High-accuracy waveforms for black hole-neutron star systems with spinning black holes

Francois Foucart,[1] Alexander Chernoglazov,[1] Michael Boyle,[2] Tanja Hinderer,[3, 4, 5]
Max Miller,[1] Jordan Moxon,[6] Mark A. Scheel,[6] Nils Deppe,[6] Matthew D. Duez,[7]
Francois Hébert,[6] Lawrence E. Kidder,[2] William Throwe,[2] and Harald P. Pfeiffer[8]

[1]*Department of Physics, University of New Hampshire, 9 Library Way, Durham NH 03824, USA*
[2]*Cornell Center for Astrophysics and Planetary Science, Cornell University, Ithaca, New York, 14853, USA*
[3]*GRAPPA Institute of High-Energy Physics, University of Amsterdam,*
Science Park 904, 1098 XH Amsterdam, Netherlands
[4]*Delta Institute for Theoretical Physics, Science Park 904, 1090 GL Amsterdam, The Netherlands*
[5]*Institute for Theoretical Physics, Utrecht University,*
Princetonplein 5, 3584 CC Utrecht, The Netherlands
[6]*TAPIR, Walter Burke Institute for Theoretical Physics, MC 350-17,*
California Institute of Technology, Pasadena, California 91125, USA
[7]*Department of Physics & Astronomy, Washington State University, Pullman, Washington 99164, USA*
[8]*Max-Planck-Institut fur Gravitationsphysik, Albert-Einstein-Institut, D-14476 Potsdam, Germany*

The availability of accurate numerical waveforms is an important requirement for the creation and calibration of reliable waveform models for gravitational wave astrophysics. For black hole-neutron star binaries, very few accurate waveforms are however publicly available. Most recent models are calibrated to a large number of older simulations with good parameter space coverage for low-spin non-precessing binaries but limited accuracy, and a much smaller number of longer, more recent simulations limited to non-spinning black holes. In this paper, we present long, accurate numerical waveforms for three new systems that include rapidly spinning black holes, and one precessing configuration. We study in detail the accuracy of the simulations, and in particular perform for the first time in the context of BHNS binaries a detailed comparison of waveform extrapolation methods to the results of Cauchy Characteristic Extraction. The new waveforms have $< 0.1\,\mathrm{rad}$ phase errors during inspiral, rising to $\sim (0.2 - 0.4)\,\mathrm{rad}$ errors at merger, and $\lesssim 1\%$ error in their amplitude. We compute the faithfulness of recent analytical models to these numerical results, and find that models specifically designed for BHNS binaries perform well ($F > 0.99$) for binaries seen face-on. For edge-on observations, particularly for precessing systems, disagreements between models and simulations increase, and models that include precession and/or higher-order modes start to perform better than BHNS models that currently lack these features.

I. INTRODUCTION

Over the last 5 years, our ability to observe gravitational waves from merging compact objects has grown at an impressive rate. The first observation of two merging black holes (GW150914) [1] was followed by nine more black hole mergers during the "O1" and "O2" observing runs of the LIGO-Virgo Collaboration (LVC) [2], as well as the first detection of a binary neutron star merger (GW170817) [3]. The O3 observing run is still being analyzed by the LVC, but public alerts indicate that dozens of additional mergers have been observed.[1]

The third type of binary merger likely to be detected by current observatories, black hole-neutron star (BHNS) mergers, remains the most elusive. At the time of this writing, five O3 alerts are classified as likely BHNS mergers. However, the robust classification of a specific event as a BHNS binary remains a difficult task, due to uncertainties about the mass range of black holes and neutron stars, and the inability of current observatories to directly demonstrate the presence of a neutron star in the absence of an electromagnetic counterpart to the gravitational wave signal. The classification of an event as a BHNS merger in public alerts only indicates the likely presence of an object of mass $M < 3M_\odot$, that could also be a low-mass black hole. Most notably, one of the five "BHNS" alerts is the now published GW190814 event [4]. In that system, the lower mass object has a mass $M \in [2.51, 2.67]M_\odot$ (at 90% confidence), and could be either the lowest mass black hole or the highest mass neutron star observed to date. On the other hand, the second likely binary neutron star system published by the LVC (GW190425) [5] has a most massive object with mass $M \in [1.61, 2.52]$ (allowing for non-negligible spins), and could potentially be a BHNS merger if $(2 - 3)M_\odot$ black holes exist. It is quite likely that we will only be sure of the detection of a BHNS binary when we observe either a system where both masses provide unambiguous information about the nature of the merging objects (a statement that clearly depends on one's priors for the possible distribution of black hole and neutron star masses), or when we observe an electromagnetic counterpart to a GW event for which the most massive object is guaranteed to be a black hole (e.g. with $M \gtrsim 3M_\odot$).

Nearly all GW observations performed so far relied on

---

[1] See e.g. https://gracedb.ligo.org/superevents/public/O3/

the availability of accurate signal template banks.[2] This may lead to some complications and observation biases when observing BHNS binaries. Indeed, mixed binaries are likely to have larger mass asymmetries and to exhibit more orbital plane and spin precession than black hole binaries and neutron star binaries. The use of aligned-spin templates in detection pipelines can then lead to the loss of a significant fraction of events [6]. GW templates including precession, tidal effects, and the potential disruption of the neutron star may help alleviate these issues, and will certainly be valuable to perform parameter estimation. Existing models for the GW signals emitted by BHNS binaries, however, have only recently begun to include both phase and amplitude corrections associated with tidal distortion and with the disruption of the neutron star (see e.g. [7–9]), and do not so far account for precession.

This is where numerical simulations play an important role: analytical models are tested and calibrated on numerical simulations, to make sure that the models properly capture the late-time inspiral and non-linear merger phase. Numerical waveforms hybridized with analytical models at early times can also be injected into parameter estimation pipelines to estimate model biases (see e.g. [10, 11] for BHNS mergers). There is, however, a limited number of available numerical waveforms to perform these tests. The 134 SACRA simulations used by Lackey *et al* [7] still provide the most extensive parameter space coverage of BHNS mergers, yet these waveforms are now quite old. The limited length and accuracy of the simulations make them most useful to calibrate amplitude corrections at the time of merger, but not as useful to the modeling of the GW phase. Additionally, these waveforms are limited to aligned BH spins with dimensionless spins $\chi < 0.75$. We have recently published a much smaller set of 5 longer, more accurate SpEC simulations [12], publicly available as part of the SXS catalogue,[3] but these simulations are limited to non-spinning black holes and obviously do not come close to the SACRA waveforms in term of parameter space coverage.[4] Two of the five have very high neutron star spins, to help efforts to model dynamical tides in BHNS mergers [14], but are otherwise less useful to calibrate models within the most likely range of BHNS parameters.

In this paper, we present a new set of 3 BHNS binaries [15–17] performed with the SpEC code. These simulations complement our existing set of long, accurate simulations. All of these systems have more realistic mass ratios than most of our public waveforms ($Q = 3, 4$). Two have higher BH spins than existing public wave-forms ($\chi_{\mathrm{BH}} = 0.9$), aligned with the orbital angular momentum of the binary. The third has a significant BH spin ($\chi_{\mathrm{BH}} = 0.75$) misaligned with the orbital angular momentum by $45°$, leading to significant precession of the orbital plane. All three waveforms are long by the standard of BHNS simulations ($26 - 33$ cycles). They also have accuracy comparable to the BHNS waveforms in our current catalogue, despite the use of higher mass ratios and higher black hole spins. In fact, the numerical accuracy of these waveforms is high enough that we have to more carefully analyze the uncertainty associated with the extrapolation of the GW signal to null infinity. We thus perform a detailed study of waveform extrapolation errors, and compare for the first time the results of waveform extrapolation to the waveforms obtained using Cauchy Characteristic Extraction (CCE) methods.

We describe our numerical methods in Sec. II, the resulting waveforms in Sec. III A, and our numerical accuracy in Sec. III B. A comparison between waveform extrapolation and CCE is provided in Sec. III C. In the rest of this paper, we use units such that $G = c = 1$, and define $M_{\mathrm{BH}}, M_{\mathrm{NS}}$ as the ADM masses of the black hole and neutron star at infinite separation, $M = M_{\mathrm{BH}} + M_{\mathrm{NS}}$ as the total mass of the system, and χ_{BH} as the dimensionless black hole spin. All neutron stars in our simulations are initially non-spinning.

II. METHODS

A. Evolution methods

The simulations presented here are performed with the SpEC numerical relativity code.[5] SpEC evolves Einstein's equations in the Generalized Harmonic formalism [18] on a pseudospectral grid. The grid rotates and contracts to follow the evolution of the binary, and is distorted so that the apparent horizon of the black hole remains nearly spherical [19]. A sphere of constant grid-frame radius is excised from the grid to avoid evolving the interior of the black hole. The general relativistic equations of hydrodynamics are solved on a separate cartesian grid [20, 21]. Our latest algorithm follows the prescriptions of Radice *et al* [22] to obtain high-order convergence in smooth regions while capturing shocks.

The neutron star matter is described by a $\Gamma = 2$ ideal gas equation of state, with an ad-hoc thermal component: $P = 101.45\rho_0^{\Gamma} + \Gamma\rho_0 T$, with P the pressure and ρ_0 the baryon density. We choose the central density of the neutron star to get a small but reasonable compactness $C_{\mathrm{NS}} = GM_{\mathrm{NS}}/(R_{\mathrm{NS}}c^2) = 0.144$ for $M_{\mathrm{NS}} = 1.4 M_{\odot}$. The use of such a simple equation of state has a few advantages, including lower simulation costs and higher numerical accuracy than for more realistic models, and

the possibility to rescale the result of the simulations with the total mass of the system. Hence, in this paper, we typically report all masses, times, and distances as dimensionless numbers. Its main disadvantage is that while the dimensionless tidal deformability of the neutron star is reasonable ($\Lambda = 791$, around the upper bound allowed by current observations for low-mass neutron stars), the internal structure of the star and the inferred mass-radius relationship are not. This is less of an issue for BHNS systems than for binary neutron star systems (as the latter require us to construct two neutron stars, potentially of different masses, with physically consistent tidal properties), but would certainly be a major limitation if we wanted to study the formation of a post-merger accretion disk, or any microphysics. For GW modeling, the tidal deformability has been shown to be the main parameter setting the properties of GW signals both before and during merger [7]. Further tests of the impact of the equations of state beyond the dimensionless tidal deformability would however be desirable in the future.

We use a third-order Runge-Kutta method for the time evolution. At the end of each time step, the metric and its derivatives are interpolated from the pseudospectral grid to the finite difference grid, while the fluid variables are interpolated from the finite difference grid to the pseudospectral grid. At other times (e.g. at intermediate steps of the Runge-Kutta algorithm, or when evaluating variables using dense output), variables evolved on another grid are evaluated using linear extrapolation, using their last two communicated values. Overall, the simulations presented here use the exact same numerical methods as the simulations published in Foucart *et al* [12].

B. Grid structure

The finite difference grid used in our simulations has a relatively simple structure. Before merger, we use a Cartesian grid with constant grid spacing $\Delta x_{\rm grid}$ in the grid coordinates. As the grid contracts during the binary inspiral, this would lead to a significant decrease in the grid spacing in the inertial frame, to $0.4\Delta x_{\rm grid}$ by the time of merger, dramatically increasing the cost of the simulations. Instead, whenever the grid spacing decreases by 20% in the inertial frame, we reset it to the grid spacing at the initial time, interpolating onto a coarser finite difference grid. This operation is performed $\sim 4-5$ times per simulation, and keeps the resolution roughly constant during the evolution. Once the neutron star disrupts, we use fixed mesh refinement. The finest level of refinement is a grid of 324^3 cells, centered on the black hole, and with the same spacing as the pre-merger grid in the inertial frame. Additional levels of refinement are added as needed, each new level a 324^3 grid centered on the black hole and twice the grid spacing of the previous level. We note that to save computational resources, these grids are divided into 27^3 cells blocks that can be fully ignored by the evolution if no matter

is present in the region that they cover. After the gravitational wave signal from the merger has left the finite difference grid, we save computational resources by reducing the resolution of that grid: by that point, following the gravitational waves to the radius at which they are extracted is our only concern, and this only requires evolution of Einstein's equations. We evolve each configuration at 3 resolutions, summarized in Table I. If we assume $M_{\rm NS} = 1.4M_\odot$, these correspond to initial grid spacings $\Delta x_{\rm FD}^0 = (295, 236, 189)\,$m.

Before merger, the pseudospectral grid is constructed from 8 spherical shells surrounding the black hole, 1 ball and 8 spherical shells covering the neutron star and its surroundings, and 32 spherical shells covering the wave region (centered on the center of mass of the binary, and with radii ranging from 2.5 times the binary separation to $R_{\rm out} = 500M$). Between these 3 regions, we use distorted cylinders, with the line connecting the compact objects as their axis. After merger, the region inside of the 32 outer shells is covered with "CubedSphere" subdomains, i.e. cubes distorted so that one coordinate is constant at constant radius (defined as the distance to the center of the remnant black hole). The number of basis functions within each of these subdomains is chosen adaptively, to obtain a user-specified accuracy estimated from the spectral coefficients of the evolved variables [21, 23]. The target accuracy on the pseudospectral grid scales as $(\Delta x_{\rm FD}^0)^5$. Errors on the pseudospectral grid should thus converge to zero faster than the errors on the finite difference grid.

Visualizations of the matter density, black hole apparent horizon, finite difference grid, and pseudospectral grid are provided in Fig. 1 for simulation Q4S9 (see next section), around the time of merger. The figure illustrates the adaptivity of both numerical grids.

C. Initial Conditions

Initial conditions for all simulations are obtained using our in-house Spells initial data solver [24–26]. Spells solves for the constraints in Einstein's equations and for an irrotational velocity profile inside the neutron star, while imposing hydrostatic equilibrium. We first find initial data on a quasi-circular trajectory [26, 27], then reduce the eccentricity according to the iterative procedure described in Pfeiffer *et al* [28]. Each iteration requires the evolution of the binary for ~ 3 orbits. We target eccentricities of $\lesssim 0.002$ for non-precessing systems. For the precessing system in this paper, the eccentricity of the quasi-circular initial data was already very small ($e \sim 0.003$), and could not be reduced using our standard procedure.

We consider 3 initial configurations, summarized in Table I. Figure 2 also provides an overview of these simulations and of existing public BHNS waveforms. The first configuration is a system with mass ratio $Q = 3$ and aligned BH spin $\chi_{\rm BH} = 0.9$, hereafter named Q3S9. The
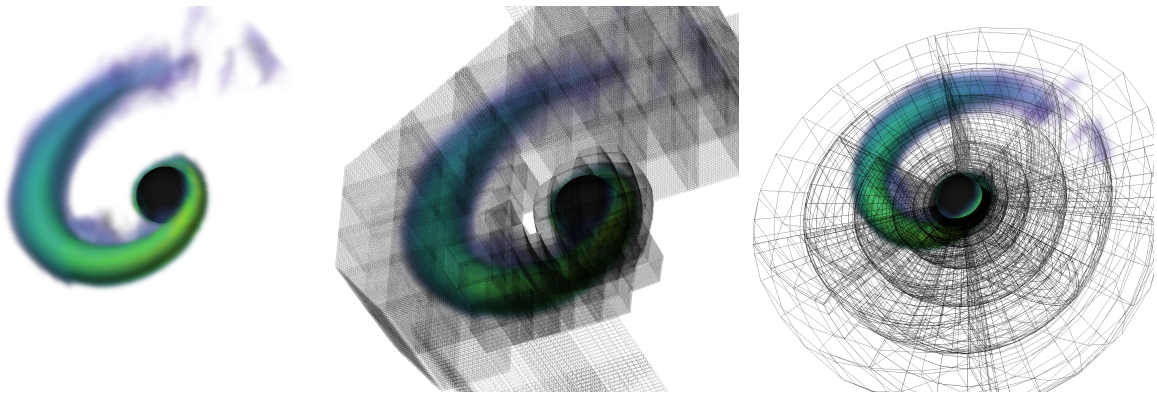
FIG. 1. Simulation Q4S9 at a time when $\sim 30\%$ of the neutron star mass has been accreted by the black hole. *Left*: Matter with density above $6 \times 10^7 \, \mathrm{g/cm^3}$, and apparent horizon of the black hole. *Center*: Finite difference grid, showing both mesh refinement and the fact that the grid does not cover vacuum regions. *Right*: Inner region of the pseudospectral grid, showing mesh refinement close to the black hole and in regions where dense matter is present. We only show points below the orbital plane and within $\sim 15M$ of the black hole's center. The pseudospectral grid extends to much larger distances $(500M)$, using spherical shells not shown on the figure. All three figures are taken from our lowest resolution simulation.
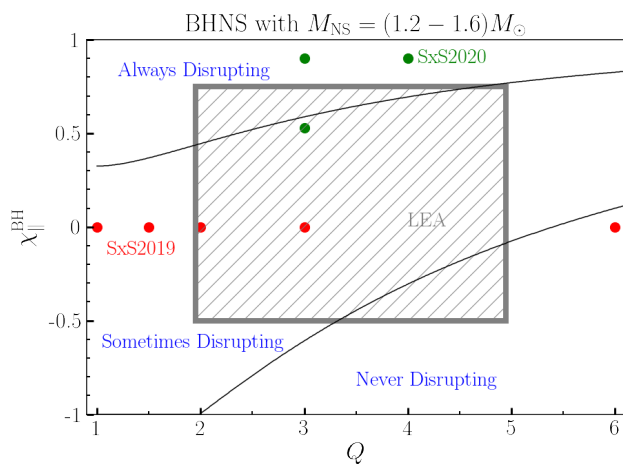

FIG. 2. Distribution of the BHNS simulations in the SXS catalogue projected in the Q-$\chi_\parallel^{\mathrm{BH}}$ plane. Simulations from [12] are red dots (SXS2019), while simulations from this work are green dots (SXS2020). The region of parameter space covered by 134 short waveforms from [7] is shown in grey (LEA). For context, we also show the regions of parameter space where $(1.2 - 1.6)M_\odot$ neutron stars satisfying the equations of state constraints of [29] always disrupt, disrupt for some equations of state and/or neutron star masses only, or never disrupt, according to the fitting formula from [30].

second is identical except for the choice of a mass ratio $Q = 4$, and is named Q4S9. The last configuration has a dimensionless BH spin $\chi_{\mathrm{BH}} = 0.75$, misaligned by $45°$ with the orbital angular momentum, and initially in the plane formed by the orbital angular momentum vector and the line connecting the center of the compact objects. The misalignment of the spin leads to significant precession of the orbital plane of the binary. We label this simulation Q3S75p. The initial orbital frequencies are chosen to provide more than ~ 12 orbits before merger,

TABLE I. Overview of the simulations presented in this paper. $M_{\mathrm{BH,NS}}$ are the ADM masses of the BH and NS in isolation, χ_{BH} the initial dimensionless BH spin, e the initial eccentricity, i_{BH} the initial inclination of the BH spin with respect to the orbital angular momentum vector, $\tilde\Lambda$ the effective dimensionless tidal deformability of the binary, Ω_0 the initial angular velocity, $M = M_{\mathrm{BH}} + M_{\mathrm{NS}}$ the total mass, t_{peak} the time at which the $(2,2)$ mode of the GW signal reaches its maximum amplitude, and $\Delta x_{\mathrm{FD}}^{t=0}$ the initial spacing of the finite volume grid.

Name	$\frac{M_{\mathrm{BH}}}{M_{\mathrm{NS}}}$	χ_{BH}	e	i_{BH}	$\tilde\Lambda$	$\Omega_0 M$	$\frac{t_{\mathrm{peak}}}{M}$	$\frac{\Delta x_{\mathrm{FD}}^{t=0}}{M_{\mathrm{NS}}}$
Q3S9-L0	3	0.9	0.0004	$0°$	35.2	0.0236	2342.6	0.143
Q3S9-L1	3	0.9	0.0005	$0°$	35.2	0.0236	2346.8	0.114
Q3S9-L2	3	0.9	0.0005	$0°$	35.2	0.0236	2346.5	0.091
Q4S9-L0	4	0.9	0.0017	$0°$	15.3	0.0243	2661.7	0.143
Q4S9-L1	4	0.9	0.0018	$0°$	15.3	0.0243	2660.1	0.114
Q4S9-L2	4	0.9	0.0017	$0°$	15.3	0.0243	2660.2	0.091
Q3S75p-L0	3	0.75	0.0031	$45°$	35.2	0.0197	3473.5	0.143
Q3S75p-L1	3	0.75	0.0031	$45°$	35.2	0.0197	3469.9	0.114
Q3S75p-L2	3	0.75	0.0031	$45°$	35.2	0.0197	3472.3	0.091

and we find indeed that simulation Q3S9 evolves for 13.2 orbits before the peak of the dominant $(2,2)$ mode of the GW signal, Q4S9 for 15.6 orbits, and Q3S75p for 16.3 orbits. By this metric, these 3 simulations are longer than all but one of the existing public BHNS waveforms.[6] [12] In terms of the number of time steps required, which may be more relevant to the growth of numerical errors, the simulations presented here are significantly longer than any public BHNS waveform.

As shown on Fig. 2, the initial conditions for our simulations are all in the regime where the neutron star is

————

[6] The longest public BHNS waveform is a non-spinning, $Q = 1.5$ simulation evolved for 16.6 orbits

strongly disrupted by the black hole – and would be disrupted even for softer equations of state. The two high-spin simulations are also out of the range of the SACRA simulations used by Lackey *et al* (LEA) [7], which makes them useful to test models of both the phase evolution and tidal disruption of BHNS mergers. Our lower spin waveform, on the other hand, has the advantage of being the only precessing system shown on Fig. 2. Less effort has gone into the production of waveforms for non-disrupting binaries, in part because our best simulation in that regime [31], the $Q = 6$ simulation on Fig. 2, showed that the resulting waveform could not be distinguished from an equivalent BBH waveform, at least within our numerical errors (or with any existing GW detector).

D. Waveform extraction

We use two independent methods to estimate the gravitational wave signal at infinity from the values of the metric at finite radii. The first, used in all of our previous BHNS publications, follows the procedure outlined by Boyle & Mroue [32]. The Newman-Penrose scalar Ψ_4 and metric perturbation h are estimated on spheres of constant inertial radii R_i (the latter using Regge-Wheeler-Zerilli techniques), and decomposed into spin$=-2$ spherical harmonics components. For each R_i, we then compute a retarded time $t_{\rm ret}(t, R_i)$ approximately accounting for the travel time of the wave from the merging compact objects to R_i. We then fit the ansatz

$$A_{lm}(t_{\rm ret}, r) = \sum_{j=0}^{N} A_{lm,j}(t_{\rm ret}) r^{-j} \tag{1}$$

$$\phi_{lm}(t_{\rm ret}, r) = \sum_{j=0}^{N} \phi_{lm,j}(t_{\rm ret}) r^{-j} \tag{2}$$

to the amplitude A_{lm} and phase ϕ_{lm} of the (l, m) component of our spherical harmonics decomposition, at a fixed set of retarded times. The (l, m) mode at infinity is then estimated to be $A_{lm,0} e^{i\phi_{lm,0}}$. This procedure can be applied to either Ψ_4 or h. In our simulations, we fit this ansatz to the estimated values of h, Ψ_4 at 20 radii between $100M$ and $400M$, equally spaced in r^{-1}.

The second method is used here for the first time in our fluid simulations. We perform Cauchy Characteristic Evolution (CCE) [33–37] using the methods described in Moxon *et al* [38], and implemented in the open-source SpECTRE code.[7] In CCE, we use the evolution data on a surface of constant inertial radius ($R = 200M, 300M, 400M$ here) as boundary condition for a non-linear evolution on a null foliation of the spacetime outside of that surface, propagating the signal to null infinity. CCE carries to null infinity the Bondi news,

and the SpECTRE implementation of CCE also provides estimates of all five Weyl scalars and of the gravitational wave strain – although with some caveats related to initial data and Bondi-Metzner-Sachs (BMS) freedom for the latter, discussed in Sec. III C.

With these two methods at our disposal, we can more carefully study the reliability of our estimates of the signal at null infinity.

The waveforms publicly released as part of the SXS catalogue include two versions of the extrapolated signal. The first contains the waveform in the inertial frame of the simulation. The second corrects for the motion of the center-of-mass of the binary, and provides the waveform in the rest frame of the system (before merger) [39, 40]. This procedure avoids some mode mixing in all cases, and is particularly useful for the precessing system: precessing BHNS binaries generated with our initial data solver have non-zero velocity in the direction perpendicular to the initial orbital plane ($\sim 0.001c$) that can lead to significant phase errors at merger ($\sim 0.3\,\rm rad$) simply due to the change in the time-of-flight of the waveforms from the binary to the observer. Comparisons between numerical results and analytical models presented in this paper are performed after removal of the center-of-mass motion.

III. RESULTS

A. Overview

The evolution of all three configurations proceed as is typical for BHNS systems with moderate mass ratios $Q \sim 2-4$ and significant spins: the neutron star disrupts well out of the innermost stable circular orbit of the black hole, leading to mass ejection and the formation of a massive accretion disk (see Figs. 1-2). In these simulations, however, we do not attempt to follow the evolution of the post-merger remnant with enough accuracy to properly measure the masses of the ejecta and post-merger accretion disks, in part due to the high computational cost of following rapid accretion by the black hole in BHNS systems, and in part because the post-merger evolution likely lacks realism when using a simple $\Gamma = 2$ ideal gas equation of state. Once the neutron star disrupts, stopping GW emission, we focus on following the propagation of the GW signal to large distances.

The extrapolated GW signals obtained from our simulations are publicly available as part of the SXS catalogue, for extrapolation orders $N = 2 - 5$ and for multipoles up to $l = 8$. Figure 3 shows the h_+ and $h_\times$ polarizations of the GW signals, for observers who are, at $t = 0$, in the direction of the total angular momentum of the binary, or perpendicular to that direction and in the orbital plane. The clearest difference between these BHNS waveforms and equivalent BBH waveforms is the rapid cutoff in GW emission when the neutron star disrupts. The other features of the signal are similar to BBH systems: the signal in the direction of the total angular

_____

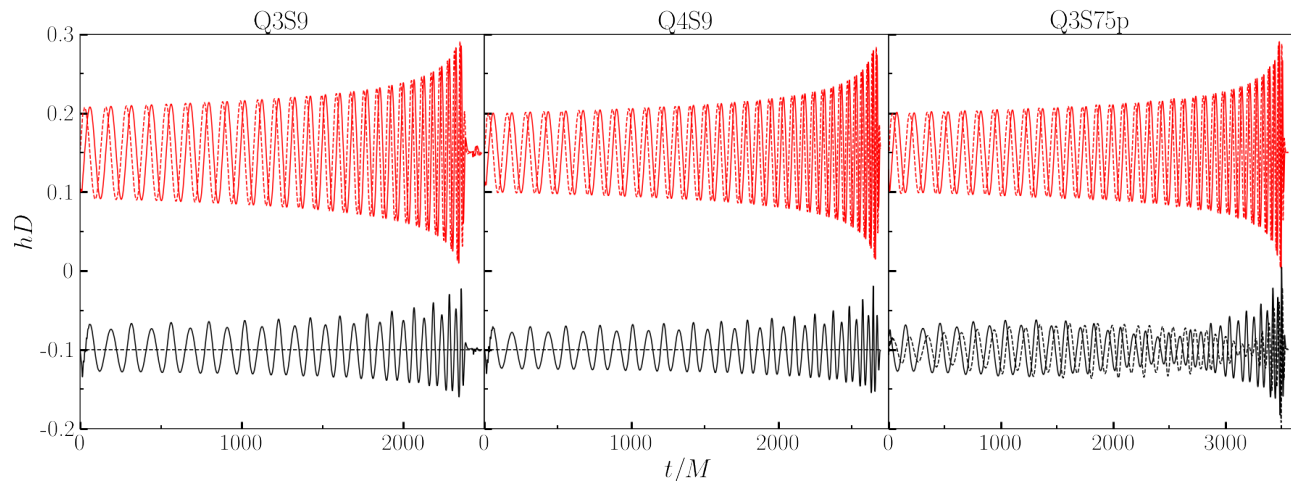[7] https://github.com/sxs-collaboration/spectre

FIG. 3. GW strain for simulations Q3S9 (left), Q4S9 (center) and Q3S75p (right). Each plot shows results for an observer in the direction of the total angular momentum (top, 'face-on' for non-precessing system) and from a direction that, at $t = 0$, is orthogonal to the total angular momentum and within the orbital plane of the binary (bottom, 'edge-on' at $t = 0$). We show both polarizations h_+ (solid lines) and $h_\times$ (dashed lines), and shift the average value of the strain for readability. The edge-on signals clearly show the impact of mass asymmetry (oscillations in the maximum of the strain, due to a significant $l = 3, m = 3$ mode) and, for Q3S75p, precession (slow oscillation in the amplitude of the signal, non-zero value of $h_\times$). Q3S75p goes through slightly more than half a precession cycle over the course of the simulation.

momentum of the system is entirely dominated by the $l = 2, m = \pm 2$ modes and is thus a nearly feature-less chirping signal. For non-precessing systems, the edge-on signal shows more clearly the impact of unequal masses. For Q3S75p, in addition to the mass asymmetry, the precession of the orbital plane is clearly visible for the initially 'edge-on' observer. The $h_\times$ signal, in particular, clearly shows that the Q3S75p system goes through slightly more than half a precession cycle during the simulation.

An important aspect of the BHNS binary systems presented here, as opposed to those previously published in the SXS catalogue [12], is the impact of subdominant modes on the signal. For the Q3S9 (resp. Q4S9) configuration, the peak amplitude of the $(3, 3)$ mode of the strain is 18% (resp. 21%) of the peak amplitude of the dominant $(2, 2)$ mode. The $(4, 4)$ mode has 7% (8%) of the peak amplitude of the dominant mode, while other modes remain below 5% of the amplitude of the $(2, 2)$ mode. For Q3S75p, the precession of the orbital plane additionally leads to the mixing of modes with the same l but different m. All $l = 2$ and $l = 3$ modes then have significant amplitude. Accordingly, these new public waveforms should be particularly useful to test the effect of higher-order modes on detection and parameter estimation for BHNS binaries.

B.   Numerical accuracy

The main expected uses of our waveforms are testing and calibrating analytical waveform models, and injection in detection and parameter estimation pipelines used

by GW observers. It is thus critical to provide careful, conservative estimates of our errors, to avoid introducing systematic biases in these studies, and in upcoming observations of BHNS binaries. In Foucart et al [12], we proposed a standardized method to estimate phase errors in SpEC BHNS waveforms that accounts for 3 main sources of error: finite resolution, extrapolation of the signal to infinity, and mass loss at the boundary of the finite difference grid used to evolve neutron stars. We summarize this method here, and plot the resulting error estimates on Fig. 4, for the dominant $(2, 2)$ mode of the signal.

Finite-resolution errors are estimated by comparing the results of our low (L0), medium (L1), and high (L2) accuracy simulations. We first measure the phase differences between L0 and L2, and estimate the error in the L2 waveform using Richardson extrapolation to infinite resolution. The extrapolation is performed assuming 2nd-order convergence, a relatively conservative estimate considering that the methods used in our simulations are typically at least 3rd order convergent. We then perform the same calculation, but using the results of the L1 and L2 simulations. At any time, the worst of these two error estimates is assumed to be the finite-resolution error. We note that we need this comparison because in the hybrid spectral/finite volume algorithm used in SpEC, different parts of the evolution may dominate the error budget at different times. Unfortunately, as a result the sign of the phase difference between two simulations may change over the course of the evolution, leading to occasional cancellations of the error estimates based on 2 resolutions only. On the other hand, we have found that the more complex error estimate described here has
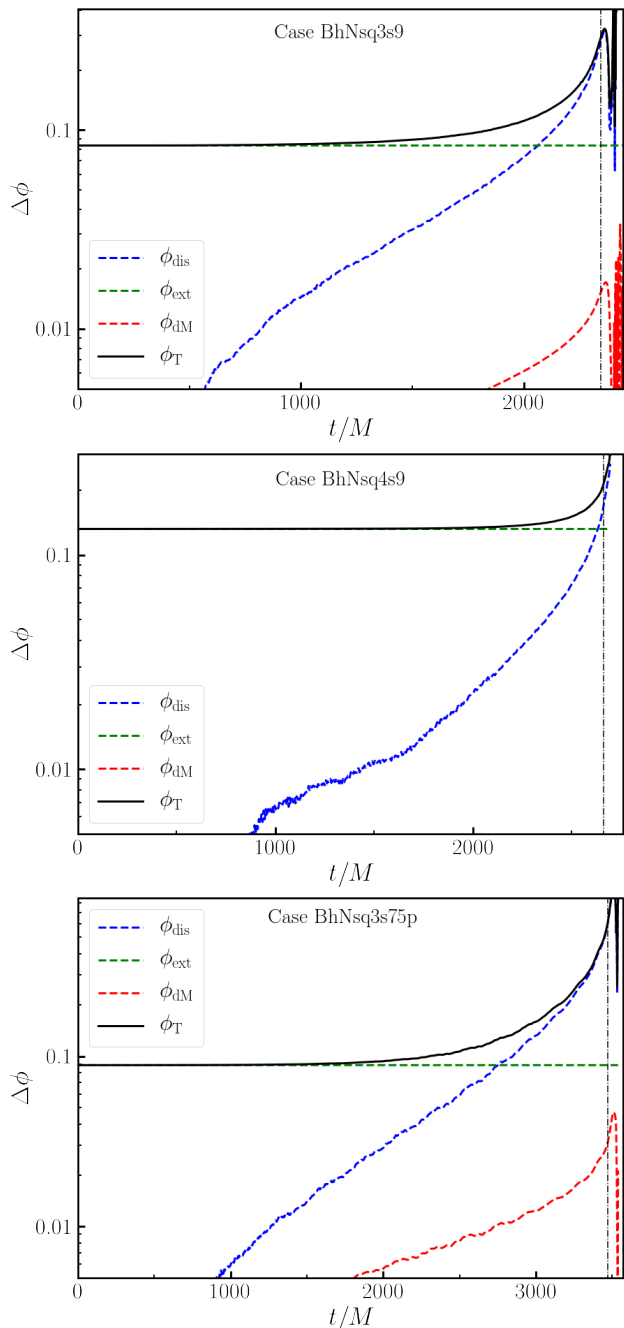
FIG. 4. Estimates of the phase error for the $(2, 2)$ mode of the GW signal in each simulation. We include the impact of mass losses (Φ_{dM}), extrapolation error (Φ_{ext}), and grid resolution (Φ_{dis}), following the methods described in [12] and Sec. III B. The vertical dot-dashed curves indicate t_{peak}.

provided us with conservative estimates of the numerical error whenever improved numerical methods / increased computational resources have allowed us to test it against higher accuracy results.

Extrapolation errors are estimated by measuring the phase difference between 2nd order and 3rd order extrapolation between $t = 0$ and t_{peak} (the time when the amplitude of the $(2, 2)$ mode of the GW signal is maximum, see Table I). The maximum phase difference over that time span is taken as the extrapolation error. This is a very conservative choice that was made largely because, as opposed to BBH simulations, BHNS simulations do not show clear convergence of the extrapolated waveform with the chosen extrapolation order. This has not been much of an issue so far, as extrapolation errors remained much smaller than finite resolution errors [12]. Fig. 4 shows that this is no longer the case in these new simulations. We thus perform a more in-depth study of extrapolation errors in the following section, that indicates that extrapolation using a 2nd order polynomial in $1/r$ leads to extrapolation errors that are significantly smaller than those shown in Fig. 4.

Finally, a small mass loss at the boundary of our finite difference grid could lead to an error in the mass of the neutron star, and thus in the phase evolution of the system. However, none of the simulations presented here loses more than $10^{-4} M_\odot$ over the course of the binary inspiral, and the phase error due to mass loss at grid boundaries is thus negligible.

Overall, we note that the phase error at t_{peak} is $(0.2 - 0.4)$ rad and limited by the finite resolution of the simulation, while during inspiral it is $\lesssim 0.1$ rad and limited by the estimated extrapolation error. We will however see in Sec. III C that the true extrapolation error is nearly certainly significantly smaller than what is shown on Fig. 4. We keep the estimate on Fig. 4 to allow for direct comparisons with our previous waveforms [12]. Only one simulation from [12] has smaller phase errors, and it is an equal mass, non-spinning configuration that is significantly easier for our code to evolve, and slightly shorter (in number of orbits) than the simulations presented here.

We can also estimate the uncertainty in the amplitude of the GW signal in our simulations. Fig. 5 shows relative differences between the amplitude of the $(2, 2)$ mode of the waveforms at different resolutions, and using different extrapolation orders. We see that the errors are small (typically $\lesssim 1\%$), especially when compared to current calibration uncertainties in GW detectors (e.g. $7\% - 10\%$ for GW170817 [3]). Finite-resolution errors often appear negligible when compared to extrapolation errors. This is because the main error due to finite resolution is a small shift of the time required for binaries to orbit / inspiral. This has a much larger effect on the phase of the gravitational wave signal than on its slowly-varying amplitude.

Very similar results are found for higher-order (l, m) modes, except that the phase error is multiplied by $l/2$, i.e. the ratio of the frequency of the (l, m) mode and the frequency of the $(2, 2)$ mode. This is once more a consequence of the fact that the dominant source of error is a slight change in the evolution timescale of the system. The resulting time shift in the waveform is the same for all modes, and the associated phase error is thus proportional to the frequency of the mode. Relative errors
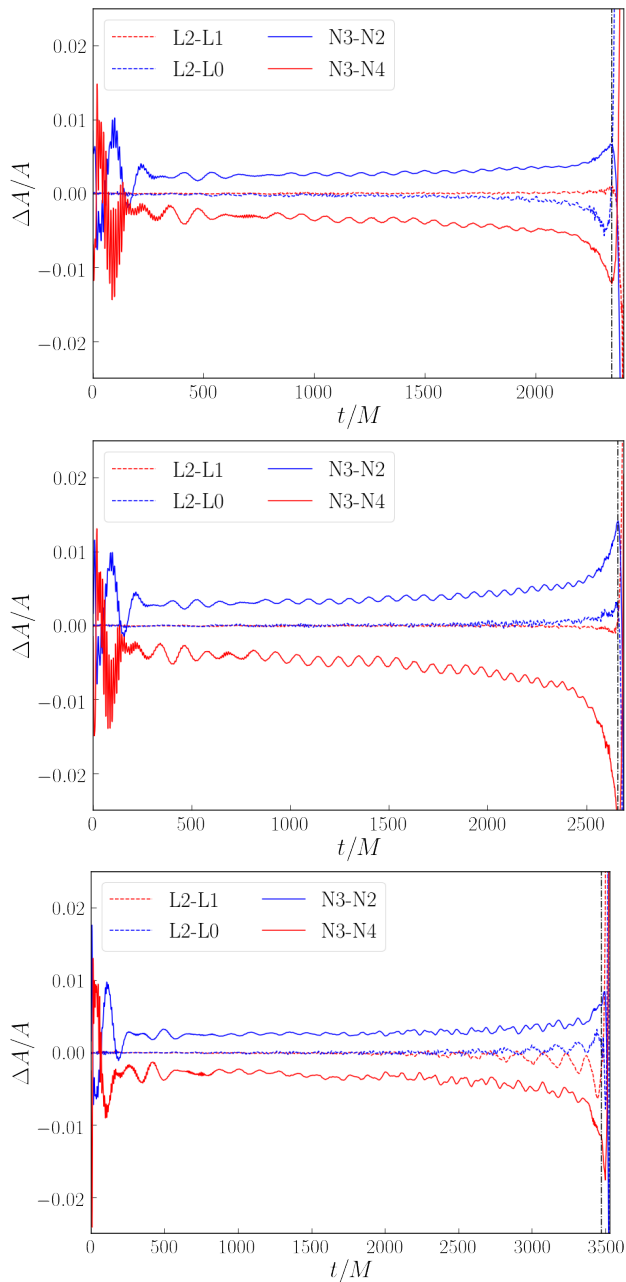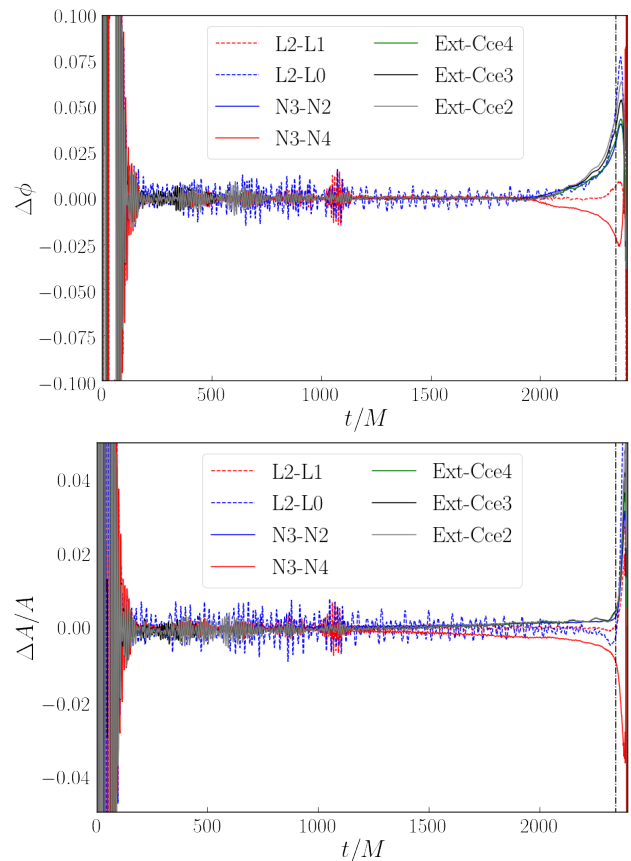
FIG. 6. Difference in the phase of the (2,2) mode for various estimates of Ψ_4 in system Q3S9, after application of a time and phase shift minimizing the phase error in the window $t \in [500, 2300]$. We show phase differences between our numerical resolutions (L0,L1,L2), between different extrapolation orders (N2,N3,N4, for the highest resolution simulation), and between the highest resolution simulation with $N = 3$ and the waveforms obtained using CCE from data extracted at $R = 200M$ (Cce2), $R = 300M$ (Cce3), and $R = 400M$ (Cce4). We see very good agreement between the various CCE waveforms, and between $N = 2$ extrapolation (solid blue curve) and CCE.

FIG. 5. Estimates of the relative error in the amplitude of the (2,2) mode of the GW signal for each simulation. From top to bottom, we show results for Q3S9, Q4S9, and Q3S75p. Dashed curves show comparisons between the amplitude obtained with different numerical resolutions ($L0, L1, L2$), while the solid curves show comparisons between the amplitude for different extrapolation orders ($N2, N3, N4$). The vertical dot-dashed curves indicate t_{peak}. We clearly see that extrapolation is the main source of error when estimating the GW amplitude. Amplitude errors are $\lesssim 1\%$, except for simulation Q4S9 at the time of merger.

in the amplitude of the signal only increase slightly for higher order modes (see e.g. Figs 9-11 in the next section). The absolute error in the amplitude of the signal is thus dominated by the error in the dominant $l = 2$ mode(s) for most binary orientations.

C. Extrapolation errors and Cauchy Characteristic Evolution

The previous section showed that extrapolation errors may significantly contribute to our error budget. However, extrapolation errors are difficult to assess: there is no clear improvement as the order of extrapolation increases, and in fact errors tend to grow beyond $N = 4$ extrapolation. In this section, we will argue that $N = 2$
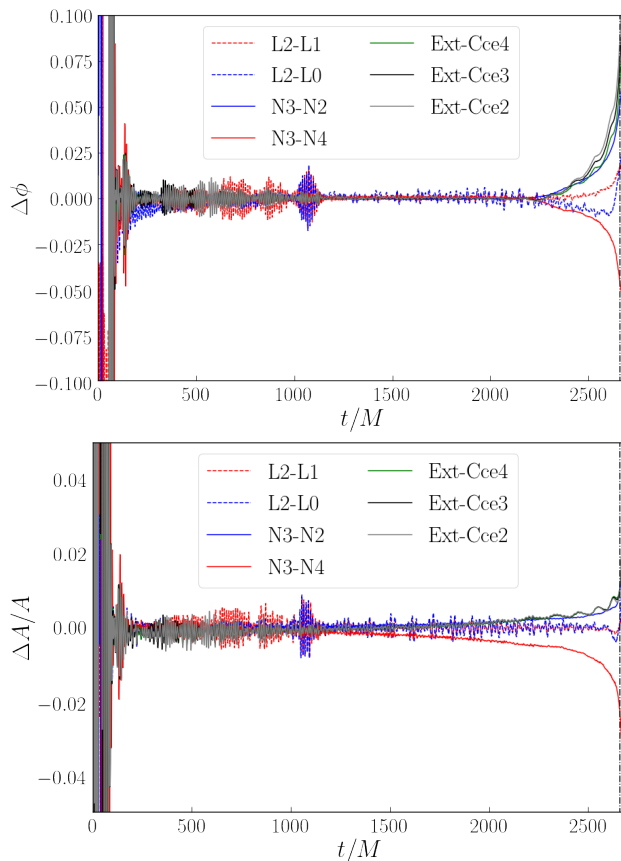
FIG. 7. Same as Fig. 6, but for the $(2, 2)$ mode in the Q4S9 configuration. We find very similar results, including very good agreement between all CCE results and the $N = 2$ extrapolation results.

FIG. 8. Imaginary part of the (2,2) mode of the GW strain ($h_\times$ polarization) using CCE extraction from $R = 400M$ and extrapolation with $N = 3$. The dotted line show the amplitude of the signal ($\sqrt{|h|^2}$). In the CCE results, the average value of the strain is shifted away from 0 during the first $\Delta t \sim 1000M$.

extrapolation is more accurate than our previous estimates would indicate. The main argument to that effect is a direct comparison of the GW signals obtained using extrapolation, and those obtained using CCE. We begin with a study of the Weyl scalar Ψ_4 (proportional to the second time derivative of the strain), for reasons that will become clear below. Fig. 6 shows the phase and amplitude differences between all wave extraction methods for case Q3S9, and Fig. 7 provides the same information for case Q4S9. Differences between simulations at 3 resolutions are also shown for reference. We note that CCE does not provide an absolute value of the time that can be used consistently for all waveforms, due to the unknown travel time between the inner boundaries of the various CCE evolution systems, and the differences between the simulation time of extrapolation and the asymptotically inertial time of CCE. Accordingly, all comparisons in this section are performed after application of a time and phase shift to the waveforms, chosen to minimize the phase difference before merger. Phase differences between waveforms using different resolution are naturally smaller than without alignment, to the point that after alignment differences between extrapolation
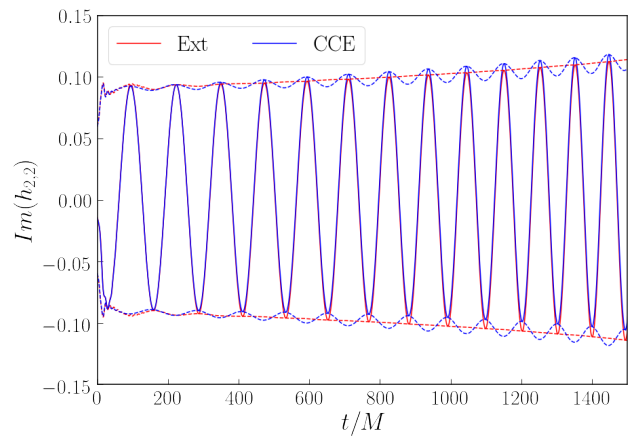
methods are at least of the same magnitude as differences between numerical resolutions. However, both case show two important results: all CCE waveforms are in very good agreement with each other ($\lesssim 0.01$ rad phase difference and $\lesssim 0.1\%$ amplitude error), and all CCE waveforms agree well with $N = 2$ extrapolation. As Ψ_4 is expected to be recovered to high accuracy by CCE, this is a first indication that $N = 2$ extrapolation provides accurate predictions for the GW signal at infinity.

Gravitational wave detectors, however, measure a projection of the complex strain $h = h_+ + ih_\times$, not Ψ_4. Unfortunately, CCE predictions for h suffer from a small drift in the average value of the strain over the first $1000M$ of evolution, illustrated in Fig. 8. In Fig. 8, the average value of $h_\times$ increases, eventually leading to a constant shift between the CCE and extrapolated results for $t > 1000M$, and visible oscillations in the inferred amplitude and phase of the GW signal.

This issue remains under investigation, although the likely cause of the early slow drift and late constant shift in the CCE strain is incomplete CCE initial data and a corresponding maladapted BMS frame. The evolution performed by CCE on outgoing null slices has a similar, but somewhat less dramatic, initial-data problem as the central Cauchy simulation: the accumulated effect of the past inspiral is difficult to estimate, and therefore transient effects appear during the early stage of the simulation. These initial data transients occur on a longer timescale than the Cauchy junk, due to the larger characteristic scale of the system. The CCE scale is set by the extraction radius, rather than the orbital scale of the compact merger. Finally, the initial transient leaves a lasting imprint on the CCE strain as an erroneous 'memory' contribution.

This issue can be partially negated by applying a constant offset to $h_+, h_\times$, chosen to zero the average of the
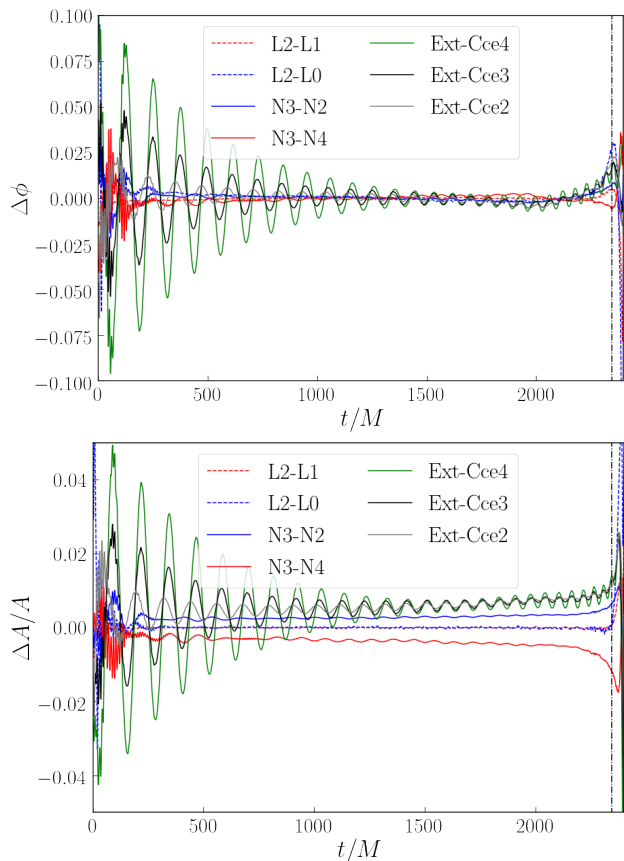
FIG. 9. Same as Fig. 6, but for the $(2,2)$ mode of the GW strain. For all CCE waveforms, we subtract a constant value from the strain (see text). Early-time oscillation in the CCE results are due to the initial drift in the average value of the GW strain.

FIG. 10. Same as Fig. 9, but after applying a highpass filter to all signals to remove the drift in the average value of the strain (10th order Butterworth filter with critical frequency of $0.002M^{-1} \approx 72\,\text{Hz}$). We note that this filter leads to changes in the phase of the signal that are larger than any of the errors displayed here; this plot indicates that most of the error observer in Fig. 9 is indeed due to a slow drift in the average value of the strain, but the filtering does not provide us with a better waveform template.

strain over a given time interval. Fig. 9 shows phase and amplitude differences for the strain for case Q3S9, with the averaging performed over 8 GW cycles immediately following $t = 1000M$. Large oscillations due to the drift in the strain dominate the errors, but once we make abstraction of this issue, results are similar to what we obtained for Ψ_4: the various CCE waveforms are very consistent with each other, and agree well with $N = 2$ extrapolation. The oscillations observed at the frequency of the GW signal in Fig. 9 can be avoided by applying a highpass filter on all compared signals, as shown in Fig. 10. While this confirms the origin of the phase error, this filtered signal cannot be substituted for the original signal when performing model comparisons; in our attempts to filter the signal, no filter could remove the phase error due to the drift in the average value of the strain without introducing larger phase differences as a result of the filtering itself.

Finally, we can follow the same procedure, but for the $(3,3)$ mode of the strain. Fig. 11 shows the resulting differences between extrapolated and CCE waveforms. Higher-order modes have very low amplitude early in the
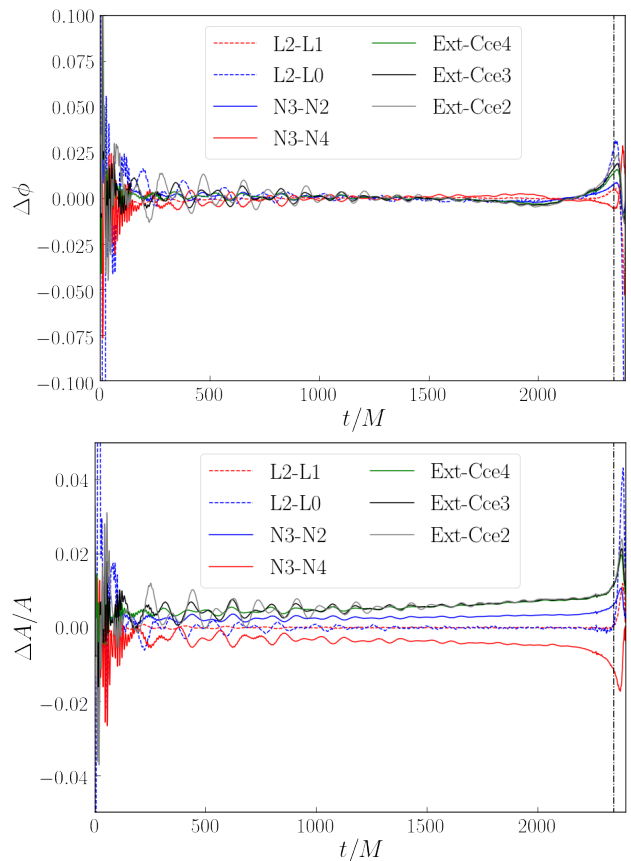
evolution of the binary, when the drift in the CCE strain occurs, and this appears to mitigate issues with the CCE method. As for Ψ_4 and the dominant mode of the strain, we find good agreement between all CCE waveforms and $N = 2$ extrapolation.

We can also gain some confidence in the accuracy of $N = 2$ extrapolation by studying how robust extrapolation results are when the finite radii used by the fits change. This is illustrated in Fig. 12, where we show the results of $N = 2, 3, 4, 5$ extrapolation at a given retarded time ($t = 3000M$) using all 20 finite radii, dropping the 4 radii farthest from the binary, and dropping the 8 radii farthest from the binary. High-order extrapolation appears to be significantly impacted by noise in the finite radius measurements, leading to visibly problematic extrapolation functions (e.g. non-monotonous behavior of $A(r)$). On the other hand, $N = 2$ extrapolation provides very consistent results. While we find that the exact behavior of the higher-order extrapolation methods de-
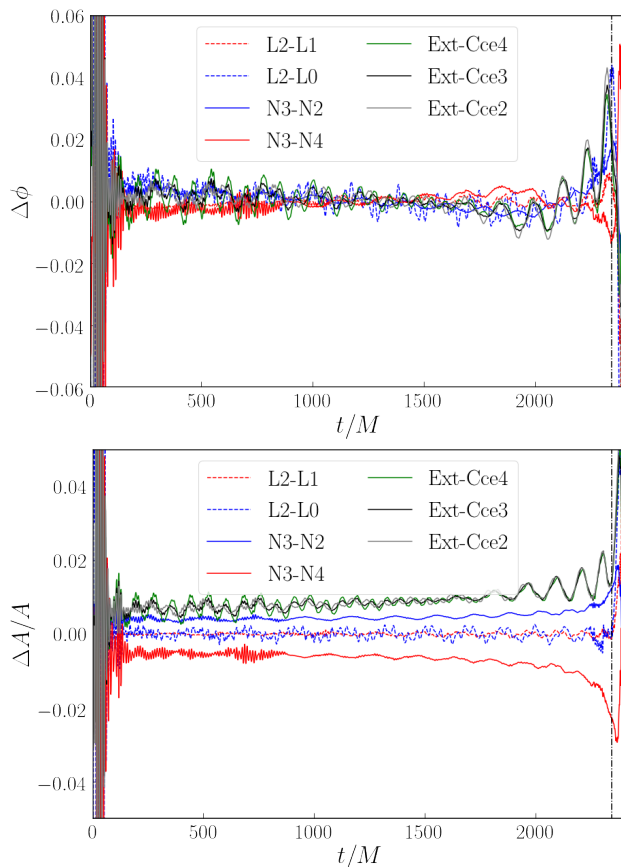
FIG. 11. Same as Fig. 9, but for the $(3, 3)$ mode of the GW strain. The CCE method provides much better result for this mode, with great self-consistency between CCE results using different extraction radii, and small differences between CCE and $N = 2$ extrapolation.

pends on the choice of retarded time under consideration, the robustness of $N = 2$ extrapolation does not.

As a result of this exploration of CCE and waveform extrapolation, we can revisit our estimates for extrapolation errors, and recommendation for the 'best' extrapolated waveform to use. The CCE waveforms (evolved from 3 different radii) and $N = 2$ extrapolation provide results consistent to better than 1% in the amplitude of the waveform, and better than 0.01 radian in its phase *after alignment of the waveforms using an appropriate time and phase shift, and ignoring oscillations due to the drift in the average value of the strain in CCE.* For applications where waveform alignment is necessary (e.g. comparisons with analytical models, hybridization,...), this should provide us with appropriate estimates of the extrapolation error *as long as we use $N = 2$ extrapolation, and not higher order methods.* Comparing $N = 2$ and $N = 3$ extrapolation provides a reasonable upper bound on the extrapolation error with $N = 2$ if that comparison is performed over the entire duration of the simulation.

D. Comparison with analytical models

With these error estimates in mind, we now move to a short investigation of the agreement between our numerical waveforms and commonly used waveform models in data analysis. We note than an in-depth study of modeling uncertainties goes beyond the scope of this paper. Here, we are mostly interested in providing a broad overview of the impact that various modeling choices and the physical effects of precession, higher modes, tidal effects during inspiral, and tidal disruption have on the agreement between numerical and analytical waveforms. We consider the following models, where the string 'IMRPhenom' indicates a model in the family of phenomenological inspiral-merger-ringdown BBH models in the frequency-domain [41–49], and 'SEOBNR' a model in the family of time-domain inspiral-merger-ringdown BBH models based on the Effective-One-Body formalism and calibrated to numerical simulations [50–65]:

- IMRPhenomXAS [66]: an aligned-spin BBH model that only includes the dominant $(2, \pm 2)$ modes of the strain

- IMRPhenomXP [67]: a BBH model that only includes the dominant $(2, \pm 2)$ modes of the strain and accounts for the main features of precession

- IMRPhenomXHM [68]: an aligned-spin BBH model that includes higher-order modes

- IMRPhenomXPHM [67]: a BBH model that includes higher-order modes and the main features of precession

- SEOBNRv4 [56]: an aligned-spin BBH model that uses only the dominant $(2, \pm 2)$ modes to determine the strain

- SEOBNRv4P [69]: a BBH model that includes precession by describing all six spin degrees of freedom throughout the BBH coalescence, but uses only the dominant $(2, \pm 2)$ modes to determine the strain

- SEOBNRv4HM [70]: an aligned-spin BBH model that includes higher-order modes

- SEOBNRv4PHM [69]: a BBH model that includes higher-order modes and precession by describing all six spin degrees of freedom throughout the BBH coalescence

- IMRPhenomPv2_NRTidalv2 [71]: a model whose BBH baseline includes the main features of precession that also incorporates tidal effects based on calibrating analytical results to numerical simulations of NSNS binaries [71–74], but not higher order modes. This model only describes the inspiral, with the signals tapered to zero at the NSNS merger frequency predicted by numerical simulations, and does not attempt to model the disruption of a neutron star by a black hole companion.
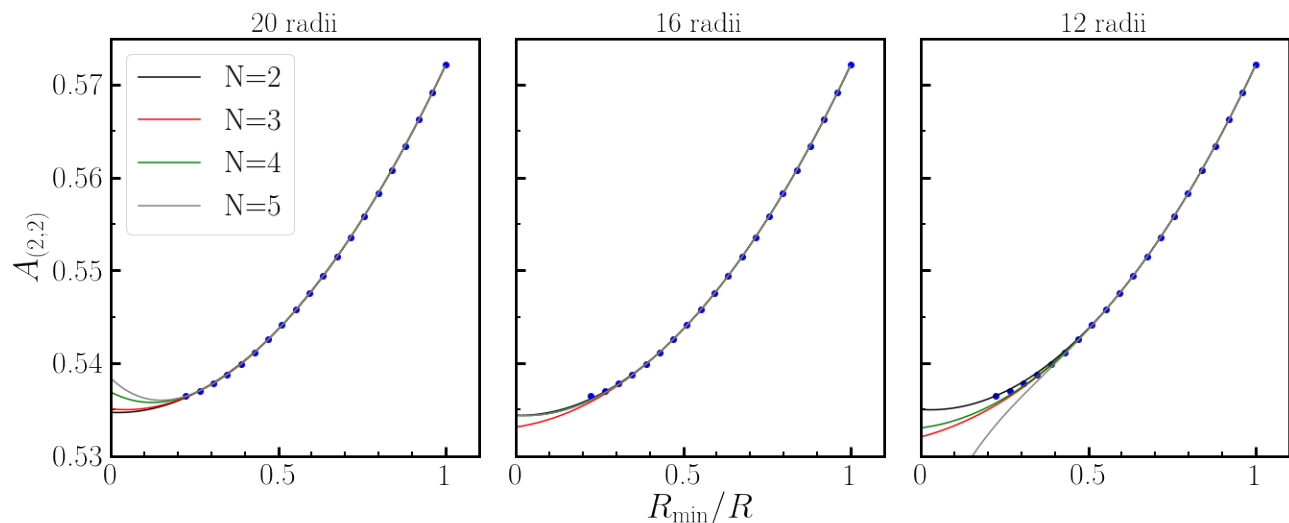
FIG. 12. Amplitude of the $(2,2)$ mode of the GW strain at a retarded time $t = 3000M$ for simulation Q3S9-L2. We show the values estimated at finite radii (normalized to $R_{\min} = 100M$), and the extrapolation polynomials of order $N = 2, 3, 4, 5$. From left to right, we show extrapolation functions fitted to data at all 20 radii used by our simulation, the 16 smallest radii of that set, and the 12 smallest radii of that set (only the $N = 3$ results is distinguishable from the others on the middle panel). While low-order extrapolation provides consistent and visually reasonable results, higher order extrapolation is less reliable. This can be contrasted with results in vacuum simulations, where results converge to a well-defined answer as N increases.

TABLE II. Faithfulness of analytical models to the numerical results, for the + polarization of the gravitational wave signal. For each case, we consider face-on observation (F suffix) and edge-on observation (E) suffix. All faithfulnesses are computed using the pyCBC library, as discussed in more detail in the text. We also provide the frequency f_{low} used as a lower bound for the calculation of the faithfulness, and the SNR of the part of the signal above f_{low} for systems at a distance of 100 Mpc. All calculations are performed using the Zero-Detuned High-Power noise power spectrum of LIGO.

Model	Q3S9-F	Q3S9-E	Q4S9-F	Q4S9-E	Q3S75p-F	Q3S75p-E
f_{low} [Hz]	300	300	250	250	300	300
SNR [100Mpc]	12.5	5.9	16.7	7.9	11.7	5.1
IMRPhenomXAS	0.972	0.957	0.981	0.955	0.988	0.917
IMRPhenomXP	0.969	0.955	0.976	0.951	0.986	0.959
IMRPhenomXHM	0.971	0.964	0.979	0.967	0.987	0.928
IMRPhenomXPHM	0.968	0.961	0.974	0.962	0.987	0.964
SEOBNRv4	0.966	0.952	0.976	0.950	0.987	0.915
SEOBNRv4P	0.966	0.952	0.976	0.951	0.970	0.955
SEOBNRv4HM	0.966	0.956	0.976	0.946	0.987	0.929
SEOBNRv4PHM	0.966	0.957	0.976	0.962	0.971	0.932
IMRPhenomPv2_NRTidalv2	0.985	0.972	0.969	0.947	0.984	0.957
SEOBNRv4T	0.977	0.963	0.983	0.956	0.976	0.872
IMRPhenomNSBH	0.991	0.977	0.994	0.969	0.992	0.904
SEOBNRv4_ROM_NRTidalv2_NSBH	0.988	0.973	0.991	0.966	0.992	0.900

- SEOBNRv4T [14, 75]: an aligned-spin model that includes analytical descriptions of tidal effects, but not higher-order modes. This model only describes the inspiral, with the signals tapered to zero at the NSNS merger frequency predicted by numerical simulations, and does not attempt to model the disruption of a neutron star by a black hole companion.

- IMRPhenomNSBH [8]: an aligned-spin model specifically designed for BHNS binaries: it includes both tidal effects from [71] and the disruption of the neutron star by the black hole. The model does not include higher-order modes.

- SEOBNRv4_ROM_NRTidalv2_NSBH [9]: an aligned-spin model specifically designed for BHNS binaries based on a reduced-order-model approximation to the frequency-domain BBH signals predicted by the SEOBNRv4 model: it includes both tidal effects from [71] and the disruption of the neutron star by the black hole. The model does not include higher-order modes.

These models are generally representative of the latest

iteration of the IMRPhenom and SEOBNR models (for recent models within another family of effective one body models see e.g. [76–78]). To determine the agreement between a model and numerical simulation, we calculate the faithfulness

$$F(h_1, h_2) = \max_{t_c, \phi_0} \left(\frac{\langle h_1, h_2 \rangle}{\sqrt{\langle h_1, h_1 \rangle \langle h_2, h_2 \rangle}} \right) \qquad (3)$$

with

$$\langle h_1, h_2 \rangle = 4\mathrm{Re} \int_{f_{\mathrm{low}}}^{f_{\mathrm{high}}} df \frac{h_1^*(f) h_2(f)}{S_n(f)} \qquad (4)$$

and $S_n(f)$ the one-sided power spectral density of the detector noise. Here, we take $f_{\max}$ to be very large ($\sim 100\,\mathrm{kHz}$), and set f_{low} to a value sufficiently large to avoid artifacts due to the finite length of the numerical simulations ($250\,\mathrm{Hz}$ for the $Q = 4$ system, and $300\,\mathrm{Hz}$ for the $Q = 3$ systems). The faithfulness is calculated using the pyCBC library [79], and that same library is used to generate the waveform models. Table II shows the faithfulness of the various analytical models to the highest-resolution numerical simulation (using $N = 2$ extrapolation). We calculate F for the $+$ polarization of the waveform and for observers located along the direction of the total angular momentum of the system (face-on) and in a direction orthogonal to the total angular momentum (edge-on). We note that the faithfulness of the low-resolution simulation to the high-resolution simulation is $F > 0.9999$ for all configurations and orientations, much larger than the faithfulness of any of the models.

For the precessing simulations, defining the initial spins require a few additional assumptions. For models that only include aligned spins, we define the black hole spin as the component of the spin aligned with the orbital angular momentum at the beginning of the numerical simulation. For models that do include precession, we also have to maximize F over the phase of the precession of the spin at a reference frequency, or over the reference frequency at which the spin is defined (depending on the inputs of the model).

We also calculate the SNR $\sqrt{\langle h, h \rangle}$ of the numerical waveform above f_{low} for a binary located at $100\,\mathrm{Mpc}$, to put the faithfulness numbers into context. We note that this is not the SNR of the full BHNS waveform, as a large fraction of the SNR is at frequency $f < f_{\mathrm{low}}$. All faithfulness and SNR results are summarized in Fig. II.

We can see some clear trend in these tabulated results. For systems observed face-on, the two models specifically designed for BHNS systems perform noticeably better, with $F \gtrsim 0.99$. Including tidal effects without accounting for the disruption of the neutron star helps for the Q3S9 system ($F \sim 0.98$ for the tidal models, $F \sim 0.97$ for the BBH models), but has no noticeable impact on the faithfulness for Q4S9 and Q3S75p. The use of higher-order modes and/or precession does not seem to impact F when a system is observed face-on. The faithfulness is generally lower for systems observed edge-on rather than

face-on. The IMRPhenomX BBH models also inidicates that for edge-on systems, F improves when including higher-order modes and, for the precessing binary, when including precession. With the SEOB models, higher-order modes help us improve F for the non-precessing systems, and including precession helps with the precessing system – but the model that include both effects actually perform worse than the precession-only and higher-mode only models. Finally, going from Q3S9 to Q4S9 to Q3S75p, the impact of tides decreases while the impact of higher-order modes / precession increases. As a result, the non-precessing tidal models (SEOBNRv4T, IMRPhenomNSBH, SEOBNRv4_ROM_NRTidalv2_NSBH) become less faithful, down to just $F \sim 0.90$ for Q3S75p. We note however that it is perfectly possible that an analytical waveform with parameters reasonably close to Q3S75p would provide a good match to the numerical results – we do not here attempt to find the best matching waveform, but only look at the faithfulness of the waveform for fixed initial conditions.

We can glean more information about the strengths and limits of various models by looking at Figs 13-14. The amplitude plots best capture the impact of the models specifically designed for BHNS systems. The amplitude of the signal is very well captured for Q3S75p. For Q3S9, a system with a black hole spin higher than those used to calibrate BHNS models, the agreement is a little worse, but still noticeably better than for systems that do not account for the disruption of the neutron star. The phase errors for Q3S9 show the importance of including tidal phase corrections in the models. This is very different from our results for Q3S75p, for which the inclusion of precession and higher-order modes improves the phase accuracy far more than the inclusion of tidal effects.

We also note that nearly identical results are obtained when comparing the second time derivative of the strain from analytical models to the CCE results for Ψ_4, indicating that one way to get around the drift in the strain when using CCE results to calibrate models could be to directly use Ψ_4 when performing these calibrations.

IV. CONCLUSIONS

We presented a new set of long, high-accuracy BHNS waveforms generated using the SpEC code, which are now publicly available. These waveforms sample regions of parameter space not covered by existing waveforms: high black hole spins, and one precessing system. All simulations are quite long by the standard of BHNS evolutions (> 13 orbits), and are of high accuracy in both phase ($0.2 - 0.4$ rad at merger) and amplitude ($\sim 1\%$ errors). They should thus be particularly helpful for testing and calibrating future BHNS waveform models.

In previous BHNS simulations, errors due to the extrapolation of the waveform to infinity were typically negligible even when using conservative error estimates. We find that this is no longer the case with our latest sim-
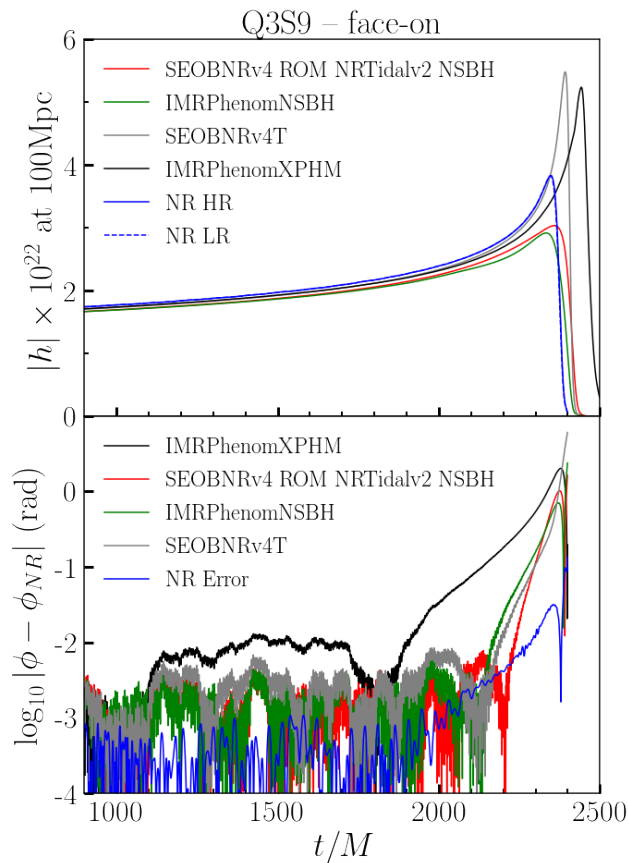
FIG. 13. *Top*: Amplitude of the complex strain $h = h_+ + ih_\times$ for system Q3S9 seen face-on. We show results for our low- and high-resolution simulations, for the two BHNS models, as well as for one tidal model that does not include neutron star disruption and one BBH model that includes precession and higher-order modes.*Bottom*: Phase difference between the models and the high-resolution numerical result. All signals are aligned by minimizing the phase difference in the time interval $t \in [500, 2000]M$.

FIG. 14. Same as Fig. 13, but for system Q3S75p.

ulations. Accordingly, we perform a more careful study of extrapolation errors by comparing waveform extrapolation to CCE. Results for Ψ_4 at null infinity indicate very good agreement between CCE initialized from different simulation radii, and between CCE and low-order (quadratic) extrapolation. Similar results appear to hold for the strain h, up to a drift in the time-averaged value of h that appears when using CCE. Higher-order extrapolation appears less reliable. Overall, we conclude that while extrapolation errors remain small (compared to finite-resolution errors) in our simulations when using quadratic extrapolation, one should avoid the use of higher-order methods in BHNS SpEC simulations. More practically, we thus recommend users of our waveform catalogue to take the highest resolution simulation with quadratic extrapolation ($N = 2$) as our 'best' waveform when multiple resolutions and/or extrapolation orders are available.
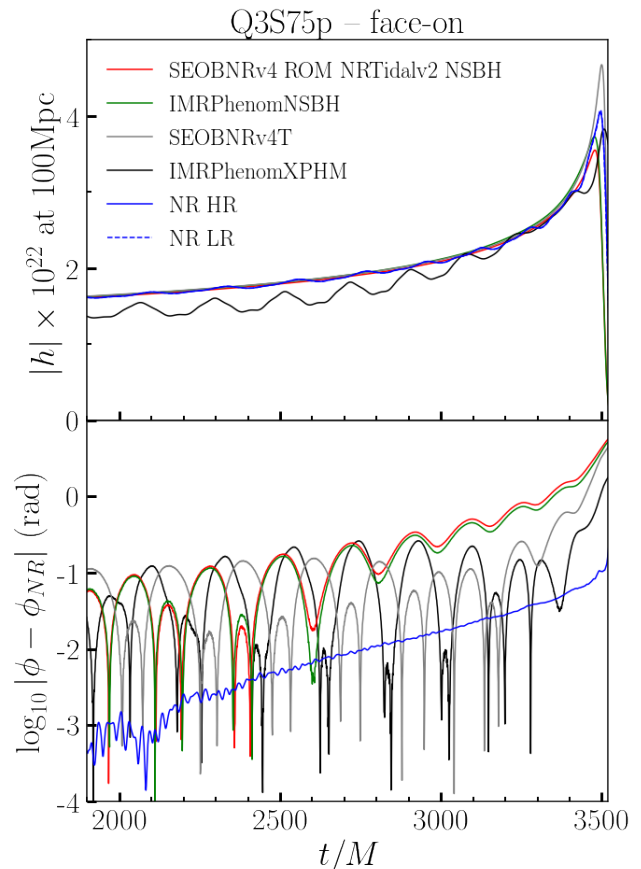
Finally, we compute the faithfulness to our simulation of a range of existing binary black hole models, binary neutron star models, and BHNS models. We focus on the high frequency portion of the signal that can be studied directly with our numerical waveforms. We find that for face-on observations, two recent BHNS models perform quite well – as already demonstrated in [9] for our non-precessing systems. The inclusion of higher-order modes and/or precession effects is less crucial to high faithfulness at high frequency. A more careful study of phase errors however indicates that even for observations along the total angular momentum of the system, the inclusion of precessional effects can help reduce phase differences with respect to our precessing system. For non-precessing systems observed edge-on, higher-order modes and tidal effects are both significant. Finally, for the precessing system observed edge-on, only a few of the models used here manage to reach faithfulness $\gtrsim 0.95$ (all of them precessing models), and no model reaches faithfulness > 0.97. Thus, there certainly remain important improvements that could be made to BHNS models by combining recent progress in the modeling of finite size effects with state-of-the-art results for precession in black hole binaries.

[1] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. Lett. **116**, 061102 (2016), arXiv:1602.03837 [gr-qc].
[2] B. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. X **9**, 031040 (2019), arXiv:1811.12907 [astro-ph.HE].
[3] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and et al., Astrophys. J. Lett. **848**, L13 (2017), arXiv:1710.05834 [astro-ph.HE].
[4] R. Abbott *et al.* (LIGO Scientific, Virgo), Astrophys. J. **896**, L44 (2020), arXiv:2006.12611 [astro-ph.HE].
[5] B. Abbott *et al.* (LIGO Scientific, Virgo), Astrophys. J. Lett. **892**, L3 (2020), arXiv:2001.01761 [astro-ph.HE].
[6] I. Harry, A. Nitz, D. A. Brown, A. Lundgren, E. Ochsner, *et al.*, Phys. Rev. D **89**, 024010 (2014), arXiv:1307.3562 [gr-qc].
[7] B. D. Lackey, K. Kyutoku, M. Shibata, P. R. Brady, and J. L. Friedman, Phys. Rev. D **89**, 043009 (2014), arXiv:1303.6298 [gr-qc].
[8] J. E. Thompson, E. Fauchon-Jones, S. Khan, E. Nitoglia, F. Pannarale, T. Dietrich, and M. Hannam, Phys. Rev. D **101**, 124059 (2020), arXiv:2002.08383 [gr-qc].
[9] A. Matas *et al.*, (2020), arXiv:2004.10001 [gr-qc].
[10] K. Chakravarti *et al.*, Phys. Rev. **D99**, 024049 (2019), arXiv:1809.04349 [gr-qc].
[11] Y. Huang, C.-J. Haster, S. Vitale, V. Varma, F. Foucart, and S. Biscoveanu, (2020), arXiv:2005.11850 [gr-qc].
[12] F. Foucart *et al.*, Phys. Rev. **D99**, 044008 (2019), arXiv:1812.06988 [gr-qc].
[13] K. Kiuchi, K. Kawaguchi, K. Kyutoku, Y. Sekiguchi, M. Shibata, and K. Taniguchi, Phys. Rev. D **96**, 084060 (2017), arXiv:1708.08926 [astro-ph.HE].
[14] T. Hinderer *et al.*, Phys. Rev. Lett. **116**, 181101 (2016), arXiv:1602.00599 [gr-qc].
[15] SXS Collaboration, "Black-hole neutron-star binary simulation SXS:BHNS:0008," (2020), DOI:10.5281/zenodo.4139881.
[16] SXS Collaboration, "Black-hole neutron-star binary simulation SXS:BHNS:0009," (2020), DOI:10.5281/zenodo.4139890.
[17] SXS Collaboration, "Black-hole neutron-star binary simulation SXS:BHNS:0010," (2020), DOI:10.5281/zenodo.4139871.
[18] L. Lindblom, M. A. Scheel, L. E. Kidder, R. Owen, and O. Rinne, Class. Quantum Grav. **23**, S447 (2006), arXiv:gr-qc/0512093v3 [gr-qc].
[19] D. A. Hemberger, M. A. Scheel, L. E. Kidder, B. Szilágyi, G. Lovelace, N. W. Taylor, and S. A. Teukolsky, Class. Quantum Grav. **30**, 115001 (2013), arXiv:1211.6079 [gr-qc].

[20] M. D. Duez, F. Foucart, L. E. Kidder, H. P. Pfeiffer, M. A. Scheel, and S. A. Teukolsky, Phys. Rev. D **78**, 104015 (2008), arXiv:0809.0002 [gr-qc].
[21] F. Foucart, M. B. Deaton, M. D. Duez, L. E. Kidder, I. MacDonald, C. D. Ott, H. P. Pfeiffer, M. A. Scheel, B. Szilágyi, and S. A. Teukolsky, Phys. Rev. D **87**, 084006 (2013), arXiv:1212.4810 [gr-qc].
[22] D. Radice and L. Rezzolla, Astron.Astrophys. **547**, A26 (2012), arXiv:1206.6502 [astro-ph.IM].
[23] B. Szilágyi, Int. J. Mod. Phys. D **23**, 1430014 (2014), arXiv:1405.3693 [gr-qc].
[24] H. P. Pfeiffer, *Initial Data for Black Hole Evolutions*, Ph.D. thesis, Cornell University (2003).
[25] H. P. Pfeiffer, L. E. Kidder, M. A. Scheel, and S. A. Teukolsky, Comput. Phys. Commun. **152**, 253 (2003), gr-qc/0202096.
[26] F. Foucart, L. E. Kidder, H. P. Pfeiffer, and S. A. Teukolsky, Phys. Rev. D **77**, 124051 (2008), arXiv:arXiv:0804.3787 [gr-qc].
[27] H. P. Pfeiffer, S. A. Teukolsky, and G. B. Cook, Phys. Rev. D **62**, 104018 (2000).
[28] H. P. Pfeiffer, D. A. Brown, L. E. Kidder, L. Lindblom, G. Lovelace, and M. A. Scheel, Class. Quantum Grav. **24**, S59 (2007), gr-qc/0702106.
[29] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, and et al., ArXiv e-prints (2018), arXiv:1805.11581 [gr-qc].
[30] F. Foucart, T. Hinderer, and S. Nissanke, Phys. Rev. **D98**, 081501 (2018), arXiv:1807.00011 [astro-ph.HE].
[31] F. Foucart, L. Buchman, M. D. Duez, M. Grudich, L. E. Kidder, I. MacDonald, A. Mroue, H. P. Pfeiffer, M. A. Scheel, and B. Szilágyi, Phys. Rev. D **88**, 064017 (2013), arXiv:1307.7685 [gr-qc].
[32] M. Boyle and A. H. Mroué, Phys. Rev. D **80**, 124045 (2009), arXiv:0905.3177 [gr-qc].
[33] N. T. Bishop, R. Gómez, L. Lehner, and J. Winicour, Phys. Rev. D **54**, 6153 (1996), arXiv:9705033 [gr-qc].
[34] N. T. Bishop, R. Gomez, L. Lehner, M. Maharaj, and J. Winicour, Phys. Rev. D **56**, 6298 (1997), arXiv:gr-qc/9708065.
[35] J. Winicour, Living Rev. Rel. **8** (2005).
[36] M. C. Babiuc, B. Szilágyi, J. Winicour, and Y. Zlochower, Phys. Rev. D **84**, 044057 (2011), arXiv:1011.4223 [gr-qc].
[37] C. J. Handmer and B. Szilágyi, Class. Quantum Grav. **32**, 025008 (2015), arXiv:1406.7029.
[38] J. Moxon, M. A. Scheel, and S. A. Teukolsky, (2020), arXiv:2007.01339 [gr-qc].

[39] C. J. Woodford, M. Boyle, and H. P. Pfeiffer, Phys. Rev. D **100**, 124010 (2019), arXiv:1904.04842 [gr-qc].

[40] M. Boyle *et al.* (SXS collaboration), (2019), arXiv:1904.04831 [gr-qc].

[41] P. Ajith, S. Babak, Y. Chen, M. Hewitson, B. Krishnan, J. T. Whelan, B. Brügmann, P. Diener, J. Gonzalez, M. H. S. Husa, M. Koppitz, D. Pollney, L. Rezzolla, L. Santamaría, A. M. Sintes, U. Sperhake, and J. Thornburg, Class. Quantum Grav. **24**, S689 (2007), arXiv:0704.3764 [gr-qc].

[42] P. Ajith, S. Babak, Y. Chen, M. Hewitson, B. Krishnan, *et al.*, Phys. Rev. D **77**, 104017 (2008), arXiv:0710.2335 [gr-qc].

[43] P. Ajith, M. Hannam, S. Husa, Y. Chen, B. Bruegmann, N. Dorband, D. Mueller, F. Ohme, D. Pollney, C. Reisswig, L. Santamaria, and J. Seiler, Phys. Rev. Lett. **106**, 241101 (2011), arXiv:0909.2867 [gr-qc].

[44] L. Santamaría, F. Ohme, P. Ajith, B. Brügmann, N. Dorband, M. Hannam, S. Husa, P. Mösta, D. Pollney, C. Reisswig, E. L. Robinson, J. Seiler, and B. Krishnan, Phys. Rev. D **82**, 064016 (2010), arXiv:1005.3306 [gr-qc].

[45] 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].

[46] 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].

[47] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, *et al.*, Phys. Rev. Lett. **113**, 151101 (2014), arXiv:1308.3271 [gr-qc].

[48] P. Schmidt, M. Hannam, and S. Husa, Phys. Rev. D **86**, 104063 (2012), arXiv:1207.3088 [gr-qc].

[49] P. Schmidt, F. Ohme, and M. Hannam, Phys. Rev. D **91**, 024043 (2015), arXiv:1408.1810 [gr-qc].

[50] A. Buonanno and T. Damour, Phys. Rev. D **59**, 084006 (1999), arXiv:gr-qc/9811091 [gr-qc].

[51] A. Buonanno and T. Damour, Phys. Rev. D **62**, 064015 (2000), arXiv:gr-qc/0001013 [gr-qc].

[52] T. Damour and A. Nagar, Phys. Rev. D **76**, 064028 (2007), arXiv:0705.2519 [gr-qc].

[53] T. Damour, B. R. Iyer, and A. Nagar, Phys. Rev. D **79**, 064004 (2009), arXiv:0811.2069 [gr-qc].

[54] A. Nagar and A. Shah, Phys. Rev. D **94**, 104017 (2016), arXiv:1606.00207 [gr-qc].

[55] Y. Pan, A. Buonanno, R. Fujita, E. Racine, and H. Tagoshi, Phys. Rev. D **83**, 064003 (2011), arXiv:1006.0431 [gr-qc].

[56] A. Bohé, L. Shao, A. Taracchini, A. Buonanno, S. Babak, I. W. Harry, I. Hinder, S. Ossokine, M. Pürrer, V. Raymond, T. Chu, H. Fong, P. Kumar, H. P. Pfeiffer, M. Boyle, D. A. Hemberger, L. E. Kidder, G. Lovelace, M. A. Scheel, and B. Szilágyi, Phys. Rev. D **95**, 044028 (2017), arXiv:1611.03703 [gr-qc].

[57] 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), arXiv:1202.0790 [gr-qc].

[58] A. Taracchini *et al.*, Phys. Rev. **D89**, 061502 (2014), arXiv:1311.2544 [gr-qc].

[59] 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].

[60] Y. Pan, A. Buonanno, A. Taracchini, M. Boyle, L. E. Kidder, A. H. Mroué, H. P. Pfeiffer, M. A. Scheel, B. Szilágyi, and A. Zenginoglu, Phys. Rev. D **89**, 061501 (2014), arXiv:1311.2565 [gr-qc].

[61] E. Barausse and A. Buonanno, Phys. Rev. D **81**, 084024 (2010), arXiv:0912.3517 [gr-qc].

[62] E. Barausse, E. Racine, and A. Buonanno, Phys. Rev. D **80**, 104025 (2009), arXiv:0907.4745 [gr-qc].

[63] S. Babak, A. Taracchini, and A. Buonanno, Phys. Rev. **D95**, 024010 (2017), arXiv:1607.05661 [gr-qc].

[64] E. Barausse and A. Buonanno, Phys. Rev. D **84**, 104027 (2011), arXiv:1107.2904 [gr-qc].

[65] Y. Pan, A. Buonanno, M. Boyle, L. T. Buchman, L. E. Kidder, *et al.*, Phys. Rev. D **84**, 124052 (2011), arXiv:1106.1021 [gr-qc].

[66] G. Pratten, S. Husa, C. Garcia-Quiros, M. Colleoni, A. Ramos-Buades, H. Estelles, and R. Jaume, (2020), arXiv:2001.11412 [gr-qc].

[67] G. Pratten *et al.*, (2020), arXiv:2004.06503 [gr-qc].

[68] C. García-Quirós, M. Colleoni, S. Husa, H. Estellés, G. Pratten, A. Ramos-Buades, M. Mateu-Lucena, and R. Jaume, (2020), arXiv:2001.10914 [gr-qc].

[69] S. Ossokine *et al.*, (2020), arXiv:2004.09442 [gr-qc].

[70] R. Cotesta, A. Buonanno, A. Bohé, A. Taracchini, I. Hinder, and S. Ossokine, Phys. Rev. **D98**, 084028 (2018), arXiv:1803.10701 [gr-qc].

[71] T. Dietrich, A. Samajdar, S. Khan, N. K. Johnson-McDaniel, R. Dudi, and W. Tichy, Phys. Rev. D **100**, 044003 (2019), arXiv:1905.06011 [gr-qc].

[72] T. Dietrich, S. Bernuzzi, and W. Tichy, Phys. Rev. **D96**, 121501 (2017), arXiv:1706.02969 [gr-qc].

[73] T. Dietrich *et al.*, (2018), arXiv:1804.02235 [gr-qc].

[74] T. Dietrich, S. Bernuzzi, B. Bruegmann, and W. Tichy, in *26th Euromicro International Conference on Parallel, Distributed and Network-based Processing* (2018) pp. 682–689, arXiv:1803.07965 [gr-qc].

[75] J. Steinhoff, T. Hinderer, A. Buonanno, and A. Taracchini, Phys. Rev. **D94**, 104028 (2016), arXiv:1608.01907 [gr-qc].

[76] A. Nagar, G. Riemenschneider, G. Pratten, P. Rettegno, and F. Messina, Phys. Rev. D **102**, 024077 (2020), arXiv:2001.09082 [gr-qc].

[77] S. Akcay, S. Bernuzzi, F. Messina, A. Nagar, N. Ortiz, and P. Rettegno, Phys. Rev. D **99**, 044051 (2019), arXiv:1812.02744 [gr-qc].

[78] A. Nagar, G. Pratten, G. Riemenschneider, and R. Gamba, Phys. Rev. D **101**, 024041 (2020), arXiv:1904.09550 [gr-qc].

[79] "pycbc library," DOI: 10.5281/zenodo.3993665.