit to equal-spin cases only, both including extreme-mass-ratio constraints. Orange: Difference ΔE_{rad} of the 2D part of the 3D fit to the full data set minus the 2D-only fit to equal-spin data.
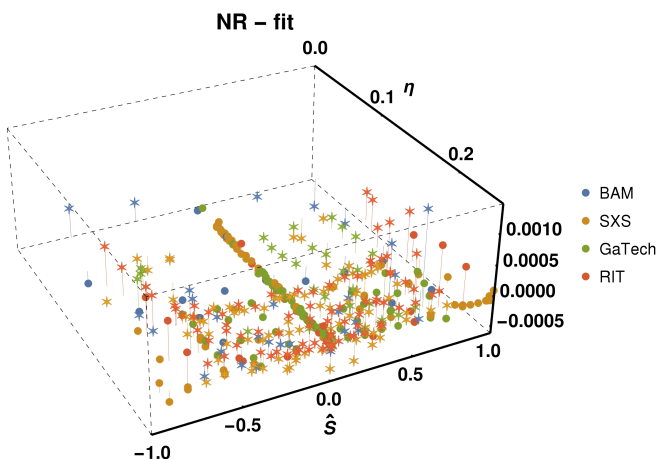

FIG. 22. Residuals of the radiated-energy fit, projected to the 2D parameter space of η and $\widehat{S}$. Stars: unequal-spin points.

The residuals of the new E_{rad} fit are shown over the $(\eta, \widehat{S})$ parameter space in Fig. 22 and as histograms in Fig. 23, compared with several previously published fits [9, 19, 52]. Comparison statistics are also summarized in Table XII. Again we find significant improvement, by a factor of 5 in residual standard deviation over the simple two-coefficient fit of [19] and about 40% improvement over the previous PhenomD fit of [9] and the RIT fits of [52, 61]. As the maximum possible emitted energy fraction – for equal-mass BBHs with extremal positive spins – the new fit predicts $E_{\text{rad}} \approx 0.1142$, consistent with the fit from [70] which was specifically calibrated to equal-mass-
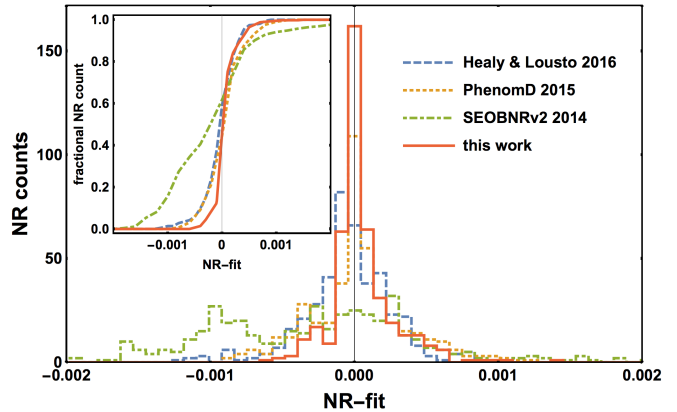

FIG. 23. Fit residuals of the radiated energy E_{rad}, for this work (cf. Table X) and for previously published fits (SEOBNRv2 2014 [19], Healy & Lousto 2016 [61], PhenomD 2015 [9]), evaluated over the set of 427 NR simulations shown in Fig. 2. Main panel: histograms, with 10 outliers for SEOBNRv2 (a recalibration of the fit from [51]) with $|\text{NR} - \text{fit}| > 0.002$ outside of the plot range. Inset: cumulative distributions over the same range.

	N_{coef}	mean	stdev	AICc	BIC
HLZ2014 [52]	19	-5.4×10^{-5}	3.4×10^{-4}	-5802.5	-5723.2
HL2016 [61]	19	-4.4×10^{-5}	3.0×10^{-4}	-5909.8	-5830.5
PhenomD	10	2.5×10^{-5}	3.4×10^{-4}	-5914.9	-5870.8
(refit)	10	6.1×10^{-5}	3.3×10^{-4}	-5947.7	-5899.6
SEOBNRv2 [19]	2	-1.7×10^{-4}	1.0×10^{-3}	-5036.1	-5023.9
This work	15	4.7×10^{-5}	2.2×10^{-4}	-6454.8	-6391.0
(refit)	15	6.3×10^{-5}	2.1×10^{-4}	-6482.8	-6419.0
(uniform	15	-4.0×10^{-6}	2.1×10^{-4}	-5987.3	-5923.5
(uniform refit)	15	1.4×10^{-6}	2.0×10^{-4}	-6034.2	-5970.4

TABLE XII. Summary statistics for the new radiated-energy fit compared with previously published fits [9, 19, 52, 61], evaluated over the full set of 427 NR simulations shown in Fig. 2. Also listed are a refit of the PhenomD [9] ansatz to the present NR data set, a refit of our hierarchically obtained ansatz directly to the full data set, and results with the same fitting procedure, but using uniform weights.

equal-spin cases and with [9, 52], but 15% higher than [51], underlining the importance of high-order $\widehat{S}$ terms at extremal spins.

V. CONCLUSIONS

We have developed a hierarchical step-by-step fitting method for results of numerical-relativity simulations of BBH mergers, assuming nonprecessing spins and negligible eccentricity, so that the parameter space is given by the mass ratio and two spin components: (η, χ_1, χ_2). We have then applied the method to the spin χ_{f} and mass M_{f} of the remnant Kerr BH, the latter being equivalent to radiated energy E_{rad}.

An appropriate fit is constructed in simple steps which reduce the dimensionality of the problem, with the ansatz choice at each step driven by inspecting the data. The full higher-dimensional fit is then built in a bottom-up fashion, modeling each parameter's contribution in order of its importance, as illustrated in the flowchart of Fig. 1.

A key goal of our approach is to avoid overfitting. To achieve this, and to evaluate the quality of the fits, we use in-

formation criteria (AICc and BIC) described in Appendix B. At least as important, however, is also the hierarchical data-driven nature of our procedure, e.g. when modeling the sub-dominant dependence on the difference between the spins. Through a reduction to one-dimensional problems, and inspecting the data at different mass ratios, an appropriate model with a small number of parameters is easily identified.

We compare our results with previous fits in the literature, also refitting these previous models to our data set, and we find a clear preference for the new fits in terms of residuals and information criteria.

Our emphasis on inspecting the underlying data set, and comparing its quality with the statistical analysis of fit errors, highlights that for further improvement of these numerical fits, it will be essential to understand the uncertainties and systematic effects in a data set, and to cleanly combine data from different codes, rather than simply increasing the number of calibration simulations without controlling the error budget.

In addition to the quality of the fit as applied to the available numerical relativity data, we investigate extrapolation properties beyond the calibration region, in particular for extreme spins, where we check that our fit does not overshoot the extreme Kerr limit, and avoids pathological oscillations. The extreme-mass-ratio limit has previously been incorporated into fits for final mass and spin in different ways, in particular for the final spin, where the influence of radiated energy is not always accounted for (e.g. in [47, 52]). Doing so is however important to avoid oscillatory features due to an unphysical value of the derivative with respect to the symmetric mass ratio η at $\eta = 0$, and indeed we achieve a more robust behavior for close-to-extremal final spins in comparison to other recent fits [52, 55] – see Fig. 14. In order to avoid numerical errors in this derivative, we use an analytical expansion around the $(\eta = 0, \widehat{S} = 1)$ corner.

We have also verified that our fits for final mass and spin are consistent with the Hawking area theorem for black holes [71]: the area of the final horizon is larger than the sum of the individual horizons of the initial black holes. The difference in areas is much larger than the fit errors, thus the area theorem does not provide an additional constraint on the fits.

For both $\chi_{\rm f}$ and $E_{\rm rad}$, we can robustly identify the main unequal-spin contribution of the form $f(\eta)(\chi_1 - \chi_2)$. The shape of the function $f(\eta)$ is well determined for final spins, but has larger errors for the smaller unequal-spin contribution to radiated energy. We also discuss and find statistical evidence for two further unequal-spin terms, one quadratic in $\Delta\chi$, and a mixed $\widehat{S}\,\Delta\chi$ term, which however are comparable in size to errors in the numerical data, so that their η-dependent shape cannot be tightly constrained with the current data set.

The same hierarchical fitting procedure can be easily applied to other quantities such as peak luminosity [72, 73]. An important goal of our work is the accurate calibration of unequal-spin effects for full inspiral-merger-ringdown waveform models, in particular toward extending the PhenomD model [9, 10]. This will allow investigating under which conditions gravitational-wave observations can reveal the individual spins of a coalescing binary. While no fundamental modifications are necessary to the method itself, a careful analysis of the data sets will be required to make appropriate choices in combining the 1D subspace fits, including extreme-mass-ratio information, and adding spin-difference terms. A more ambitious goal for the future is the extension of our hierarchical fit construction to the higher-dimensional problems of generic precessing BBHs.

Appendix A: Data sets and NR uncertainties

For NR calibration of BBH coalescence models, it is useful to combine data sets from different research groups and numerical codes, both to increase robustness against code inaccuracies and errors in the preparation of data products (such as incorrect metadata), and to benefit from the combined computational resources of different groups. Computational cost increases significantly with mass ratio and spin, thus the high-mass-ratio and high-spin regions are still poorly sampled, and numerical errors are often higher (as we will see below). Very different spins on the two black holes will typically also increase computational cost due to the need to resolve different scales in the grids around the two black holes (higher spin leads to smaller black holes for typical coordinate gauges in numerical relativity), which is a challenge for accurately capturing small subdominant effects like the nonlinear spin-difference effects we discuss in this work.

In this Appendix, we discuss our procedures to eliminate data points of poor quality, to assign fit weights, and to check consistency between assumed error bars and our fit results. We expect two main avenues to significantly improve over the fits we have presented in this paper: (a) providing more data points with high spins and unequal masses, in order to improve the accuracy of the fit near the boundaries of the fitting region and to reduce the need for extrapolation; and (b) determining more accurate and robust error bars for NR data, which would allow one to isolate small subdominant effects.

Final spin and final mass are usually computed as surface integrals over the apparent horizon using the isolated-horizon formalism [74] (see [75] for a summary of methods and references for different codes), from surface integrals over spheres
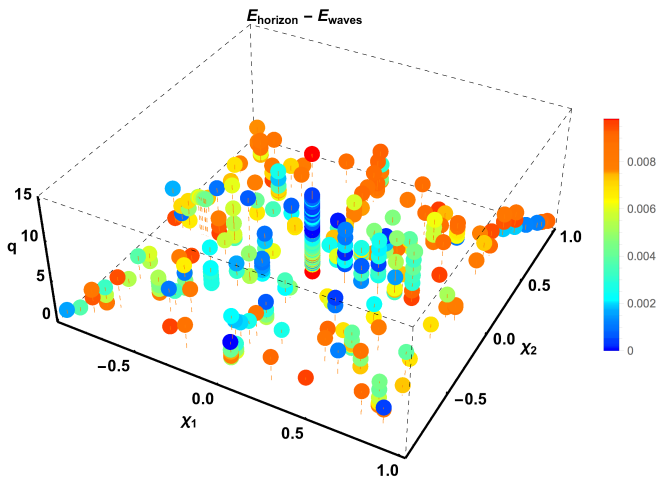
FIG. 24. Differences between radiated energy computed from either horizon or waveform data, across the parameter space. The color scale quantifies the differences between the two computations. Differences are largest for high-mass-ratio and high-spin cases, where high NR accuracy is more demanding.

at large or infinite radius (as in [64]), or from the energy or angular momentum balance computed from initial and radiated quantities (see Eq. (A1) below). Final spin and final mass can also be obtained from fits to the ringdown [76].

Final mass and spin from the apparent horizon (AH) are generally expected to be more accurate than those based on the evaluation of asymptotic quantities (such as Bondi mass, angular momentum, or radiated energy and angular momentum) at finite radius, where errors may arise due to finite radius truncation, insufficient extrapolation to infinity or numerical inaccuracies in propagating the wave content to large distances at sufficient numerical resolution. On the other hand, AH quantities may suffer from inaccuracies in finding the apparent horizon and from gauge ambiguities.

For SXS, RIT and GaTech results, we take the values provided in the catalogs [52, 59, 60, 62, 63], which in general have been computed at the horizon. For a comparison of horizon vs radiated quantities for the RIT results, see Table V in [52], where error bars for the horizon quantities are significantly smaller. For the BAM code, for some large-mass-ratio cases the AH finder fails due to the unfortunate choice of a shift condition, which results in a coordinate growth in the horizon which is roughly linear in time during the evolution. After several orbits the horizon of the larger BH is then no longer contained within the fine grid of the mesh refinement, which may trigger a failure of the horizon finder code. Due to the high computational cost of the simulations, we have not rerun these cases with improved parameters for the apparent horizon finder code. But rather, we compute the final angular momentum from the angular momentum surface integral at large radius, and the energy from the radiated GWs. For processing GaTech waveforms we have also followed [77].

For all 414 cases where we have both the waveform and AH estimate available, we perform cross-checks between AH quantities and those obtained from integrating radiated energy and angular momentum. In order to compute the final mass from the radiation quantities, we integrate

$$\frac{dE}{dt} = \lim_{r \to \infty} \left[\frac{r^2}{16\pi} \int_\Omega d\Omega \left| \int_{-\infty}^t \psi_4 d\tilde{t} \right|^2 \right] \quad \text{(A1)}$$
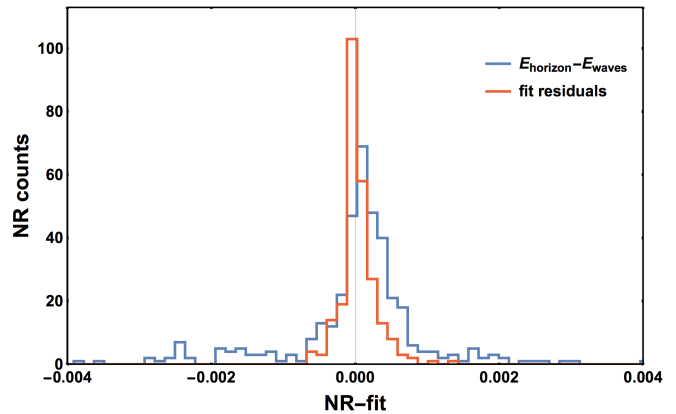

FIG. 25. Histograms of the differences between radiated energy computed from either horizon or waveform data (as in Fig. 24), compared with the residuals of the new radiated-energy fit (as in Fig. 23).

over time, starting after the initial burst of "junk radiation".

We extrapolate the result from different extraction radii to infinity at linear order in inverse radius. To account for energy radiated at separations larger than the initial separation of the NR simulation, we compute the radiated energy at 3.5 PN order [78–84] from $\omega \in [0, \omega_0]$, with ω_0 being the initial orbital frequency of the NR simulation (after junk radiation). Then we can consider the differences between radiated energy values from the horizon and from the integrated waveforms, shown in Fig. 24, as an estimate of NR errors. However, this will typically be a *pessimistic* estimate because horizon quantities are in general more reliable and thus big differences are typically caused by inaccuracies in the integrated emission. In Fig. 25 we show that the distribution of this pessimistic estimate is similar to, but much wider-tailed than, the residuals from our radiated-energy fit.

A more realistic measure of NR errors is the difference between results from different codes for equal initial parameters. With a strict tolerance requiring equal initial parameters to within numerical accuracy,

$$|\lambda_i - \lambda_j| \le \epsilon = 0.0002 \text{ with } \lambda_i = \{\eta_i, \chi_{1i}, \chi_{2i}\}, \quad \text{(A2)}$$

we find 41 such duplicate configurations out of the total of 427 cases.[1] We evaluate differences between these equal-parameter cases for final spin and radiated energy. Figure 26 shows that, with strict tolerance, these error estimates (standard deviations of 3.1×10^{-4} for χ_f and 1.6×10^{-4} for E_{rad})[2] are still on the same order but smaller than the respective fit residuals (RMSE of 5.2×10^{-4} for χ_f and 2.2×10^{-4} for E_{rad}). However, the set of true duplicates is small and mostly concentrated in equal-spin-similar-mass regions of the parameter space (cf. Fig. 27), preventing us from naively extrapolating this error estimate to the full parameter space. Hence we consider it as a somewhat *optimistic* estimate of final-state NR errors.

---

[1] With a more relaxed tolerance, $\epsilon = 0.001$ in Eq. (A2), we find 33 duplicates and 19 sets of two or more configurations with reasonably similar parameters, corresponding to a total of 131 cases (30% of the total data set), compatible with the 71 "twins" out of a data set of 248 reported in [55]. The wider-tolerance tuples are shown as open circles in Fig. 26.

[2] For the relaxed tolerance, the values are 2.8×10^{-3} for χ_f and 3.5×10^{-4} for E_{rad}, compatible with the 2×10^{-3} given in [55] for χ_f.

id	q	χ_1	χ_2	ω_0	D_0	E_{rad}	ΔE_{rad}	χ_f	$\Delta\chi_f$	tag	code
1	1.00	-0.80	-0.80	0.060	5.88	0.0325	-0.0010	0.4122	-0.0146	D6.2_q1_a-0.8_m100	GaTech
2	1.00	-0.60	-0.60	0.058	5.93	0.0349	-0.0013	0.4876	-0.0066	D6.2_q1_a-0.6_m100	GaTech
3	1.00	0.80	-0.80	0.025	10.92	0.0491	0.0002	0.6839	-0.0000	D11_a0.8_q1.00_m103_As	GaTech
4	1.00	0.80	0.80	0.024	11.07	0.0883	-0.0005	0.9086	0.0010	D11_q1.00_a0.8_m200	GaTech
5	2.50	0.60	0.60	0.051	6.27	0.0528	0.0002	0.8255	0.0004	Lq_D6.2_q2.50_a0.6_th000_m140	GaTech
6	3.50	0.00	0.00	0.015	15.90	0.0258	0.0007	0.5046	0.0005	BBH_CFMS_d15.9_q3.50_sA_0_0_0_sB_0_0_0	SXS
7	5.00	-0.73	0.00	0.030	9.53	0.0129	0.0004	0.0222	0.0460	D10_q5.00_a-0.73_0.00_m240	GaTech
8	5.00	-0.72	0.00	0.030	9.54	0.0129	0.0004	0.0164	0.0340	D10_q5.00_a-0.72_0.00_m240	GaTech
9	5.00	-0.71	0.00	0.029	9.55	0.0130	0.0005	0.0105	0.0220	D10_q5.00_a-0.71_0.00_m240	GaTech
10	5.00	0.00	0.00	0.027	10.07	0.0176	-0.0001	0.4175	0.0009	D10_q5.00_a0.0_0.0_m240	GaTech
11	5.50	0.00	0.00	0.031	9.16	0.0161	0.0001	0.3932	-0.0002	D9_q5.5_a0.0_Q20	GaTech
12	6.00	0.00	0.00	0.027	10.13	0.0145	-0.0001	0.3732	0.0007	D10_q6.00_a0.00_0.00_m280	GaTech
13	6.00	0.40	0.00	0.026	10.35	0.0195	0.0000	0.6257	-0.0000	D10_q6.00_a0.40_0.00_m280	GaTech
14	8.00	0.85	0.85	0.048	6.50	0.0248	-0.0027	0.8948	-0.0012	q8++0.85_T_80_200_-4pc	BAM
15	10.00	0.00	0.00	0.035	8.39	0.0082	-0.0000	0.2588	-0.0019	D8.4_q10.00_a0.0_m400	GaTech
16	10.00	0.00	0.00	0.035	8.39	0.0081	-0.0001	0.2665	0.0058	q10c25e_T_112_448	BAM

TABLE XIII. NR cases from the source catalogs not included in the fit calibration, for reasons detailed below.
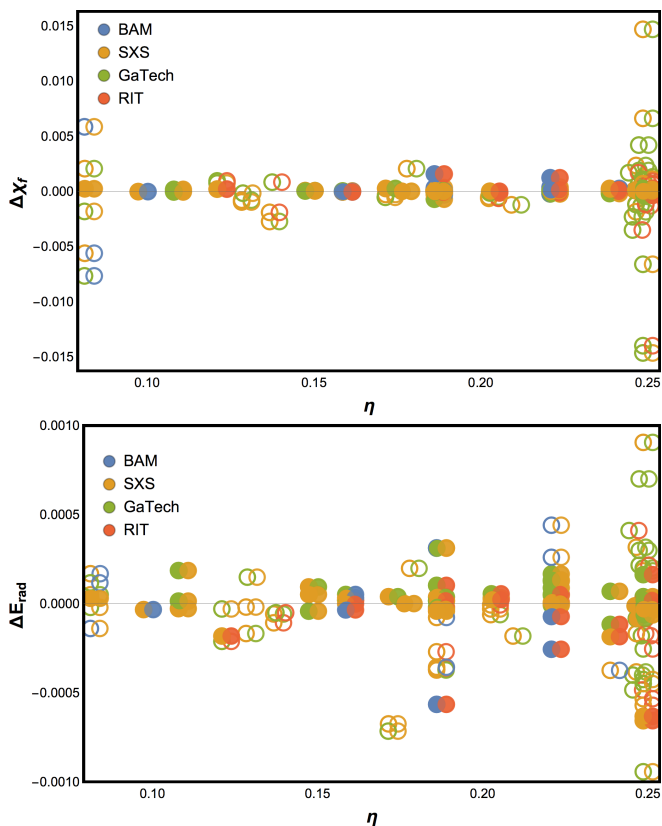

FIG. 26. Differences in final spin χ_f (top panel) and radiated energy E_{rad} (lower panel) for equal-parameter configurations but different NR codes. Solid circles: configurations with parameters equal to within numerical accuracy (narrow tolerance). Open circles: similar configurations but with some deviation in the parameters (wider tolerance, e.g. up to $\Delta\eta \approx 0.001$). Pairs of simulations are shown with a small horizontal offset for ease of visual identification.

FIG. 27. Differences in the final-state quantities for equal-parameter configurations and different NR codes. Top panel: final spin χ_f, lower panel: radiated energy E_{rad}. Points here correspond to both open and solid circles from Fig. 26 (wider tolerance).

We therefore have a rough expectation for the range of possible NR errors bracketed by these pessimistic and optimistic estimates, but no detailed information for each case over the whole parameter space. Instead, we use simple heuristic fit weights. The overall scal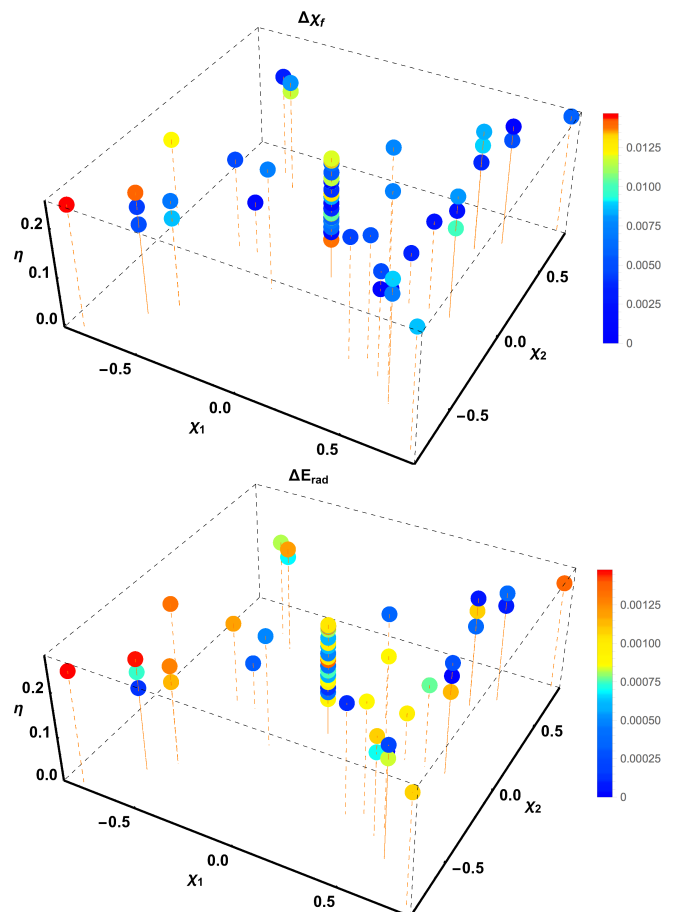e of the NR error is not relevant for determining fit weights, so we only need to assign relative weights between the cases, emulating the usual quadratic scaling with data errors which can also be deduced from Fig. 24. For SXS data we down-weight cases with $\eta < 0.1$ by a factor of 2^2; while for the puncture codes (BAM, GaTech, RIT) we expect larger inaccuracies especially at low η, and so we
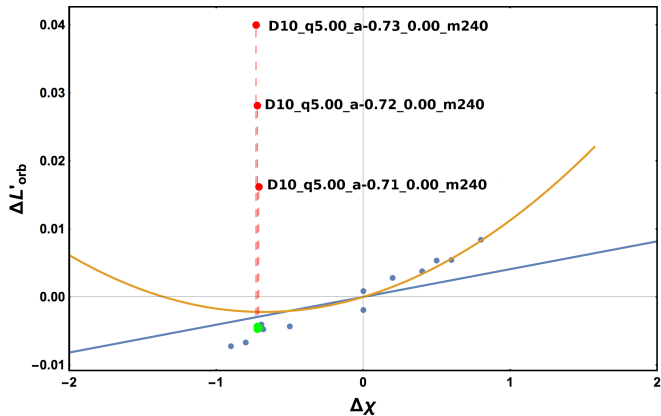
FIG. 28. Unequal-spin effects for final spin χ_f at $q = 5$, shown as residuals against the 2D equal-spin fit (cf. Fig. 9). The three points highlighted in red are similar configurations from the GaTech catalog, for which it has since been confirmed [85] that the sign of χ_f should be negative instead, making L'_{orb} fit with the trend – corrected values are shown in green.

down-weight by a factor of 2^2 above $\eta = 0.223$ and 3^2 below that mass ratio, and a factor of 5^2 below $\eta = 0.05$ (including the computationally challenging $q = 18$ cases). As mentioned before, a more detailed NR error study, leading to better-determined weights, would be a clear avenue to further improve fit results.

From the original set of NR simulations we have removed 16 cases as outliers, which are listed in Table XIII. For this decision, we have considered three main sources of outliers: cases whose NR setup is not appropriate for the purpose of this study, duplicated configurations for which the variations in the final quantities are much larger than the RMSE, and cases that are found to be drastically off the trend of otherwise smooth data sets in any of the one-dimensional plots in our hierarchical fitting procedure. Outliers 1,2,5,16 have rather short orbital evolutions, so that they can be used for ringdown-only studies, but not for our purpose of predicting the final state from initial parameters. For outliers 10–12 and 15–16 we have found large variations in the final-state values for different codes (see Fig. 26). Here we have used only the equivalent SXS configuration, in each case corresponding to longer and presumably more accurate evolutions. The remaining seven outliers have been identified after performing the step-by-step one-dimensional analysis of the data, each deviating so clearly that there must be an underlying systematic problem and not just a statistical fluctuation (in which case they could not be excised from the data set). As an example, we highlight in Fig. 28 three clear GaTech outliers found in the unequal-spin calibration step; however, it was recently confirmed [85] that these three cases should have a negative sign of their final spin, and with this change they are fully consistent with our fits. We note that the overall data quality of the omitted cases may be perfectly adequate for other studies; while for this final-state study, due to good data coverage in the corresponding parameter-space regions and clear global trends in the full data set, the consistency requirements are quite narrow.

We also summarize in Table XIV 27 new `BAM` simulations first used in this paper, including recalculated values for previously published [9] mass ratio $q = 18$ simulations.

q	χ_1	χ_2	ω_0	D_0	$e \, [\times 10^{-3}]$	E_{rad}	χ_f
1.00	0.00	-0.85	0.022	11.97	2.25	0.0392	0.5514
1.00	0.85	-0.85	0.023	11.61	2.61	0.0491	0.6854
1.00	0.50	-0.50	0.023	11.58	1.59	0.0485	0.6858
1.20	0.00	-0.85	0.020	12.79	0.74	0.0401	0.5747
1.20	0.50	-0.50	0.028	10.00	1.76	0.0503	0.7142
1.20	0.85	-0.85	0.028	10.00	5.16	0.0527	0.7359
1.50	-0.50	0.50	0.024	11.00	1.80	0.0408	0.5865
1.75	0.00	0.85	0.022	11.69	2.35	0.0484	0.7033
1.75	0.00	-0.85	0.021	12.18	1.93	0.0369	0.5810
1.75	0.85	-0.85	0.023	11.59	4.95	0.0567	0.8167
1.75	-0.85	0.85	0.021	12.35	2.66	0.0343	0.4607
1.75	0.85	0.00	0.021	12.35	1.00	0.0682	0.8724
1.75	-0.85	0.00	0.020	12.69	0.58	0.0307	0.3934
2.00	0.50	-0.50	0.024	11.10	1.76	0.0464	0.7510
2.00	0.00	-0.85	0.023	11.47	2.85	0.0347	0.5693
2.00	0.00	0.85	0.024	11.16	2.52	0.0436	0.6732
2.00	0.85	-0.85	0.024	11.00	1.78	0.0556	0.8344
2.00	-0.85	0.85	0.022	11.73	3.07	0.0310	0.4002
2.00	-0.50	0.50	0.023	11.53	2.60	0.0336	0.4925
2.00	-0.85	0.00	0.022	11.97	2.70	0.0292	0.3425
2.00	0.85	0.00	0.023	11.36	4.02	0.0646	0.8782
3.00	-0.50	0.50	0.024	11.26	1.69	0.0237	0.3339
3.00	0.50	-0.50	0.025	10.63	1.41	0.0373	0.7410
3.00	-0.85	0.00	0.023	11.74	3.25	0.0201	0.1562
4.00	0.00	0.85	0.026	10.52	14.79	0.0230	0.4900
4.00	-0.85	0.85	0.023	11.37	5.04	0.0158	0.0323
4.00	-0.50	0.50	0.024	10.99	1.68	0.0177	0.2152
18.00	0.80	0.00	0.042	9.00	5.00	0.0087	0.8308
18.00	-0.80	0.00	0.027	10.00	2.50	0.0031	-0.5270
18.00	-0.40	0.00	0.031	9.00	1.30	0.0035	-0.1813
18.00	0.00	0.00	0.040	7.58	1.30	0.0041	0.1623
18.00	0.40	0.00	0.040	7.43	5.00	0.0057	0.5030

TABLE XIV. New `BAM` simulations used in this work, with a focus on high unequal spins; and recalculated values for $q = 18$ cases from [9]. For each simulation, we list mass ratio $q = m_1/m_2$, initial spins χ_1 and χ_2, reference orbital frequency Ω_0, initial separation D_0 (after junk radiation), eccentricity e, radiated energy E_{rad} (scaled to unit initial mass) and dimensionless final spin χ_f.

Appendix B: Fit assessment and model selection criteria

Both while constructing the full ansatz in our hierarchical process, and when selecting the final fit, we rank fits by several standard statistical quantities, which are briefly summarized here for the benefit of the reader.

A basic figure of merit is the root-mean-square-error,

$$\text{RMSE} = \sqrt{\frac{1}{N_{\text{data}}} \sum_{n=1}^{N_{\text{data}}} \left[X_{\text{NR}}(\eta_n, \chi_{1,n}, \chi_{2,n}) - \text{fit}(\eta_n, \chi_{1,n}, \chi_{2,n}) \right]^2},$$

(B1)

which just checks the overall goodness of fit. One caveat here is that down-weighted NR cases are fully counted in the RMSE, so that a generalized variance estimator using weights can be more useful.

Furthermore, it is important in model selection to penalize models with too many free coefficients, as in principle the RMSE can be made arbitrarily small when the number of coefficients approaches the number of data points. A popular

figure of merit for model selection considering the number of coefficients is the Akaike information criterion [45],

$$\text{AIC} = -2\ln\mathcal{L}_{\max} + 2N_{\text{coeffs}}, \qquad (B2)$$

which intuitively can be understood as weighing up goodness of fit (measured by the maximum log-likelihood $\mathcal{L}_{\max}$) against parsimony. Standard implementations, as the one from Wolfram Mathematica, assume Gaussian likelihoods.

A generalization that corrects the AIC for low numbers of observations and reproduces it for large data sets is the AICc:

$$\text{AICc} = \text{AIC} + \frac{2N_{\text{coeffs}}(N_{\text{coeffs}} + 1)}{N_{\text{data}} - N_{\text{coeffs}} - 1}. \qquad (B3)$$

In this work, we always use AICc instead of AIC.

A related quantity, similar in form but with a completely different theoretical justification and with subtle differences in practice, is the Bayesian information criterion or Schwarz information criterion [46]:

$$\text{BIC} = -2\ln\mathcal{L}_{\max} + N_{\text{coeffs}}\ln(N_{\text{data}}). \qquad (B4)$$

Though based on an approximation to full Bayesian model selection (while the AIC is derived from information theory), the BIC in general cannot be interpreted as a direct measure of Bayesian evidence between models.

There is much literature on advantages and disadvantages of these two criteria, and several other alternatives exist – see e.g. [86] and references therein. In practice, the BIC tends to impose a slightly stronger penalty on extra parameters than the AIC(c). Both criteria have been criticized [86] for not only penalizing completely extraneous parameters, but also the existence of degeneracies between parameters. However, this is a virtue rather than a problem for our purpose of selecting parsimonious model ansätze.

For all of AIC, AICc and BIC, the model with the *lowest* value is preferred. Higher than unit differences between two models are generally required to count as significant evidence; [86] quotes ±5 as "strong" and ±10 as "decisive" evidence.

By default, we rank the one-dimensional η and $\widehat{S}$ ansätze by BIC, and apply the same criterion to judge how many $\Delta\chi$ terms to include in the final 3D model. In general these are not guaranteed to be the best by AICc or RMSE as well, but in practice we check that the three criteria give a very similar ranking of models.

Still, we find that sometimes the best fit by any of these three criteria (RMSE, AICc, BIC) can suffer from one or several parameters being not well constrained. In parameter estimation over large measured data sets, such as the CMB example in [86], this is a desirable feature of model selection, as it allows to assess which physics is actually constrained by the data. However, in our case we are well aware that our data set does not constrain all possible functions over the parameter space, and we are more interested in reporting a well-constrained model that captures the information in the data set than to dig out weak constraints on possible extensions of that model. So we augment our model selection criteria by considering also the well-constrainedness of each individual fit coefficient, and allow for picking a fit with slightly worse summary statistics if it has better-constrained coefficients; or we drop individual coefficients from a high-order ansatz and reassess the quantitative criteria for that reduced model.
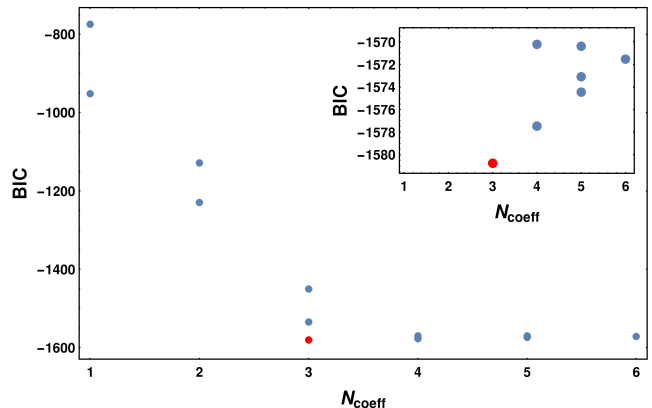

FIG. 29. BIC for the one-dimensional $L'_{\text{orb}}\left(\eta, \widehat{S} = 0\right)$ fits from Sec. III B. The inset panel is a zoom-up of the top-ranked fits. The tested set of ansätze includes all polynomials from second to seventh order in η and all rational functions of order (i, j), $j \leq i$, up to $i + j = 6$. The preferred ansatz, a rational function of order $(3,1)$ with three free coefficients, is highlighted.

As an example of how e.g. the BIC can guide model selection, we show in Fig. 29 the BIC ranking for the one-dimensional $L'_{\text{orb}}\left(\eta, \widehat{S} = 0\right)$ fits from Sec. III B. A plateau of almost constant BIC is made up of several fits with $N_{\text{coeffs}} \geq 3$, with the more complex fits yielding no additional improvement, so that we choose the simplest fit among this group. Still, even if it had not come up actually top-ranked, as in this case, choosing a low-N_{coeffs} fit from within the high-ranked group would be preferable over some slightly higher-ranked, but less-well-constrained fit.

Appendix C: Spin parameter selection

The results of the main text are given in terms of the spin parameter $\widehat{S}$. However, there is no unique definition of an "effective spin", and alternative parametrizations have been used in the literature [9, 55]. We have tested the robustness of our hierarchical approach for two additional spin parameters:

$$\widehat{S} = \frac{S}{m_1^2 + m_2^2}, \qquad S = m_1^2\chi_1 + m_2^2\chi_2, \qquad \chi_{\text{eff}} = m_1\chi_1 + \chi_2 m_2.$$

We redid the hierarchical ansatz construction and fitting for S and χ_{eff}, making the same ansatz choices for χ_{eff} as we did for $\widehat{S}$ in the main text, but changing the 1D spin ansatz to a polynomial P(7) for S (instead of R(3,1) for $\widehat{S}$ and χ_{eff}) because rational functions in S tend to yield singularities. Checking other possible choices, we have not found any ansatz combination that makes these alternatives match or exceed the performance of the $\widehat{S}$-based fits presented in the main part of this paper. Results in terms of the RMSE, AICc and BIC are listed in Table XV, and residual histograms shown in Fig. 30. We still obtain better results than most previous fits (see Tables VI and XII) with any parametrization, thus demonstrating the robustness of our method.

Again, we have also analyzed the fit in the extrapolation regions to detect any artifacts not reflected by the statistical criteria (which are meaningful only in the calibrated region). In Fig. 31 we check the extrapolation behavior of fits with the alternative parametrizations in the notoriously difficult $\chi_1 = \chi_2 = 1$ limit. The approach to this limit is smoother

	RMSE	AICc	BIC
$\hat{S}$	5.15×10^{-4}	-5991.5	-5923.9
S	5.24×10^{-4}	-5930.9	-5863.3
χ_{eff}	5.97×10^{-4}	-5799.6	-5731.9

	RMSE	AICc	BIC
$\hat{S}$	2.24×10^{-4}	-6454.8	-6391.0
S	6.45×10^{-4}	-5526.1	-5439.1
χ_{eff}	4.23×10^{-4}	-5962.7	-5898.9

TABLE XV. Summary statistics for fits with three different choices of effective spin parameter and ansatz choices as discussed below, evaluated over the full 427 point NR data set. Top table: Final spin, lower table: radiated energy.
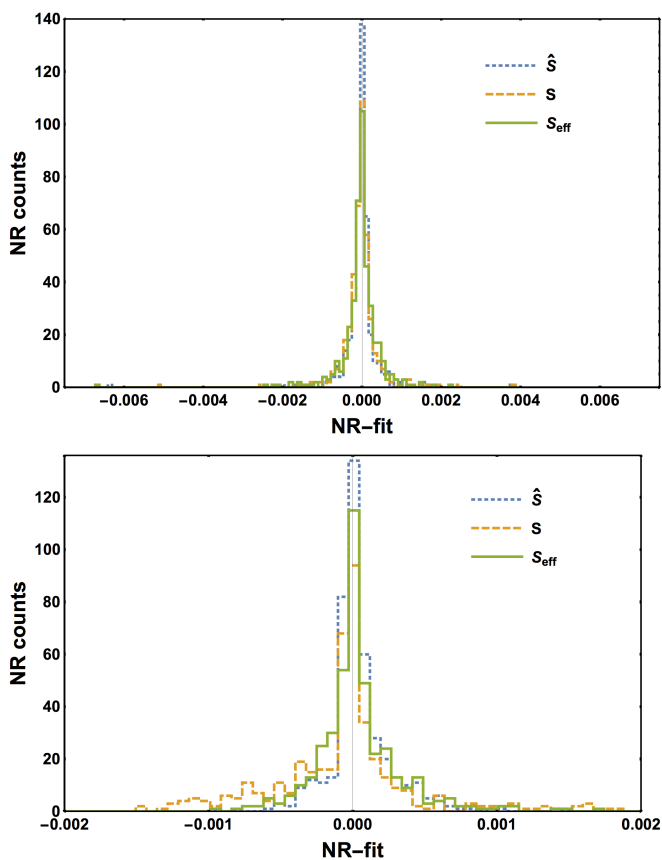

FIG. 30. Fit residuals for three different choices of effective spin parameter. Top panel: final spin; lower panel: radiated energy.

for the fits using $\widehat{S}$ and χ_{eff} than for that using S, which shows some certainly nonphysical oscillations.

The conclusion is that the hierarchical fitting method is quite robust under a change of effective-spin parametrization, and indeed we would expect full equivalence in the limit of a huge data set with small, completely known NR errors (using appropriately adapted ansätze for each parametrization). With the current data set, $\widehat{S}$ and χ_{eff} perform similarly, while when using S additional high-spin data would be even more important to ensure smooth extrapolation.
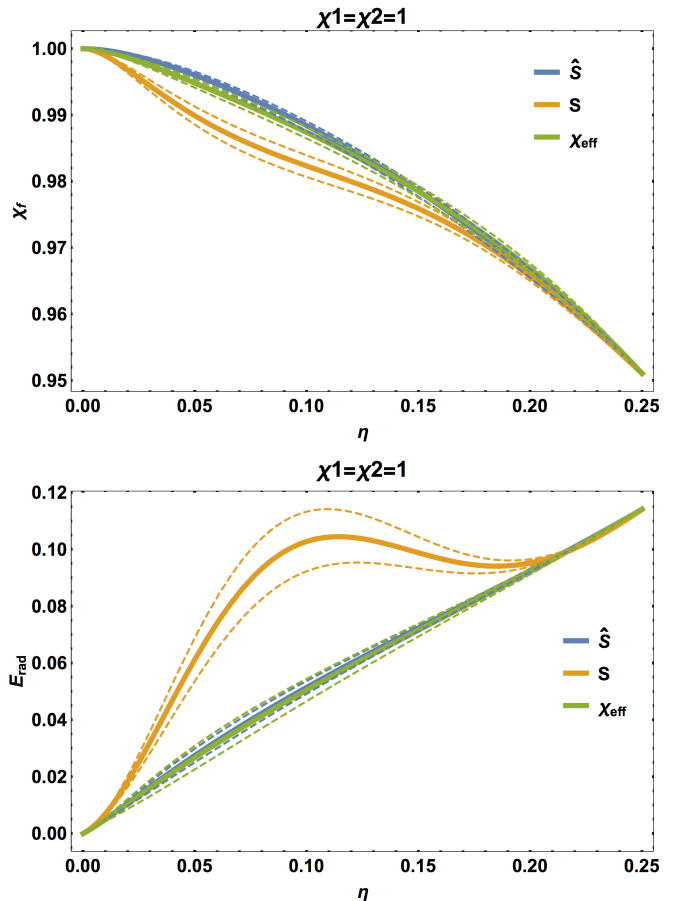

FIG. 31. Final-state quantities in the extremal $\chi_1 = \chi_2 = 1$ limit for three different choices of effective spin parameter. Top panel: final spin, lower panel: radiated energy.

Appendix D: Fit uncertainties

The uncertainty of evaluating a fitted quantity Q at a point (η, χ_1, χ_2) can be expressed through *prediction intervals* [87]

$$Q(\eta, \chi_1, \chi_2) \pm q_t(x, N_{\mathrm{data}} - N_{\mathrm{coeffs}}) \sqrt{\widehat{\sigma}^2 + \sigma_{\mathrm{fit}}^2}, \quad \text{(D1)}$$

where q_t is the student-t quantile for a confidence level x, $\widehat{\sigma}^2$ is the error variance estimator from the (weighted) mean-square error of the calibration data under the fit, and σ_{fit}^2 is the standard error estimate of the fitted model, which for a single-stage fit is

$$\sigma_{\mathrm{fit}}^2 = \mathrm{grad}^t(\eta, \chi_1, \chi_2) \cdot C_{\mathrm{fit}} \cdot \mathrm{grad}(\eta, \chi_1, \chi_2) \quad \text{(D2)}$$

with the gradient vector $\mathrm{grad}(\eta, \chi_1, \chi_2)$ of the fit ansatz in the coefficients, evaluated at this point, and the covariance matrix C_{fit} of the fit. Note that Eq. (D1) gives the uncertainty for a single additional observation, as opposed to the narrower *confidence interval* of the mean prediction, which lacks the $\widehat{\sigma}^2$ term.

In our hierarchical fitting approach, to propagate the uncertainties from the nonspinning, equal-mass and extreme-mass-ratio limits, we have to assume that the uncertainties in these regimes and that of the final fit are independent, so that we can take the full covariance matrix as a block-diagonal composition of these four contributions. The half width of a prediction

interval at confidence x is then

$$q_t\left(x, N_{\text{data}} - N_{\text{coeffs}}\right)\sqrt{\widehat{\sigma}^2 + \sigma_{\text{final}}^2 + \sigma_\eta^2 + \sigma_{\widehat{S}}^2 + \sigma_{\eta=0}^2}\,. \quad \text{(D3)}$$

As these three particular regimes are significantly better constrained than the bulk of the parameter space (which is the

main motivation for the hierarchical approach, in the first place), their uncertainty contribution is small, so that the accuracy of this approximation is not critical.

The covariance matrices for the fits to both final spin and radiated energy are provided in ASCII format as supplementary material. The estimated variances are $\widehat{\sigma}_{\chi_{\text{f}}}^2 \approx 6.225 \times 10^{-8}$ for final spin and $\widehat{\sigma}_{E_{\text{rad}}}^2 \approx 2.061 \times 10^{-8}$ for radiated energy.

[1] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 061102 (2016), arXiv:1602.03837 [gr-qc].
[2] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 241102 (2016), arXiv:1602.03840 [gr-qc].
[3] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 241103 (2016), arXiv:1606.04855 [gr-qc].
[4] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. **X6**, 041015 (2016), arXiv:1606.04856 [gr-qc].
[5] R. P. Kerr, Phys. Rev. Lett. **11**, 237 (1963).
[6] F. Pretorius, Phys. Rev. Lett. **95**, 121101 (2005), arXiv:gr-qc/0507014 [gr-qc].
[7] M. Campanelli, C. O. Lousto, P. Marronetti, and Y. Zlochower, Phys. Rev. Lett. **96**, 111101 (2006), arXiv:gr-qc/0511048 [gr-qc].
[8] J. G. Baker, J. Centrella, D.-I. Choi, M. Koppitz, and J. van Meter, Phys. Rev. Lett. **96**, 111102 (2006), arXiv:gr-qc/0511103 [gr-qc].
[9] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. Jiménez Forteza, and A. Bohé, Phys. Rev. **D93**, 044006 (2016), arXiv:1508.07250 [gr-qc].
[10] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jiménez Forteza, and A. Bohé, Phys. Rev. **D93**, 044007 (2016), arXiv:1508.07253 [gr-qc].
[11] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, Phys. Rev. Lett. **113**, 151101 (2014), arXiv:1308.3271 [gr-qc].
[12] A. Bohé, M. Hannam, S. Husa, F. Ohme, M. Pürrer, and P. Schmidt, *PhenomPv2 – technical notes for the LAL implementation*, Tech. Rep. LIGO-T1500602 (LIGO Scientific Collaboration and Virgo Collaboration, 2016).
[13] A. Bohé, M. Pürrer, M. Hannam, S. Husa, F. Ohme, P. Schmidt, *et al.*, "IMRPhenomPv2 publication," in preparation.
[14] W. H. Press, Astrophys. J. **170**, L105 (1971).
[15] S. Chandrasekhar and S. L. Detweiler, Proc. Roy. Soc. Lond. **A344**, 441 (1975).
[16] S. L. Detweiler, Astrophys. J. **239**, 292 (1980).
[17] K. D. Kokkotas and B. G. Schmidt, Living Rev. Rel. **2**, 2 (1999), arXiv:gr-qc/9909058 [gr-qc].
[18] P. Ajith *et al.*, Phys. Rev. Lett. **106**, 241101 (2011), arXiv:0909.2867 [gr-qc].
[19] A. Taracchini *et al.*, Phys. Rev. **D89**, 061502 (2014), arXiv:1311.2544 [gr-qc].
[20] Y. Pan, A. Buonanno, A. Taracchini, L. E. Kidder, A. H. Mroué, H. P. Pfeiffer, M. A. Scheel, and B. Szilágyi, Phys. Rev. **D89**, 084006 (2014), arXiv:1307.6232 [gr-qc].
[21] A. Ghosh *et al.*, Phys. Rev. **D94**, 021101 (2016), arXiv:1602.02453 [gr-qc].
[22] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 221101 (2016), arXiv:1602.03841 [gr-qc].
[23] C. G. Lacey and S. Cole, Mon. Not. Roy. Astron. Soc. **262**, 627 (1993).

[24] B. F. Roukema, B. A. Peterson, P. J. Quinn, and B. Rocca-Volmerange, Mon. Not. Roy. Astron. Soc. **292**, 835 (1997), arXiv:astro-ph/9707294 [astro-ph].
[25] M. Volonteri, F. Haardt, and P. Madau, Astrophys. J. **582**, 559 (2003), arXiv:astro-ph/0207276 [astro-ph].
[26] T. Bogdanovic, C. S. Reynolds, and M. C. Miller, Astrophys. J. Lett. **661**, L147 (2007), arXiv:astro-ph/0703054 [astro-ph].
[27] W. Tichy and P. Marronetti, Phys. Rev. **D78**, 081501 (2008), arXiv:0807.2985 [gr-qc].
[28] E. Barausse, Mon. Not. Roy. Astron. Soc. **423**, 2533 (2012), arXiv:1201.5888 [astro-ph.CO].
[29] A. Sesana, E. Barausse, M. Dotti, and E. M. Rossi, Astrophys. J. **794**, 104 (2014), arXiv:1402.7088 [astro-ph.CO].
[30] A. Klein, E. Barausse, A. Sesana, A. Petiteau, E. Berti, S. Babak, J. Gair, S. Aoudia, I. Hinder, F. Ohme, and B. Wardell, Phys. Rev. **D93**, 024003 (2016), arXiv:1511.05581 [gr-qc].
[31] J. E. McClintock, R. Narayan, S. W. Davis, L. Gou, A. Kulkarni, J. A. Orosz, R. F. Penna, R. A. Remillard, and J. F. Steiner, *General relativity and gravitation. Proceedings, 19th International Conference, GR19, Mexico City, Mexico, July 4-9, 2010*, Class. Quant. Grav. **28**, 114009 (2011), arXiv:1101.0811 [astro-ph.HE].
[32] A. B. Nielsen, *Proceedings, 11th Edoardo Amaldi Conference on Gravitational Waves (Amaldi 11): Gwangju, South Korea, June 21-26, 2015*, J. Phys. Conf. Ser. **716**, 012002 (2016), arXiv:1604.00778 [astro-ph.HE].
[33] L. Brenneman, *Measuring the Angular Momentum of Supermassive Black Holes*, SpringerBriefs in Astronomy (Springer, 2013) arXiv:1309.6334 [astro-ph.HE].
[34] C. S. Reynolds, Class. Quant. Grav. **30**, 244004 (2013), arXiv:1307.3246 [astro-ph.HE].
[35] B. M. Peterson, Space Science Reviews **183**, 253 (2014).
[36] T. Damour, Phys. Rev. **D64**, 124013 (2001), arXiv:gr-qc/0103018 [gr-qc].
[37] E. Racine, Phys. Rev. **D78**, 044021 (2008), arXiv:0803.1820 [gr-qc].
[38] A. Bohé, S. Marsat, and L. Blanchet, Class. Quant. Grav. **30**, 135009 (2013), arXiv:1303.7412 [gr-qc].
[39] S. Marsat, A. Bohé, L. Blanchet, and A. Buonanno, Class. Quant. Grav. **31**, 025023 (2014), arXiv:1307.6793 [gr-qc].
[40] L. Santamaria *et al.*, Phys. Rev. **D82**, 064016 (2010), arXiv:1005.3306 [gr-qc].
[41] P. Ajith, Phys. Rev. **D84**, 084037 (2011), arXiv:1107.1267 [gr-qc].
[42] B. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. **X6**, 041014 (2016), arXiv:1606.01210 [gr-qc].
[43] S. Babak, A. Taracchini, and A. Buonanno, Phys. Rev. **D95**, 024010 (2017), arXiv:1607.05661 [gr-qc].
[44] J. M. Bardeen, W. H. Press, and S. A. Teukolsky, Astrophys. J. **178**, 347 (1972).
[45] H. Akaike, IEEE Trans. Autom. Control **19**, 716 (1974).
[46] G. E. Schwarz, Annals of Statistics **6**, 461 (1978).
[47] A. Buonanno, L. E. Kidder, and L. Lehner, Phys. Rev. **D77**, 026004 (2008), arXiv:0709.3839 [astro-ph].

[48] L. Rezzolla, P. Diener, E. N. Dorband, D. Pollney, C. Reisswig, E. Schnetter, and J. Seiler, Astrophys. J. **674**, L29 (2008), arXiv:0710.3345 [gr-qc].

[49] L. Rezzolla, E. Barausse, E. N. Dorband, D. Pollney, C. Reisswig, J. Seiler, and S. Husa, Phys. Rev. **D78**, 044002 (2008), arXiv:0712.3541 [gr-qc].

[50] E. Barausse and L. Rezzolla, Astrophys. J. **704**, L40 (2009), arXiv:0904.2577 [gr-qc].

[51] E. Barausse, V. Morozova, and L. Rezzolla, Astrophys. J. **758**, 63 (2012), [Erratum: Astrophys. J.786,76(2014)], arXiv:1206.3803 [gr-qc].

[52] J. Healy, C. O. Lousto, and Y. Zlochower, Phys. Rev. **D90**, 104004 (2014), arXiv:1406.7295 [gr-qc].

[53] Y. Zlochower and C. O. Lousto, Phys. Rev. **D92**, 024022 (2015), arXiv:1503.07536 [gr-qc].

[54] Y. Zlochower and C. O. Lousto, Phys. Rev. **D94**, 029901 [Erratum] (2016).

[55] F. Hofmann, E. Barausse, and L. Rezzolla, Astrophys. J. **825**, L19 (2016), arXiv:1605.01938 [gr-qc].

[56] X. Jiménez-Forteza, D. Keitel, S. Husa, M. Hannam, S. Khan, and M. Pürrer, "Phys. Rev. D95 064024 Supplemental Material," (); "arXiv:1611.00332v2 ancillary files," ().

[57] J. Veitch et al., Phys. Rev. **D91**, 042003 (2015), arXiv:1409.7215 [gr-qc].

[58] LIGO Scientific Collaboration, "LSC Algorithm Library - LALSuite," free software.

[59] A. H. Mroué et al., Phys. Rev. Lett. **111**, 241104 (2013), arXiv:1304.6077 [gr-qc].

[60] The SXS Collaboration, "SXS Gravitational Waveform Database," (2016).

[61] J. Healy and C. O. Lousto, Phys. Rev. **D95**, 024037 (2017), arXiv:1610.09713 [gr-qc].

[62] K. Jani, J. Healy, J. A. Clark, L. London, P. Laguna, and D. Shoemaker, Class. Quant. Grav. **33**, 204001 (2016), arXiv:1605.03204 [gr-qc].

[63] K. Jani, J. Healy, J. A. Clark, L. London, P. Laguna, and D. Shoemaker, "Georgia Tech catalog of binary black hole simulations," (2016).

[64] B. Bruegmann, J. A. Gonzalez, M. Hannam, S. Husa, U. Sperhake, and W. Tichy, Phys. Rev. **D77**, 024027 (2008), arXiv:gr-qc/0610128 [gr-qc].

[65] S. Husa, J. A. Gonzalez, M. Hannam, B. Bruegmann, and U. Sperhake, Class. Quant. Grav. **25**, 105006 (2008), arXiv:0706.0740 [gr-qc].

[66] M. Davis, R. Ruffini, W. H. Press, and R. H. Price, Phys. Rev. Lett. **27**, 1466 (1971).

[67] S. A. Hughes and R. D. Blandford, Astrophys. J. **585**, L101 (2003), arXiv:astro-ph/0208484 [astro-ph].

[68] L. Rezzolla, E. Barausse, E. N. Dorband, D. Pollney, C. Reisswig, J. Seiler, and S. Husa, Phys. Rev. D **78**, 044002 (2008).

[69] N. K. Johnson-McDaniel et al., Determining the final spin of a binary black hole system including in-plane spins: Method and checks of accuracy, Tech. Rep. LIGO-T1600168 (LIGO Scientific Collaboration and Virgo Collaboration, 2016).

[70] D. A. Hemberger, G. Lovelace, T. J. Loredo, L. E. Kidder, M. A. Scheel, B. Szilágyi, N. W. Taylor, and S. A. Teukolsky, Phys. Rev. **D88**, 064014 (2013), arXiv:1305.5991 [gr-qc].

[71] S. W. Hawking, Phys. Rev. Lett. **26**, 1344 (1971).

[72] X. Jiménez Forteza, D. Keitel, S. Husa, M. Hannam, S. Khan, L. London, and M. Pürrer, Phenomenological fit of the peak luminosity from non-precessing binary-black-hole coalescences, Tech. Rep. LIGO-T1600018-v4 (LIGO Scientific Collaboration and Virgo Collaboration, 2016).

[73] D. Keitel, X. Jiménez-Forteza, S. Husa, L. London, et al., (2016), arXiv:1612.09566 [gr-qc].

[74] O. Dreyer, B. Krishnan, D. Shoemaker, and E. Schnetter, Phys. Rev. **D67**, 024018 (2003), arXiv:gr-qc/0206008 [gr-qc].

[75] I. Hinder et al., Class. Quant. Grav. **31**, 025012 (2014), arXiv:1307.5307 [gr-qc].

[76] E. Berti, V. Cardoso, J. A. Gonzalez, U. Sperhake, M. Hannam, S. Husa, and B. Bruegmann, Phys. Rev. **D76**, 064034 (2007), arXiv:gr-qc/0703053 [GR-QC].

[77] J. C. Bustillo, P. Laguna, and D. Shoemaker, (2016), arXiv:1612.02340 [gr-qc].

[78] B. R. Iyer and C. M. Will, Phys. Rev. Lett. **70**, 113 (1993).

[79] B. R. Iyer and C. M. Will, Phys. Rev. **D52**, 6882 (1995).

[80] P. Jaranowski and G. Schaefer, Phys. Rev. **D55**, 4712 (1997).

[81] M. E. Pati and C. M. Will, Phys. Rev. **D65**, 104008 (2002), arXiv:gr-qc/0201001 [gr-qc].

[82] C. Konigsdorffer, G. Faye, and G. Schaefer, Phys. Rev. **D68**, 044004 (2003), arXiv:gr-qc/0305048 [gr-qc].

[83] L. Blanchet, Comptes Rendus Physique **8**, 57 (2007), arXiv:gr-qc/0611142 [gr-qc].

[84] S. Nissanke and L. Blanchet, Class. Quant. Grav. **22**, 1007 (2005), arXiv:gr-qc/0412018 [gr-qc].

[85] D. Shoemaker et al., private communication (2016).

[86] A. R. Liddle, Mon. Not. Roy. Astron. Soc. **377**, L74 (2007), arXiv:astro-ph/0701113 [astro-ph].

[87] J. J. Faraway, "Linear models with R," (CRC Press, Boca Raton, 2005) Chap. 3.5, pp. 41–44.