

## Systematic effects from black hole-neutron star waveform model uncertainties on the neutron star equation of state

Kabir Chakravarti,<sup>1</sup> Anuradha Gupta,<sup>2,1</sup> Sukanta Bose,<sup>1,3,\*</sup> Matthew D. Duez,<sup>3</sup> Jesus Caro,<sup>3</sup> Wyatt Brege,<sup>3</sup> Francois Foucart,<sup>4,5</sup> Shaon Ghosh,<sup>6</sup> Koutarou Kyutoku,<sup>7,8,9,10</sup> Benjamin D. Lackey,<sup>11</sup> Masaru Shibata,<sup>11,10</sup> Daniel A. Hemberger,<sup>12</sup> Lawrence E. Kidder,<sup>13</sup> Harald P. Pfeiffer,<sup>11,14</sup> and Mark A. Scheel<sup>12</sup>

<sup>1</sup>*Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411 007, India*

<sup>2</sup>*Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802, USA*

<sup>3</sup>*Department of Physics & Astronomy, Washington State University,  
1245 Webster, Pullman, Washington 99164-2814, USA*

<sup>4</sup>*Department of Physics, University of New Hampshire,  
9 Library Way, Durham, New Hampshire 03824, USA*

<sup>5</sup>*Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, California 94720, USA*

<sup>6</sup>*Department of Physics, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201, USA*

<sup>7</sup>*Theory Center, Institute of Particle and Nuclear Studies, KEK, Tsukuba 305-0801, Japan*

<sup>8</sup>*Department of Particle and Nuclear Physics, the Graduate University for Advanced Studies (Sokendai),  
Tsukuba 305-0801, Japan*

<sup>9</sup>*Interdisciplinary Theoretical and Mathematical Sciences Program (iTHEMS),  
RIKEN, Wako, Saitama 351-0198, Japan*

<sup>10</sup>*Center for Gravitational Physics, Yukawa Institute for Theoretical Physics,  
Kyoto University, Kyoto 606-8502, Japan*

<sup>11</sup>*Max Planck Institute for Gravitational Physics (Albert Einstein Institute),  
D-14476 Potsdam-Golm, Germany*

<sup>12</sup>*TAPIR, Walter Burke Institute for Theoretical Physics, California Institute of Technology,  
MC 350-17, Pasadena, California 91125, USA*

<sup>13</sup>*Cornell Center for Astrophysics and Planetary Science, Cornell University,  
Ithaca, New York, 14853, USA*

<sup>14</sup>*Canadian Institute for Theoretical Astrophysics, University of Toronto,  
Toronto, Ontario M5S 3H8, Canada*



(Received 19 September 2018; published 31 January 2019)

We identify various contributors of systematic effects in the measurement of the neutron star (NS) tidal deformability and quantify their magnitude for several types of neutron star—black hole (NSBH) binaries. Gravitational waves from NSBH mergers contain information about the components' masses and spins as well as the NS equation of state. Extracting this information requires comparison of the signal in noisy detector data with theoretical templates derived from some combination of post-Newtonian (PN) approximants, effective one-body (EOB) models, and numerical relativity (NR) simulations. The accuracy of these templates is limited by errors in the NR simulations, by the approximate nature of the PN/EOB waveforms, and by the hybridization procedure used to combine them. In this paper, we estimate the impact of these errors by constructing and comparing a set of PN-NR hybrid waveforms, for the first time with NR waveforms from two different codes, namely, SPEC and SACRA, for such systems. We then attempt to recover the parameters of the binary using *two* non-precessing template approximants. As expected, these errors have negligible effect on detectability. Mass and spin estimates are moderately affected by systematic errors for near equal-mass binaries, while the recovered masses can be inaccurate at higher mass ratios. Large uncertainties are also found in the tidal deformability  $\Lambda$ , due to differences in PN base models used in hybridization, numerical relativity NR errors, and inherent limitations of the hybridization method. We find that systematic errors are too large for tidal effects to be accurately characterized for any realistic NS equation of state model. We conclude that NSBH waveform models must be significantly improved if they are to be useful for the extraction of NS equation of state information or even for distinguishing NSBH systems from binary black holes.

DOI: [10.1103/PhysRevD.99.024049](https://doi.org/10.1103/PhysRevD.99.024049)

\*Corresponding author.  
sukanta@iucaa.in

## I. INTRODUCTION

Compact object binaries are of paramount importance in gravitational wave (GW) physics because their signals are loud, have a significant rate of occurrence, and are amenable to theoretical modelling. It is perhaps only a matter of time before neutron star-black hole (NSBH) binaries are discovered [1,2]. The phenomenological modeling of such systems is thus an important problem today. From the point of view of tidal effects, an NSBH binary constitutes a simplified version of the binary neutron star (BNS) system, as it involves only one tidal parameter. Consequently, any inference drawn on the NS equation of state (EOS) from their GW signal would be cleaner than that from BNSs.

A few phenomenological (Phenom) models of NSBH binary waveforms already exist. Post-Newtonian (PN) and numerical relativity (NR) waveforms were first combined to make complete NSBH waveform families by Lackey *et al.* [3]. They created two template families, based on different analytical waveform models. One used the aligned-spin IMRPhenomC binary black hole (BBH) model [4], while the other used a time-domain effective one-body (EOB) model [5]. The merger phase and amplitude were then modified by analytic correction functions calibrated to 134 NR simulations produced by the SACRA code [6]. We will henceforth call these waveforms the “LEA” model. As an extension of LEA, Pannarale *et al.* [7] introduced a more detailed frequency-domain model for the waveform amplitude by distinguishing systems with various degrees of expected tidal disruption. Finally, Kumar *et al.* [8], as part of a study of systematic errors in mass and spin estimates due to tidal effects, produced an enhanced version of the LEA templates, using SEOBNRv2 [9] as the underlying BBH model, with the same tidal corrections as LEA. We call the resulting templates the “LEA+” model.

In this paper, we identify key systematic errors that can affect the measurement of the NS tidal deformability parameter,  $\Lambda$ , in possible NSBH detections. Cross-code comparisons are common for BBH [10–12]. There have also been some previous attempts at comparing NR simulations with models for BNS systems [13–18] as well as for NSBH systems [19,20]. Very recently, waveforms for inspiral of compact binaries were also proposed from the EOB formalism [21]. The added complexities of the micro physics make cross-code comparisons of NSBH/BNS systems arguably even more important. This is the first time that results impacting NSBH parameter estimation are obtained by comparing waveforms based on simulations carried out with two different NR codes, namely, SPEC [22] and SACRA [6]. The paper is organized as follows: In Sec. II we summarize the different sources of systematic errors in the tidal deformability estimation. We also describe the NR late-inspiral-merger-ringdown waveforms and the procedure for constructing hybrid waveforms by combining them with PN inspiral cycles. In Sec. III, we present the

results of parameter estimation for these hybrid waveforms using the LEA and LEA+ models, before concluding in Sec. IV.

## II. WAVEFORM SYSTEMATICS

The characterization of NSBH binaries involves the estimation of a NS tidal parameter, apart from the masses and spins of the binary components. Here we will take the dimensionless tidal deformability [23] to represent the former. We follow the LEA convention and denote this parameter by  $\Lambda \equiv (2/3)k_2(c^2R_{\text{NS}}/GM_{\text{NS}})^5$ , where  $k_2$ ,  $R_{\text{NS}}$  and  $M_{\text{NS}}$  are the second Love number, radius and mass, respectively, of the neutron star. There are multiple sources of systematics that will affect the error budget for  $\Lambda$ . At the outset they can be grouped into the following broad categories.

- (1) NR modeling: To begin with, there will be systematics due to errors in the NR waveforms, namely, the effects of finite resolution and of the extrapolation of waveforms to infinity. In the limit of infinite resolution and perfect initial data, all evolution methods are equivalent, but with the introduction of a finite grid (or finite extraction radius), the equivalence is broken and so numerical errors become dependent on the process chosen to evolve the field equations. There will also be errors due to imperfect initial data leading to, for example, nonzero residual eccentricities or “junk” GW radiation.
- (2) Choice of EOS model: The choice of EOS model may also influence parameter estimation. To first order,  $\Lambda$  is the best measured matter property of the NS from the premerger phase of the GW signal from an NSBH. Yet, the space of EOSs is not one dimensional, and two different stars with the same mass and  $\Lambda$  can be constructed from two different EOS models, e.g., (somewhat simplistically) a single polytrope and a piecewise polytrope. In this sense, the choice of the EOS model used by NR can distort the GW waveform computed because the true EOS presumably does not fall perfectly within the EOS family used in the NR survey, albeit possibly only during the last few orbits of inspiral and during/after merger.
- (3) Choice of PN waveform: When constructing hybrid waveforms, the choice of the PN approximant used is also a potential source of error.
- (4) The PN-NR bridge: The method for combining the PN and NR cycles to produce complete waveforms in the detector band—also termed as “hybridization”—can be a source of systematic error. This is true even if the technique itself is perfect, in that it produces an accurate complete waveform if the PN and NR parts themselves are accurate. This is because as long as even one set of cycles (PN or NR) is erroneous or the number of overlapping PN and NR cycles is too few,

the hybridization procedure can create a waveform that is inaccurate even at frequencies where the original PN or NR parts were in themselves accurate.

- (5) Sampling the ambiguity function in  $\Lambda$ : Accurate parameter estimation depends on the knowledge and application of accurate waveform models. Estimating the value of a signal parameter then involves cross-correlating normalized templates based on such waveforms with the GW data containing a signal. As detailed here, the rate at which this cross-correlation (or, more precisely, the *match*, as defined below) changes with  $\delta\Lambda \equiv \Lambda_{\text{template}} - \Lambda_{\text{signal}}$  is slow: When the signal itself is a unit-norm “neighboring” template, the match drops by a few tenths of a percent even when  $\delta\Lambda$  is as high as several tens to a few hundreds. This is just another way of stating that the ambiguity function [24] in  $\Lambda$  is diffuse or not sharply peaked. For strong signals, this spread in  $\Lambda$  will reduce, and its estimation will be more *precise*. It does not impact estimation accuracy by itself. However, we show here that even for reasonably accurate templates if the sampling rate of the data (and the templates) is not considerably higher than 4096 Hz, the match can have multiple closely-spaced local maxima, which can cause parameter estimation algorithms to miss the global maximum. This effect can make  $\Lambda$  estimation inaccurate.
- (6) The mass-ratio effect: The mass ratio of the binary (i.e.,  $q \equiv M_{\text{BH}}/M_{\text{NS}}$ , where  $M_{\text{BH}}$  and  $M_{\text{NS}}$  are the black hole and neutron star masses, respectively) inversely dictates the magnitude of the tidal effect in its signal. With increasing  $q$ , tidal effects on the waveform are reduced. Systematic errors like the ones we discussed start to dominate over the physical tidal effects, and this means that it is harder to estimate the tidal parameter accurately. Note that this is in addition to the *statistical* effect whereby a smaller value of a parameter will incur a larger spread in its measured value. The high mass ratios also tend to make NR simulations harder.

The first five systematic errors can be reduced through improved modeling or better parameter estimation algorithms and their implementation in data in the future. The last effect is related to actual physical processes, and tells us where modeling errors will impact parameter estimation the most (i.e., at high mass ratios).

### A. Systematics from numerics

Since tidal effects are expected to be small, *numerical errors* can significantly impact measurements of the tidal deformability in NSBH binaries. In this paper, we employ waveforms from two of the major NR codes used to model NSBH mergers: SACRA [6] and SPEC [22]. These two codes were written independently of each other and use different formulations of general relativity, grid structures,

and numerical methods. The SACRA waveforms utilized here are from the same set that was used to calibrate LEA. Some of the SPEC waveforms come from simulations recently used to study the impact of dynamical tides on NSBH waveforms [25], while others are presented for the first time here (see below). Both SPEC and SACRA simulations have drastically improved in phase accuracy in recent years with the introduction of sophisticated techniques such as higher-order hydrodynamics [26] and constraint propagation [27]. Since we use waveforms made over a period of several years, numerical errors vary significantly from system to system.

There are several possible sources of NR error. One source is residual eccentricity in the initial data. SPEC simulations use an iterative approach [28] to reduce the initial eccentricity, resulting in residual eccentricities  $e \sim 0.0005\text{--}0.003$  for the simulations presented here. The SACRA simulations employed to calibrate LEA, on the other hand, used quasi-circular initial data, with typical residual eccentricities  $e \sim 0.01$ . Another source of error is the extrapolation to future null infinity. In SPEC, this error is estimated by comparing extrapolation methods of different orders [29], and is found to be small compared to other numerical errors (phase errors of  $\sim 0.01$  rad) [30]. For the SACRA simulations used in LEA, gravitational waves were extracted at a fixed radius of  $\sim 1000$  km instead. We note that the eccentricity reduction and extrapolation to null infinity are both performed in BNS simulations by SACRA [31], and that updated NSBH simulations are ongoing.

At the current accuracy of numerical simulations, however, the largest source of error is truncation error. The magnitude of that error can be estimated by comparing simulations at different resolutions, and in some cases by extrapolating results to infinite resolution. The most recent SPEC simulations (cf. Cases 1 and 2 in Tables I and II) accumulate phase difference of about 0.1–0.3 radian at merger over 20–30 waveform cycles.

These errors must be compared to the expected impact of finite size effects on the phase of the GW signal. This is a very strong function of the mass ratio of the binary. Finite size effects are negligible for  $q \gtrsim 6$ , non-spinning binary. For  $q = 6$ , the result of BBH simulations falls within the error bars of NSBH simulations [32]. On the other hand, even our most conservative error estimates predict that finite size effects are resolved with  $\sim 20\%$  relative errors at the time of merger for the lowest mass ratios considered here ( $q = 1\text{--}2$ , nonspinning binaries [25]). Finally, for high mass-ratio binaries with rapidly spinning black holes, the rapid falloff of the GW amplitude due to the disruption of the NS is the strongest finite-size effect on the waveform, and is well captured by simulations.

Analytical waveform templates may also have errors induced by their calibration to a set of NR simulations using a restricted family of EOS, i.e., EOS with a fixed functional form and a set of freely specifiable parameters

TABLE I. The parameter estimates and fitting factors for a set of hybridized NSBH waveforms, constructed from NR late-inspiral and merger cycles at different resolutions from SpEC and SACRA, when match-filtered with LEA templates in aLIGO noise PSD. For example, Case 1 shows three resolutions of SpEC waveforms with  $q = 1.5$  and  $\Lambda = 791$  recovered against LEA templates. Asterisk-marked cases were hybridized using the step method.

Injected waveform		Best matched parameters				FF %
Resolution	Code	$M_{\text{BH}}$	$M_{\text{NS}}$	$\chi_{\text{BH}}$	$\Lambda$	
Case 1. $q = 1.5$ ; $\chi_{\text{BH}} = 0$ ; $M_{\text{NS}} = 1.4$ ; EOS $\Gamma = 2$ ; $\Lambda = 791$						
Low	SpEC	2.06	1.42	0.001	579.0	99.80
Medium	SpEC	2.06	1.41	-0.005	546.3	99.80
High	SpEC	2.10	1.39	-0.006	551.3	99.77
Case 2. $q = 2$ ; $\chi_{\text{BH}} = 0$ ; $M_{\text{NS}} = 1.4$ ; EOS $\Gamma = 2$ ; $\Lambda = 791$						
Low	SpEC	2.76	1.41	-0.008	-785.0	99.72
Medium	SpEC	2.79	1.40	-0.002	-659.1	99.71
High	SpEC	2.71	1.43	-0.021	-742.9	99.75
Case 3. $q = 2$ ; $\chi_{\text{BH}} = 0.75$ ; $M_{\text{NS}} = 1.2$ ; EOS 2H; $\Lambda = 4382$						
Medium	SpEC	2.35	1.21	0.813	4097	98.94
High	SACRA	2.53	1.22	0.816	3977	99.86
Case 4. $q = 3$ ; $\chi_{\text{BH}} = 0$ ; $M_{\text{NS}} = 1.4$ ; EOS $\Gamma = 2$ ; $\Lambda = 620$						
Low	SpEC	4.00	1.46	-0.031	-107.7	98.97
Medium	SpEC	4.00	1.46	-0.027	-103.8	98.95
High	SpEC	4.00	1.46	-0.034	-2288	99.49
Case 5. $q = 3$ ; $\chi_{\text{BH}} = 0$ ; $M_{\text{NS}} = 1.35$ ; EOS H; $\Lambda = 607$						
Medium	SpEC*	3.86	1.40	-0.038	-6358	99.41
High	SpEC*	3.99	1.37	-0.000	594.1	97.70
Low	SACRA	3.79	1.43	-0.042	-426.7	99.05
Medium	SACRA	3.86	1.40	-0.036	-3750	99.61
High	SACRA	3.86	1.40	-0.029	-891	99.09
Case 6. $q = 5$ ; $\chi_{\text{BH}} = 0.75$ ; $M_{\text{NS}} = 1.35$ ; EOS 2H; $\Lambda = 2324$						
Low	SACRA*	5.70	1.53	0.859	2367	97.46
Medium	SACRA*	5.92	1.49	0.861	16784	97.31
High	SACRA*	5.82	1.51	0.849	121.5	97.58
Case 7. $q = 5$ ; $\chi_{\text{BH}} = 0.5$ ; $M_{\text{NS}} = 1.4$ ; EOS $\Gamma = 2$ ; $\Lambda = 791$						
Low	SpEC	5.28	1.73	0.475	1008	99.27
Medium	SpEC	5.30	1.72	0.476	997	99.28

that can capture some but not all of the properties of real NSs. Errors from this effect are expected to be small because, to leading order, finite size effects only depend on the tidal deformability of NSs, and the same range of  $\Lambda$  can be covered with many EOS families. The SACRA waveforms used for LEA assume a two-component piecewise polytropic EOS, i.e., an equation of state in which the pressure  $P$  is related to the baryon density  $\rho$  by

$$P = \begin{cases} \kappa_0 \rho^{\Gamma_0} & \text{if } \rho < \rho_t \\ \kappa_1 \rho^{\Gamma_1} & \text{if } \rho > \rho_t, \end{cases} \quad (1)$$

where  $\kappa_0 = 3.5966 \times 10^{13}$  in cgs units and  $\Gamma_0 = 1.3569$ . This EOS family has two free variables, which are taken to be  $\Gamma_1$ , the high-density polytropic exponent, and  $P_1$ , the pressure at a fiducial density  $\rho_{\text{fidu}} = 10^{14.7} \text{ g cm}^{-3}$ . It has been suggested in Ref. [33] that the pressure at this numerical value of  $\rho_{\text{fidu}}$  is correlated with the NS radius. The value of  $\rho_t$  is determined from these quantities by requiring the continuity of the pressure at the interface as  $\rho_t = (\kappa_0/\kappa_1)^{1/(\Gamma_1-\Gamma_0)}$  with  $\kappa_1 = P_1/\rho_{\text{fidu}}^{\Gamma_1}$ . In this paper, we will use NR waveforms generated with this EOS family as

TABLE II. Same as Table I, but now employing LEA + templates instead of LEA.

Injected waveform		Best matched parameters				FF %
Resolution	Code	$M_{\text{BH}}$	$M_{\text{NS}}$	$\chi_{\text{BH}}$	$\Lambda$	
Case 1. $q = 1.5$ ; $\chi_{\text{BH}} = 0$ ; $M_{\text{NS}} = 1.4$ ; EOS $\Gamma = 2$ ; $\Lambda = 791$						
Low	SpEC	2.08	1.41	0.000	587.6	99.40
Medium	SpEC	2.07	1.41	-0.000	723.6	99.39
High	SpEC	2.06	1.45	-0.015	373.5	99.41
Case 2. $q = 2$ ; $\chi_{\text{BH}} = 0$ ; $M_{\text{NS}} = 1.4$ ; EOS $\Gamma = 2$ ; $\Lambda = 791$						
Low	SpEC	2.75	1.42	-0.004	0.0	99.88
Medium	SpEC	2.78	1.40	-0.003	78.7	99.88
High	SpEC	2.78	1.40	-0.003	44.9	99.87
Case 3. $q = 2$ ; $\chi_{\text{BH}} = 0.75$ ; $M_{\text{NS}} = 1.2$ ; EOS 2H; $\Lambda = 4382$						
Medium	SpEC	2.47	1.16	0.760	4960	98.45
High	SACRA	2.38	1.20	0.770	4322	99.87
Case 4. $q = 3$ ; $\chi_{\text{BH}} = 0$ ; $M_{\text{NS}} = 1.4$ ; EOS $\Gamma = 2$ ; $\Lambda = 620$						
Low	SpEC	3.82	1.52	-0.046	663.6	98.97
Medium	SpEC	3.76	1.54	-0.059	796.2	98.95
High	SpEC	3.81	1.52	-0.042	580.4	99.49
Case 5. $q = 3$ ; $\chi_{\text{BH}} = 0$ ; $M_{\text{NS}} = 1.35$ ; EOS H; $\Lambda = 607$						
Medium	SpEC*	3.62	1.48	-0.056	278.9	99.48
High	SpEC*	3.63	1.48	-0.055	272.3	99.49
Low	SACRA	3.72	1.45	-0.041	375.8	99.49
Medium	SACRA	3.61	1.48	-0.063	366.3	99.65
High	SACRA	3.72	1.45	-0.040	456.9	99.54
Case 6. $q = 5$ ; $\chi_{\text{BH}} = 0.75$ ; $M_{\text{NS}} = 1.35$ ; EOS 2H; $\Lambda = 2324$						
Low	SACRA*	6.30	1.42	0.846	2298	97.67
Medium	SACRA*	5.79	1.52	0.846	64.1	97.93
High	SACRA*	6.25	1.43	0.845	1896	97.71
Case 7. $q = 5$ ; $\chi_{\text{BH}} = 0.5$ ; $M_{\text{NS}} = 1.4$ ; EOS $\Gamma = 2$ ; $\Lambda = 791$						
Low	SpEC	6.04	1.56	0.502	475.6	99.41
Medium	SpEC	6.11	1.54	0.509	795.2	99.35

well as those generated using a single-component polytropic EOS:  $P = \kappa \rho^\Gamma$ , with  $\Gamma = 2$ .

## B. Systematic errors from templates

In Sec. III, we compare PN-NR hybrid waveforms to the LEA and LEA + models. Our results are impacted by systematic errors in these models. For example, there exists known residual discrepancy between the LEA waveforms and the SACRA simulations used in their construction. According to Lackey *et al.* [3], systematic errors in the analytical Phenom waveform due to this fitting error are  $|\Delta\Lambda^{1/5}|/\sigma_{\Lambda^{1/5}} \sim 0.1$  (worse for  $q > 4$ ), where  $\sigma_{\Lambda^{1/5}}$ , the aLIGO statistical error for a source at 100 Mpc, is between 0.5 and 1 (see Fig. 15 of [3]). So  $|\Delta\Lambda|/\Lambda \approx 5|\Delta\Lambda^{1/5}|/\Lambda^{1/5} \approx 5\text{--}12\%$ , perhaps twice as big for  $q$  between 4 and 5. Systematic errors in the analytical EOB waveforms due to their discrepancy from their simulation input are about twice as big as those in the analytical Phenom waveforms [3]. Further, as an example of in built template systematic errors, there are known instances for LEA where the phase of the calibrating hybrid waveform is known to have significant amplitude mismatch with the analytical template (Figs. 21 and 22 of Ref. [3]), all of which add to the systematic error budget.

Systematic errors in the LEA and LEA + templates are also impacted by the choice of underlying point-particle

(BBH) waveform, as the models are constructed by adding tidal terms to BBH templates. Naturally, the performance of the tidal templates depends on the quality of the BBH model. This is the key difference between LEA and LEA +, and we will see that this non-trivially affects  $\Lambda$  estimates when we perform comparisons between LEA and LEA +.

### C. PN model and hybridization systematics

As we noted earlier, the choice of PN model and the method used to construct the full inspiral-merger-ringdown hybrid waveforms are themselves potential sources of systematic errors. In this study, the hybrid waveforms are constructed using the procedure outlined in Ref. [34], unless mentioned otherwise. We use the 3.5PN phase corrected SpinTaylorT4 PN approximant, with 1.5PN amplitude correction whereas tidal corrections are taken up to 6PN order. SpinTaylorT4 is one of the popular inspiral models and matches remarkably well with NR in the case of equal mass, non-spinning binaries [29], whereas for non-equal mass binaries, it is of comparable accuracy as other Taylor-approximants [35]. Therefore, we use the SpinTaylorT4 model for the inspiral part of the hybrid waveforms. The systematics due to the use of different PN approximants in constructing hybrids for nonprecessing NSBH systems will be studied in a future publication.

The hybridization is performed in the time domain. We minimize the integrated absolute squared difference between PN and NR waveforms in a time window where we have NR data and where the PN approximation is expected to be valid:

$$\delta = \int_0^T |h_{\text{PN}}(t) - a * h_{\text{NR}}(t, \vec{\mu})|^2 dt. \quad (2)$$

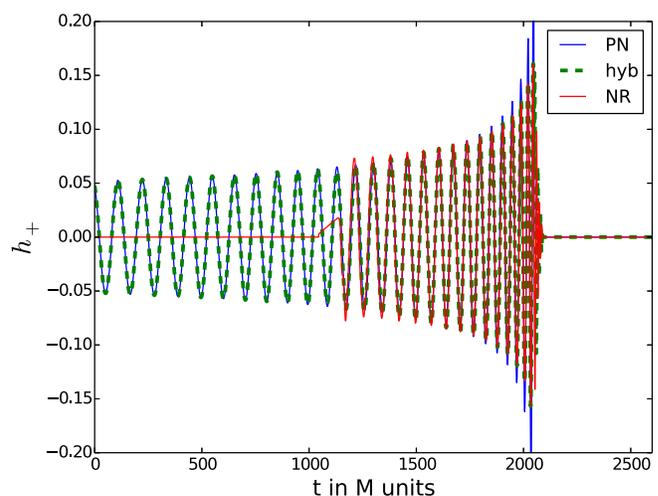
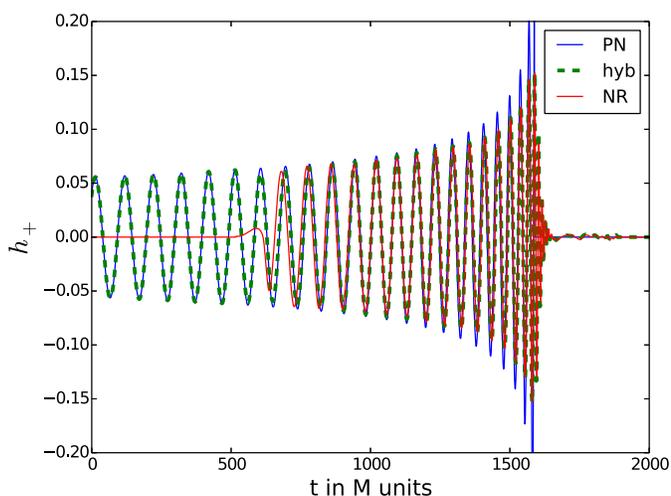


FIG. 1. Examples of hybrid waveforms generated for SPEC (right) and SACRA (left). Both have  $q = 5$ ,  $\chi_{\text{BH}} = 0.5$ . The time in total mass  $M$  at which the instantaneous frequency of the binaries crosses 10 Hz is, respectively,  $-7.143 \times 10^6$  and  $-7.863 \times 10^6$  for SPEC and SACRA.

Here  $[0, T]$  is the aforementioned overlap region of PN and NR waveforms in the time domain. We choose  $t = 0$  to be at an instantaneous signal frequency that is immediately after the point when the junk radiation in the NR signal was emitted. On the other hand,  $T$  is chosen more adaptively by allowing for as many cycles of the PN waveform into the minimization domain as possible without coming closer than a few cycles of the last stable circular orbit (LSCO). The minimization of  $\delta$  is carried out over an amplitude scaling factor  $a$  and the extrinsic parameters  $\vec{\mu}$  of the NR waveforms, namely, the initial phase  $\phi_0$  and the initial time of arrival  $t_0$ . Once the minimizing values (i.e.,  $a_{\text{min}}$  and  $\vec{\mu}_{\text{min}} = \{\phi_{0\text{min}}, t_{0\text{min}}\}$ ) are obtained, we make use of a linear interpolation joining function to produce the hybrid waveforms. The waveforms so produced can be expressed as:

$$h_{\text{hyb}}(t, \vec{\mu}) = h_{\text{PN}}(t)[1 - \tau(t)] + a_{\text{min}} h_{\text{NR}}(t, \vec{\mu}_{\text{min}})\tau(t), \quad (3)$$

where  $\tau(t)$  is the linear interpolation function, given by

$$\tau(t) = \begin{cases} 0 & \text{if } t \leq t_{0\text{min}} \\ \frac{t - t_{0\text{min}}}{t_{\text{max}} - t_{0\text{min}}} & \text{if } t_{0\text{min}} < t \leq t_{\text{max}} \\ 1 & \text{if } t_{\text{max}} < t. \end{cases} \quad (4)$$

Here,  $t_{\text{max}}$  is the time up to which we take the PN waveform to be valid; it is typically somewhat before the LSCO. Such an interpolation requires that NR cycles are available not just for the merger and ringdown phases, but also for at least several cycles of the late-inspiral phase at separations larger than the LSCO. This is the method used to produce the hybrids shown in Fig. 1.

It is important to note here that the ultimate test of the quality of these hybrid waveforms is how well the hybrid

matches a NR waveform with the same binary parameters that covers the complete band of the signal in Eq. (3). Phenomenological waveform models are approximations of such waveforms that are usually deduced by requiring high fitting factors against the hybrids. Such a criterion, however, does not guarantee that those waveform models will also estimate signal parameters highly accurately. Note that phenomenological models are calibrated to imperfect hybrid waveforms, and thus are themselves imperfect—even if they match perfectly the waveform they are calibrated against. Details on errors that can result from the hybridization procedure have been examined in Ref. [35]. Even though it analyzed only binary black hole systems, its findings have relevance for NSBH systems studied here, with the main difference being that the latter include an additional parameter, in the form of the NS tidal deformability  $\Lambda$ .

It is also worth mentioning that this hybridization method fails for our shortest NR waveforms, as the matching interval  $[0, T]$  does not contain more than a few GW cycles. When that is the case, we set  $a = 1$  in Eq. (2) and adaptively modify the function  $\tau$  in Eq. (3) to a step function. The hybrid waveform is then

$$h_{\text{hyb}}(t, \vec{\mu}) = h_{\text{PN}}(t)[1 - \tau(t)] + h_{\text{NR}}(t, \vec{\mu}_{\text{min}})\tau(t), \quad (5)$$

where

$$\tau(t) = \begin{cases} 0 & \text{if } t \leq t_{0\text{min}} \\ 1 & \text{if } t_{0\text{min}} < t, \end{cases} \quad (6)$$

and  $t_{0\text{min}}$  is defined as in Eqs. (3) and (4). Henceforth, we will refer to this method as the “step” method of hybridization.

In passing, we note that the regular method of hybridization is more likely to fail if there are large values of  $q$  or  $\chi_{\text{BH}}$  in the parameter space. This is partly due to the fact that the NR waveforms available for high  $q$  or high  $\chi_{\text{BH}}$  are usually shorter compared to the low mass, low spin waveforms.

### III. RESULTS

Our goal here is to study the effect of the various sources of systematic errors listed above on the estimation of the tidal deformability parameter. In practice, a Bayesian parameter estimation analysis is employed to infer the parameters of the source from its GW signal [36]. However, this method is computationally expensive and beyond the scope this paper (we will present the results for the Bayesian analysis in a future work). Instead we quote the maximum likelihood estimators of the parameters (i.e., parameter values of the best matched template) and the true values of the parameter; their respective differences are the errors in their measurement. Below we briefly describe how we compute the error in the tidal deformability parameter.

Let  $h(t, \vec{p}_h)$  denote a unit-norm hybrid waveform characterized by parameters  $\vec{p}_h = \{M_{\text{BH}}^h, M_{\text{NS}}^h, \chi_{\text{BH}}^h, \Lambda^h\}$ . Let  $u(t, \vec{p})$  be the generic form of a unit-norm template, where  $\vec{p} = \{M_{\text{BH}}, M_{\text{NS}}, \chi_{\text{BH}}, \Lambda\}$  are the dynamical parameters of the binary template. The inner product between  $u$  and  $h$ , i.e.,  $\langle u(f, \vec{p}) | h(f, \vec{p}_h) \rangle$ , maximized over the initial phase and time-of-arrival, is the match. The inner product itself is defined for vectors  $a$  and  $b$  as,

$$\langle a | b \rangle = 4\mathcal{R}e \left[ \int_{f_{\text{low}}}^{f_{\text{high}}} \frac{\tilde{a}(f)\tilde{b}^*(f)}{S_h(f)} df \right], \quad (7)$$

where  $S_h(f)$  is the one-sided power spectral density (PSD) of the noise of the detector. The fitting factor (FF) is the match maximized over the template parameters:

$$\text{FF} = \max_{\vec{p}, t_0, \phi_0} \langle \tilde{u}(f, \vec{p}) | \tilde{h}(f, \vec{p}_h) \rangle. \quad (8)$$

The maximization required for computing the FF was carried out using the well-known Nelder-Mead downhill simplex algorithm [37].

In this paper, we use the zero-detuned-high-power (ZDHP) noise-curve of aLIGO [38] for PSD calculation, and choose  $f_{\text{low}}$  to be 10 Hz. The systematic effects will therefore influence the bias  $\Delta p^\mu = (p_0^\mu - p_h^\mu) / p_h^\mu$ , where  $\mu$  is a parameter index and  $p_0^\mu$  denotes the maximizing values of the template parameters. We study below how large these biases are and explain the reason behind them.

The estimates of the tidal deformability parameter across different NR codes and numerical resolutions are summarized in Tables I and II. The cases where the step hybridization was applied are indicated in the same tables by asterisks. Further, estimates of the key nontidal parameters are presented in Fig. 2. The faithfulness of different tidal templates (namely, LEA vs LEA+) in estimating  $\Lambda$  can be inferred by comparing Tables I and II. Using these data, we will estimate the importance of the various systematic errors in turn.

#### A. Numerical sensitivity: Comparison of NR codes

Since SACRA and SPEC differ in gridding, formulation of Einstein’s equations (BSSN [39,40] vs generalized harmonic [41]), method of treating black holes (moving punctures vs excision), and gauge, that the two codes give consistent answers for gauge-invariant outputs, up to the truncation errors of each code, is an important test of both codes. For this test, we pick two binary configurations used in the construction of LEA and simulate them with the SPEC code. For both cases and for both codes, only the last roughly 5 orbits prior to merger are simulated. The initial data of the SPEC simulations has undergone an eccentricity reduction procedure, so the initial eccentricity for these runs (0.005–0.008) may be slightly lower than that of the SACRA runs. In this work, whenever piecewise polytropes

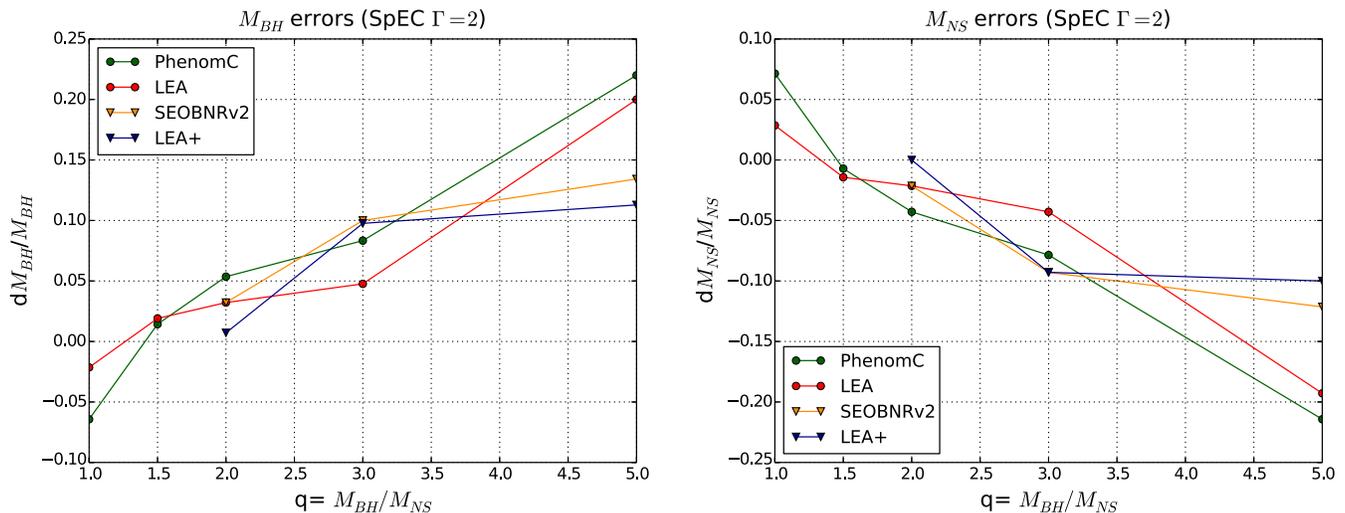


FIG. 2. Component mass estimation errors for templates based on two PN (IMRPhenomC & LEA) and two EOB (SEOBNRv2 & LEA+) models. Apart from  $q = 5$ , which has  $\chi_{\text{BH}} = 0.5$ , all other systems shown above are non-spinning.

were used for our NR simulations, be they with SACRA or SpEC, we used the same corresponding EOSs, namely H and 2H. We have employed both (relatively) old and new NR waveforms in this work. Use of old waveforms was made mainly because LEA and LEA+ templates were constructed with them. However, in order to probe the source of the systematic errors, we produced and employed some newer waveforms as well.

For our first comparison, we choose an extreme case that maximizes the importance of tidal effects: low mass-ratio ( $q = 2$ ), high BH spin ( $\chi_{\text{BH}} = 0.75$ ), and a large (even if somewhat unnatural in the light of GW170817 [42]), low-mass NS ( $M_{\text{NS}} = 1.2 M_{\odot}$ ,  $\Lambda = 4382$ ). This large NS radius is a consequence of the very stiff piecewise-polytropic EOS used in this case, labeled “2H” in [3]. Tidal effects are easy for both codes to resolve for this case. The results are listed in Case 3 in Table I, and the error in  $\Lambda$  turns out to be less than 10% for both codes.

For the second comparison, we choose a binary system close to the center of the LEA parameter space, namely, with  $q = 3$ ,  $\chi_{\text{BH}} = 0$ ,  $M_{\text{NS}} = 1.35 M_{\odot}$ ,  $\Lambda = 607$ . For this configuration, listed as Case 5 in Table I, the tidal effects are difficult to resolve, but we have as many as 3 SACRA resolutions and 2 SpEC resolutions to estimate errors. With the LEA templates, differences between the recovered  $\Lambda$  for different numerical resolutions is up to ten times the injected value of  $\Lambda$ , and we are thus entirely unable to reliably recover  $\Lambda$ . With LEA+, the recovered  $\Lambda$  is order-of-magnitude accurate, despite larger errors in the masses of the compact objects.

Next, we compare the  $l = 2$ ,  $m = 2$  mode of the decomposition of numerical waveforms in spherical harmonics in Fig. 3. In the top panel of that figure, we show for two different binaries the SACRA waveform, time-shifted relative to the SpEC results to maximize the match between

waveforms. The top left panel plots SACRA and SpEC waveforms for the first (high spin, 2H EOS) case. For this “easy” case, we find reasonable agreement, with a match around 90%. However, the agreement between SpEC and SACRA is not as good for the second (nonspinning, H EOS) case, shown on the top right of Fig. 3. For this case, visible dephasing is seen, and the match is lower than 70%. This disagreement indicates a large error in at least one set of simulations. It is only a sign of inconsistency between the two codes if it is larger than the numerical errors of the individual simulations. We thus turn to a consideration of the error of each code’s simulations, revealed by differences between resolutions.

A second SpEC simulation at higher resolution agrees with the lower-resolution to 0.1 radian through inspiral and 0.5 radian at the end of merger. However, details of SpEC’s adaptive mesh refinement algorithm can occasionally lead to significant overestimates of the numerical errors when only 2 simulations are used to estimate these errors. While the numerical errors are consistent with those measured in other recent SpEC NSBH simulations, a more rigorous error estimate would require a third simulation. On the other hand, the two available resolutions allow us at least to test the effect on parameter estimation of this level of phase error.

Convergence of the SACRA run is illustrated in the bottom panels of Fig. 3. If the time is shifted for optimal match, the agreement of different SACRA resolutions is reasonably good. For purposes of detection, this is the most important convergence check. If time shifting is not done (since different resolutions start from the same initial state), dephasing comparable to the SpEC-SACRA difference is seen. This more pessimistic comparison is more relevant for parameter estimation errors. That is, the ability to optimize match to a high value may give overly optimistic

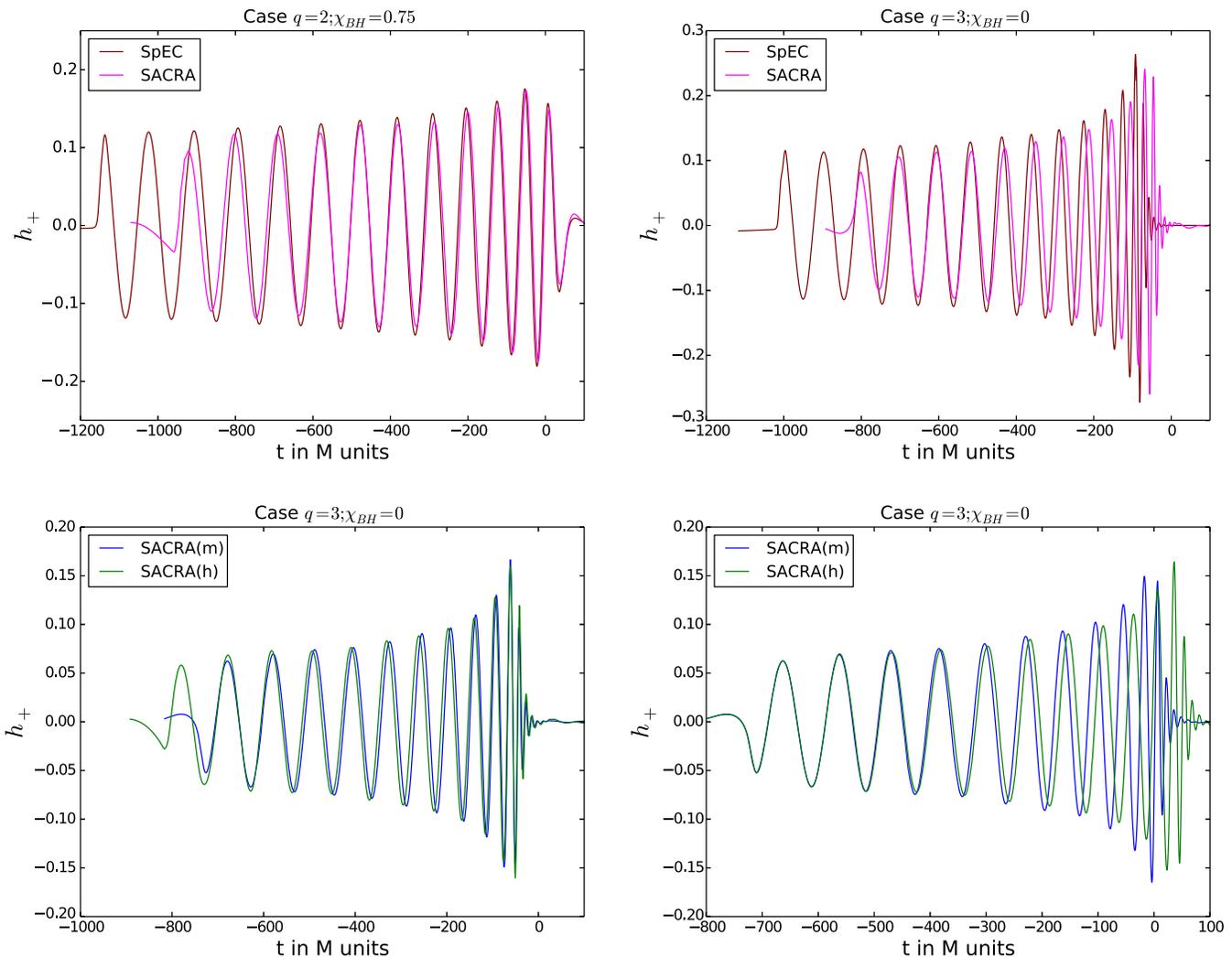


FIG. 3. In each of the two cases in the top panel, we show the relative phasing between SpEC and SACRA NR cycles when their overlap has already been maximized over initial phase and time for the interval where both types of cycles are present. The case on the left shows good overlap  $\sim 95\%$ , while that on the right gives poorer overlap  $\sim 65\%$ . In the left figure of the bottom panel the same maximization is done for two different resolutions of SACRA NR cycles. The resulting overlap between the different resolutions is seen to be  $\sim 97\%$ . In the right figure of that panel no maximization is done but two different resolutions of SACRA NR cycles are shown with their instantaneous starting frequency aligned.

expectations for parameter estimation errors. A longer waveform would also make the first comparison a lot worse than it is—the only reason this time shift works is because the numerical error happens to roughly compensate the change in inspiral rate at a different separation for a large portion of the very short simulation.

## B. Numerical sensitivity: Effect of resolution and the base BBH waveforms used

### 1. Effect of resolution on parameter estimates

One way to test the importance of numerical truncation error is to find the best template matches to numerical waveforms generated from simulations of different resolutions. We compare best matched parameters for each

binary system for which we have multiple numerical resolutions. The results are listed in Table I for best matches to LEA templates and Table II for best matches to LEA + templates.

Several conclusions are apparent. First, the  $\Lambda$  estimates are indeed often sensitive to the NR contribution, as can be seen from the effect of altering resolution. One might imagine that the many PN cycles with the correct tidal contributions would “override” the effect of numerical error in the last few cycles, but this does not turn out to be the case, in general. Including correct PN tidal terms in the inspiral part of the waveform will not guarantee that the  $\Lambda$  estimate remains unaltered with changing resolution. Second, tidal effects are seen to be recovered with somewhat small errors for the low mass case  $q = 1.5$  in LEA.

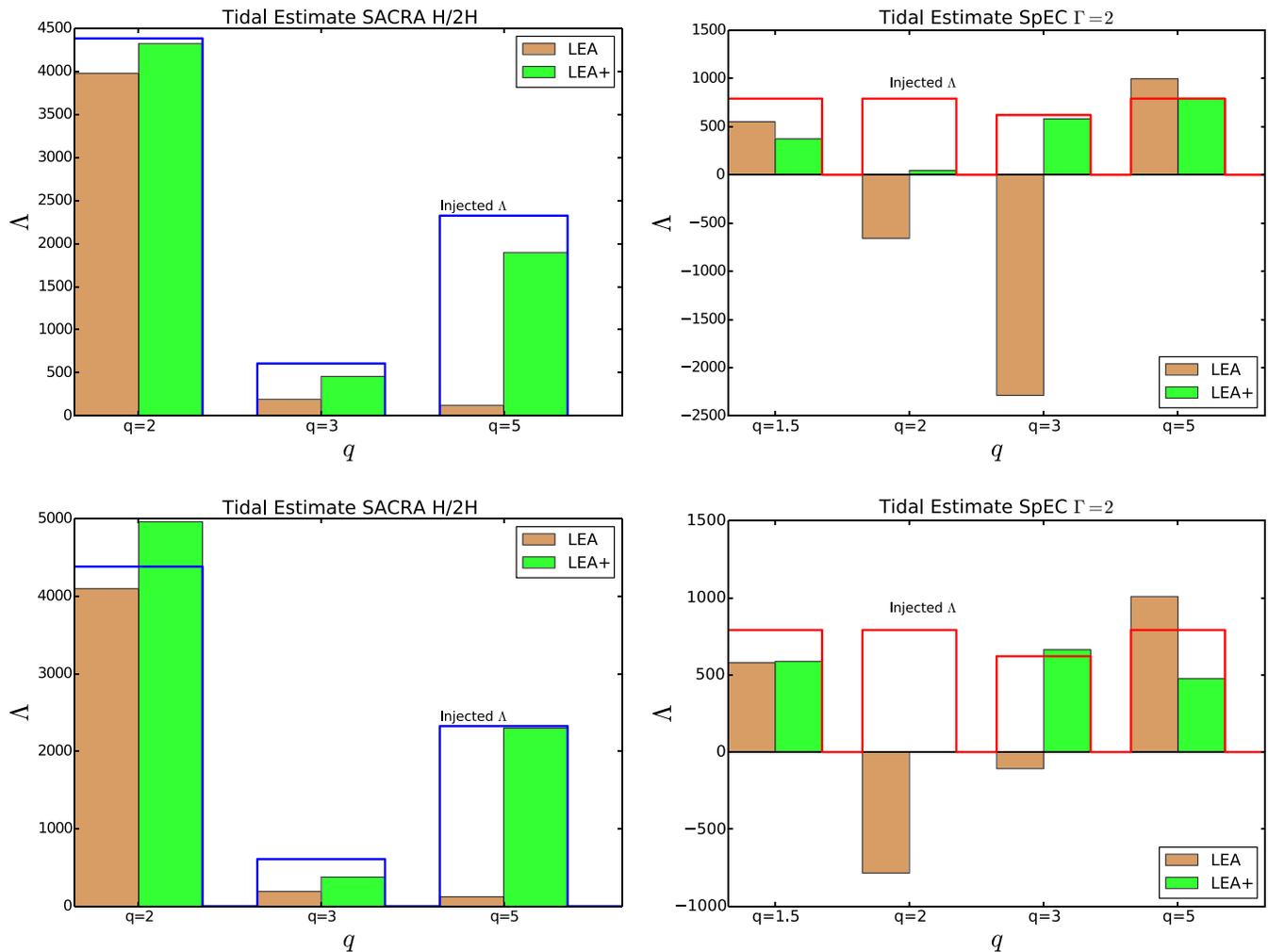


FIG. 4. Estimates of  $\Lambda$  for the highest resolution (top row) and lowest resolution (bottom row) SpEC and SACRA waveforms for the cases listed in Tables I and II. The red and the blue thick lines denote the true (or injected) value of  $\Lambda$  for those two cases, respectively. The brown and green colored bars show the estimated values of  $\Lambda$  while employing LEA and LEA+ templates, respectively.

Tidal effects are also seen to be recovered accurately for high spin systems, for which the NS disrupts well outside the innermost stable circular orbit even for moderate mass ratios, leaving a strong, easily resolved imprint in the waveform. On the other hand, for  $q \geq 3$  and low spin the deviations between resolutions, even for resolutions with lowest phase difference, swamp the measurement of  $\Lambda$ . We see that in many cases, the best-fit template does not even have positive  $\Lambda$ . For  $q = 2$ ,  $\chi_{\text{BH}} = 0$ , numerical accuracy does appear to be sufficient, in that recovered parameters change little with resolution. But here there are other sources of systematic error which again swamp the physical tidal effect and produce negative recovered  $\Lambda$ .

Why does the estimated  $\Lambda$  often vary so widely with resolution? One interpretation would be that the numerical truncation error is swamping the physical tidal effect. However, there is another possibility. If a template family has multiple members with match close to the maximum, small differences between resolutions might cause large

jumps in parameter estimates. We explore this possibility below in the following section.

## 2. Effect of the base BBH waveforms

We performed a detailed comparison of the estimates of tidal and nontidal parameters by LEA and LEA+ (see Tables I and II). We will mainly focus on the overall trend in the estimation of the tidal parameter  $\Lambda$  (see Fig. 4 as well). It is immediately clear that LEA+ does better at returning physically acceptable (non-negative) values of  $\Lambda$ . Given that LEA and LEA+ have similar tidal terms in their waveform models, we reason that the choice of the base BBH models, namely IMRPhenomC and SEOBNRv2, respectively, is likely responsible for most of these differences.

A comparison of Tables I and II indicates that the sensitivity of  $\Lambda$  estimates to numerical resolution is itself PN template-dependent. Notice that the variation in  $\Lambda$  with resolution is much larger for Case 1 when LEA

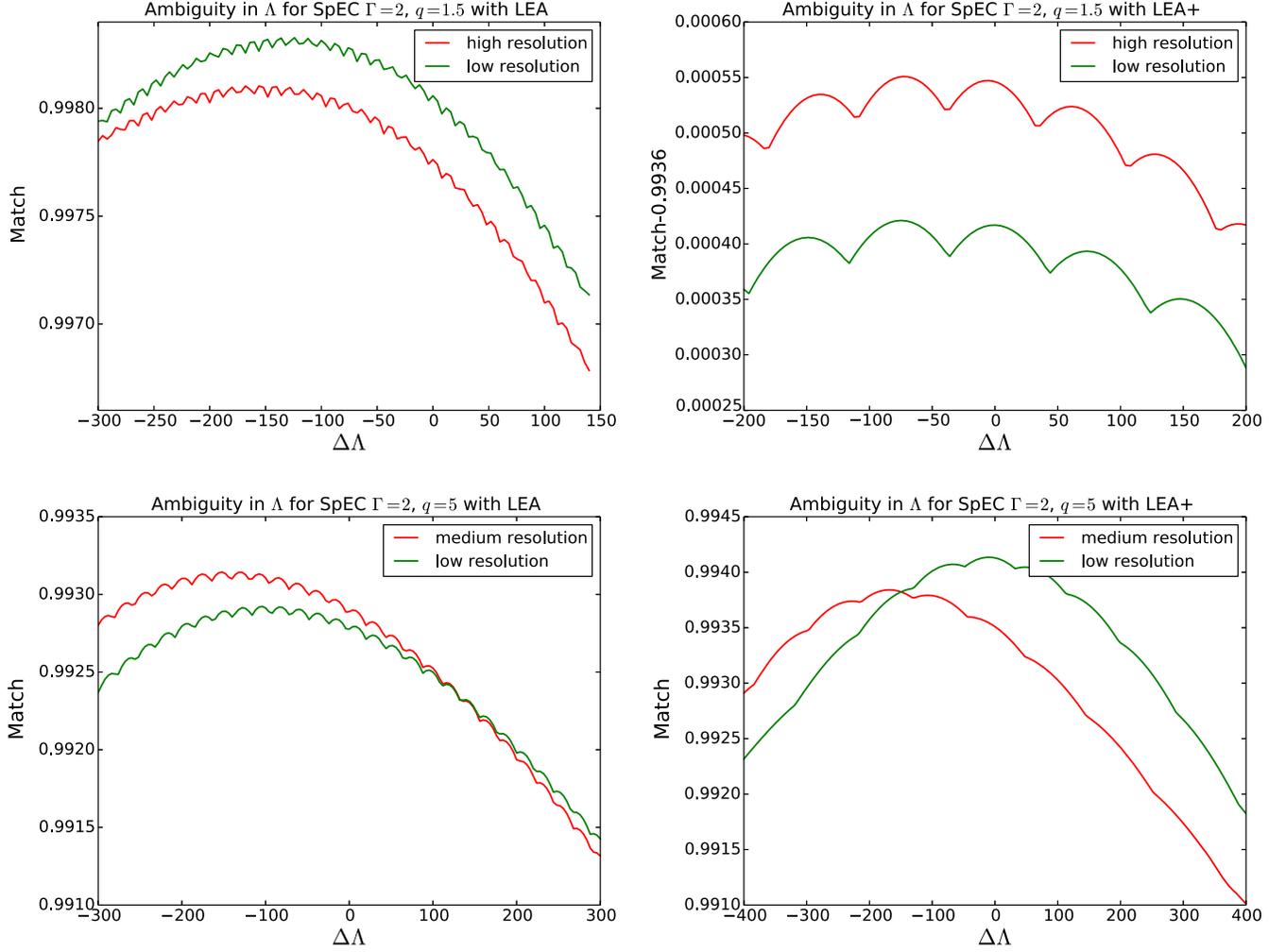


FIG. 5. Constrained ambiguity functions in  $\Lambda$  produced for SpEC hybrids matched with LEA (left panel) and LEA+ (right panel) templates for two different mass ratios, namely,  $q = 1.5$  and  $q = 5$ . Note that  $\Delta\Lambda \equiv \Lambda - \Lambda_{\text{FF}}$ , where  $\Lambda_{\text{FF}}$  is the value for which the match is maximum, and is given in Tables I and II. The sampling rate for computing the above match was chosen to be 4096 Hz. For rates as high as 32 kHz, the small oscillations go away. The concomitant local maxima prevent our fitting factor code from finding the global maximum. Let us illustrate this behavior with the plots in the top panel: The peak in the left figure (high resolution) is at  $\Delta\Lambda \approx -130$ , where  $\Lambda \approx 551.3 - 130 = 421.3$  is what the FF code should have ideally recovered instead of 551.3 if it had not got stuck at a local maximum (see Table I). That would imply a bigger error in  $\Lambda$  estimate of about 47%. (Note that error covariances of  $\Lambda$  with other NSBH parameters can change this estimate somewhat.) The peak in the top right figure (high resolution) is at  $\Delta\Lambda \approx -80$ , where  $\Lambda = 373.5 - 80 = 293.5$  is what the FF code should have ideally recovered (see Table II), once again, ignoring covariances with other NSBH parameters. That would still imply an error in  $\Lambda$  estimate of about 63%. In all of the cases we observe that  $|\Delta\Lambda| \lesssim 200$ .

+ templates are used than when LEA templates are used. For Case 5, the opposite is seen: good consistency in  $\Lambda$  for estimation using LEA+, but wide variation when LEA is used. From the behavior of Case 1 in LEA+, we see that parameters differ most at the highest resolution, a result that seems to contradict the good convergent behavior seen in the numerical waveforms. Extreme sensitivity to the waveform (or the lack thereof) can result from the shape of the match as a function of the deviation in the values of the template parameters from those of the signal.

Indeed, estimating the value of a parameter then involves maximizing the match of normalized templates

with the data containing the signal. In Fig. 5 we show that the rate at which the match changes with a deviation in the value of  $\Lambda$  from that of the signal is slow in the sense that the match drops by a few tenths of a percent even when that deviation is as high as several tens to a few hundreds. This is because the ambiguity function [24] in  $\Lambda$  is diffuse.<sup>1</sup> What is plotted, however, is not the ambiguity function itself since the hybrid waveform and

<sup>1</sup>In this case, the ambiguity function is the match between two unit-norm templates, from the same waveform family, with different values of  $\Lambda$ .

the template are not from the same waveform family. We will call this quantity the *constrained* ambiguity function since the templates, in general, may be constrained to reside in a subspace of the data-space that does not fully overlap with the subspace in which the hybrid waveform resides. Additionally, the same figure shows that this function has multiple local maxima where a parameter estimation algorithm can get stuck and miss finding the global maximum. This effect introduces a fraction of the error in  $\Lambda$  estimation. The small oscillations, accompanied by the local maxima and minima, in the constrained ambiguity function arise when the sampling rate of the data and the templates is not high enough (Fig. 5 used 4096 Hz) and has been studied in a somewhat different context in Ref. [43]. As illustrated in Fig. 5 this effect can contribute to the systematic error in the estimation of  $\Lambda$ , and its extent is somewhat different for LEA and LEA+.

### C. Challenges for hybridization

As we noted before, the regular technique of hybridization is likely to fail when there are not enough NR cycles available in the physical time-frequency “overlap” region where PN waveforms are sufficiently accurate. For NR waveforms available to us, this happens when  $q$  or  $\chi_{\text{BH}}$  are large. For such cases, we use the step method. All such cases in point have been highlighted with an asterisk in Tables I and II.

Since hybridization issues are affected by the number of NR cycles available, pursuing the simulation of sufficiently long NR waveforms is a worthwhile goal for achieving reliable hybrid waveforms. With such NR input, step hybridization will not be necessary at all, because all the waveforms will be guaranteed a relatively large patching region. These waveforms will be critical for producing more accurate NSBH templates, especially, for estimating  $\Lambda$ .

In the same way, having fewer NR cycles implies a smaller patching region thus making hybridization by method of Eqs. (2) and (3) more likely to fail. Once a stable patching region is established involving a sufficiently large number of cycles from accurate overlapping regions of PN and NR waveforms, the tidal estimate is expected to vary minimally with addition of more NR cycles—a fact that has been demonstrated conclusively in [35]. It is also known from the same work that for some of the nontidal parameters their accuracy depends on where in the frequency domain the overlapping region is chosen for hybridizing the waveforms. The important message here is this: given an SNR value, there exists an upper bound on the number of NR cycles necessary for the construction of a hybrid waveform of a desired accuracy. Going above this bound will make the systematic error arising from hybridization subdominant to the statistical error.

Finally, for a few representative cases where both the regular and step methods worked, the estimates of  $\Lambda$  are seen

to agree to within  $\sim 1\%$  of each other. Moreover, to assess how much error our hybridization method itself introduces in the estimation of  $\Lambda$ , we sliced a complete LEA waveform corresponding to Case 2, in such a way that it had at least 30 cycles in the overlapping region, and hybridized the two overlapping pieces using both the regular and the step methods. We then estimated the value of  $\Lambda$  for both hybrids using the fitting factor procedure and found the two estimates to agree within  $\sim 1\%$  of the original unsliced waveform.

### D. Binary parameter estimates

Before concluding, we state the results of the estimates for the nontidal parameters, namely,  $M_{\text{BH}}$ ,  $M_{\text{NS}}$ , and  $\chi_{\text{BH}}$ . The results of the parameter estimation of masses of the  $\Gamma = 2$  SPEC waveforms against different tidal as well as nontidal templates are shown in Fig. 2. For estimates of the masses, the templates show consistent trends. Errors in the masses are small for the low-mass binaries, at a few to several percent, and increase with  $q$  to  $\lesssim 22\%$  for  $q = 5$ . An important contributor to this error is the unfaithfulness of the base waveforms (IMRPhenomC or SEOBNRv2) relative to the NR waveforms on which they are modeled, which itself can cause a bias in the mass parameter of several percent [44]. Differences in the base PN waveform and the hybridization procedure can add to this error as well. Nevertheless, for low mass-ratios it is expected that these biases would be comparable, if not subdominant compared to the other sources of error we have highlighted.

In addition to the estimation of the individual masses, we note in Tables I and II that the error in  $\chi_{\text{BH}}$  can be as high as 14%.

## IV. CONCLUDING REMARKS

In this paper, we classified some of the sources of systematic errors in NSBH tidal waveforms and demonstrated their significant effect on accuracy of NS tidal parameter estimation. Importantly, even though the fitting factor values of the NR-based tidal NSBH templates LEA and LEA+ are very high, their best-matched values of the tidal parameter  $\Lambda$  show a bias of tens to several tens of percent relative to the true value. In addition to the magnitude of the systematic error in  $\Lambda$ , our study allows us to draw several conclusions about the nature of this error.

First, we learn that the final cycles, the portion modeled by numerical relativity, have a significant effect on the systematic error even though they are a small part of the waveform. The modeling of this portion of the evolution is confirmed to be an important endeavor.

Second, we find that the binary BH base of the template family has a large effect on the measured  $\Lambda$ . The IMRPhenomC used by LEA, for example, is insufficiently accurate for mass ratios  $q > 2$ .

Third, use of undersampled data or templates can contribute to systematic error in  $\Lambda$ , as evidenced in the appearance of multiple local maxima in the constrained ambiguity function shown in Fig. 5.

Fourth, there are nevertheless significant errors in some NR waveforms in use. Meaningful consistency between SACRA and SPEC could not be established in all cases. (Errors were too large to definitely establish inconsistency as well.) This comparison should be continued with each group's more-accurate current code. These errors have rather little effect on the match and hence on detectability, but they have a modest effect on parameter estimates.

Finally, the number of cycles in some of the available NR waveforms is inadequate to avoid problems in the hybridization process and that longer NR waveforms can mitigate the greater part of this issue. The limitations of using shorter NR inspirals were systematically studied by McDonald *et al.* [35]. Those limitations are fundamental enough that they extend beyond just BBH waveforms and can impact NSBH template construction as well.

One could, in principle, try to circumvent many of the problems by restricting efforts in the inspiral-only regime. But this quick fix suffers from obvious shortcomings, apart from losing valuable SNR, we also lose valuable information on the EOS, because the EOS effects are strongest just before merger. A further point is that the agreement of the PN and the numerical waveforms is another strong sanity check on the overall consistency of the waveforms. Resorting to PN only techniques will also rob us of this check.

Evidently the important lesson drawn from this exercise is that concerted effort among NR groups is needed to develop accurate NR waveforms with at least a few tens of cycles that can be used to construct longer NSBH waveforms, through hybridization or calibration, for reliable tidal parameter estimation. With this in mind, our recommendations for any possible future attempts on NSBH waveform construction are as follows. First, the focus of templates should crucially involve the low to intermediate

mass ratio  $2 \leq q \leq 5$  because of stronger prevalence of tidal effects. Targeting the construction of accurate waveforms with a wide range of BH spin is also desirable since stellar mass black holes are known to have low to high spins, even though there is no information available yet on how large these spins might be in NSBH systems. A part of this space of physical parameters was already explored in LEA, but with tidal pieces calibrated solely with SACRA waveforms that were old. We propose that until more accurate tidal templates are found, LEA+ should be used for tidal parameter estimation in real-data searches. The gold standard is to have templates that are tidally calibrated across multiple families of high-quality numerical waveforms with consistently high accuracy.

## ACKNOWLEDGMENTS

We thank Bhooshan Gadre and Archisman Ghosh for helpful discussions. We also thank Prayush Kumar for carefully reading the manuscript and making useful suggestions. K. C. and S. B. acknowledge support from the Navajbai Ratan Tata Trust. A. G. acknowledges support from SERB-NPDF Grant No. (PDF/2015/000263), NSF Grants No. AST-1716394 and No. AST-1708146, and the Charles E. Kaufman Foundation of The Pittsburgh Foundation. F. F. gratefully acknowledges support from NASA Grant No. 80NSSC18K0565. K. K. is supported by Japanese Society for the Promotion of Science (JSPS) Kakenhi Grant-in-Aid for Scientific Research (No. JP16H06342, No. JP17H01131, and No. JP18H04595). M. D. acknowledges support through NSF Grant No. PHY-1806207. L. K. acknowledges support from NSF Grant No. PHY-1606654, and M. S. from NSF Grants No. PHY-1708212, No. PHY-1708213, and No. PHY-1404569. L. K. and M. S. also thank the Sherman Fairchild Foundation for their support. S. G. acknowledges NSF Award PHY-1607585 that funded this work. This project has been assigned preprint number LIGO-P1800191.

- 
- [1] J. Abadie *et al.*, *Classical Quantum Gravity* **27**, 173001 (2010).
  - [2] B. P. Abbott *et al.*, *Astrophys. J. Lett.* **832**, L21 (2016).
  - [3] B. D. Lackey, K. Kyutoku, M. Shibata, P. R. Brady, and J. L. Friedman, *Phys. Rev. D* **89**, 043009 (2014).
  - [4] L. Santamaria *et al.*, *Phys. Rev. D* **82**, 064016 (2010).
  - [5] A. Taracchini, Y. Pan, A. Buonanno, E. Barausse, M. Boyle, T. Chu, G. Lovelace, H. P. Pfeiffer, and M. A. Scheel, *Phys. Rev. D* **86**, 024011 (2012).
  - [6] T. Yamamoto, M. Shibata, and K. Taniguchi, *Phys. Rev. D* **78**, 064054 (2008).
  - [7] F. Pannarale, E. Berti, K. Kyutoku, B. D. Lackey, and M. Shibata, *Phys. Rev. D* **92**, 084050 (2015).
  - [8] P. Kumar, M. Pürrer, and H. P. Pfeiffer, *Phys. Rev. D* **95**, 044039 (2017).
  - [9] A. Taracchini, A. Buonanno, Y. Pan, T. Hinderer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, A. H. Mroue, H. P. Pfeiffer, M. A. Scheel, B. Szilagyi, N. W. Taylor, and A. Zenginoglu, *Phys. Rev. D* **89**, 061502 (2014).
  - [10] M. Hannam *et al.*, *Phys. Rev. D* **79**, 084025 (2009).
  - [11] I. Hinder *et al.*, *Classical Quantum Gravity* **31**, 025012 (2014).

- [12] J. Aasi *et al.* (LIGO Scientific, VIRGO, and NINJA-2 Collaborations), *Classical Quantum Gravity* **31**, 115004 (2014).
- [13] L. Baiotti, M. Shibata, and T. Yamamoto, *Phys. Rev. D* **82**, 064015 (2010).
- [14] S. Bernuzzi, M. Thierfelder, and B. Bruegmann, *Phys. Rev. D* **85**, 104030 (2012).
- [15] S. Bernuzzi, A. Nagar, M. Thierfelder, and B. Bruegmann, *Phys. Rev. D* **86**, 044030 (2012).
- [16] J. S. Read, L. Baiotti, J. D. E. Creighton, J. L. Friedman, B. Giacomazzo, K. Kyutoku, C. Markakis, L. Rezzolla, M. Shibata, and K. Taniguchi, *Phys. Rev. D* **88**, 044042 (2013).
- [17] T. Dietrich and T. Hinderer, *Phys. Rev. D* **95**, 124006 (2017).
- [18] K. Hotokezaka, K. Kyutoku, and M. Shibata, *Phys. Rev. D* **87**, 044001 (2013).
- [19] B. D. Lackey, K. Kyutoku, M. Shibata, P. R. Brady, and J. L. Friedman, *Phys. Rev. D* **85**, 044061 (2012).
- [20] F. Pannarale, E. Berti, K. Kyutoku, and M. Shibata, *Phys. Rev. D* **88**, 084011 (2013).
- [21] A. Nagar *et al.*, *Phys. Rev. D* **98**, 104052 (2018).
- [22] L. E. Kidder, M. A. Scheel, S. A. Teukolsky, E. D. Carlson, and G. B. Cook, *Phys. Rev. D* **62**, 084032 (2000).
- [23] J. M. Lattimer, *Annu. Rev. Nucl. Part. Sci.* **62**, 485 (2012).
- [24] S. Dhurandhar, A. Gupta, B. Gadre, and S. Bose, *Phys. Rev. D* **96**, 103018 (2017).
- [25] T. Hinderer, A. Taracchini, F. Foucart, A. Buonanno, J. Steinhoff, M. Duez, L. E. Kidder, H. P. Pfeiffer, M. A. Scheel, B. Szilagyi, K. Hotokezaka, K. Kyutoku, M. Shibata, and C. W. Carpenter, *Phys. Rev. Lett.* **116**, 181101 (2016).
- [26] D. Radice, L. Rezzolla, and F. Galeazzi, *Classical Quantum Gravity* **31**, 075012 (2014).
- [27] S. Bernuzzi and D. Hilditch, *Phys. Rev. D* **81**, 084003 (2010). D. Alic, C. Bona-Casas, C. Bona, L. Rezzolla, and C. Palenzuela, *Phys. Rev. D* **85**, 064040 (2012).
- [28] H. P. Pfeiffer, D. A. Brown, L. E. Kidder, L. Lindblom, G. Lovelace, and M. A. Scheel, *Classical Quantum Gravity* **24**, S59 (2007).
- [29] M. Boyle and A. H. Mroue, *Phys. Rev. D* **80**, 124045 (2009).
- [30] F. Foucart, M. B. Deaton, M. D. Duez, L. E. Kidder, I. MacDonald, C. D. Ott, H. P. Pfeiffer, M. A. Scheel, B. Szilagyi, and S. A. Teukolsky, *Phys. Rev. D* **87**, 084006 (2013).
- [31] K. Kyutoku, M. Shibata, and K. Taniguchi, *Phys. Rev. D* **90**, 064006 (2014).
- [32] F. Foucart, L. Buchman, M. D. Duez, M. Grudich, L. E. Kidder, I. MacDonald, A. Mroue, H. P. Pfeiffer, M. A. Scheel, and B. Szilagyi, *Phys. Rev. D* **88**, 064017 (2013).
- [33] J. M. Lattimer and M. Prakash, *Astrophys. J.* **550**, 426 (2001).
- [34] P. Ajith *et al.*, *Phys. Rev. D* **77**, 104017 (2008).
- [35] I. MacDonald, S. Nissanke, H. P. Pfeiffer, and H. P. Pfeiffer, *Classical Quantum Gravity* **28**, 134002 (2011).
- [36] J. Veitch *et al.*, *Phys. Rev. D* **91**, 042003 (2015).
- [37] J. A. Nelder and R. Mead, *Comput. J.* **7**, 308 (1965).
- [38] Advanced LIGO anticipated sensitivity curves, LIGO Documents, T0900288-v3, <https://dcc.ligo.org/T1800044/public>.
- [39] M. Shibata and T. Nakamura, *Phys. Rev. D* **52**, 5428 (1995).
- [40] T. W. Baumgarte and S. L. Shapiro, *Phys. Rev. D* **59**, 024007 (1998).
- [41] L. Lindblom, M. A. Scheel, L. E. Kidder, R. Owen, and O. Rinne, *Classical Quantum Gravity* **23**, S447 (2006).
- [42] B. P. Abbott *et al.*, *Phys. Rev. Lett.* **119**, 161101 (2017).
- [43] P. Ajith *et al.*, *Classical Quantum Gravity* **29**, 124001 (2012); **30**, 199401(E) (2013).
- [44] P. Kumar, T. Chu, H. Fong, H. P. Pfeiffer, M. Boyle, D. A. Hemberger, L. E. Kidder, M. A. Scheel, and B. Szilagyi, *Phys. Rev. D* **93**, 104050 (2016).