

Up-down instability of binary black holes in numerical relativity

Vijay Varma^{1,2,3,*}, Matthew Mould⁴, Davide Gerosa⁴, Mark A. Scheel³, Lawrence E. Kidder², and Harald P. Pfeiffer⁵

¹*Department of Physics, Cornell University, Ithaca, New York 14853, USA*

²*Cornell Center for Astrophysics and Planetary Science, Cornell University, Ithaca, New York 14853, USA*

³*TAPIR 350-17, California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, USA*

⁴*School of Physics and Astronomy and Institute for Gravitational Wave Astronomy, University of Birmingham, Birmingham B15 2TT, United Kingdom*

⁵*Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlenberg 1, Potsdam 14476, Germany*



(Received 13 December 2020; accepted 5 February 2021; published 1 March 2021)

Binary black holes with spins that are aligned with the orbital angular momentum do not precess. However, post-Newtonian calculations predict that “up-down” binaries, in which the spin of the heavier (lighter) black hole is aligned (antialigned) with the orbital angular momentum, are unstable when the spins are slightly perturbed from perfect alignment. This instability provides a possible mechanism for the formation of precessing binaries in environments where sources are preferentially formed with (anti) aligned spins. In this paper, we present the first full numerical relativity simulations capturing this instability. These simulations span ~ 100 orbits and $\sim 3\text{--}5$ precession cycles before merger, making them some of the longest numerical relativity simulations to date. Initialized with a small perturbation of $1^\circ\text{--}10^\circ$, the instability causes a dramatic growth of the spin misalignments, which can reach $\sim 90^\circ$ near merger. We show that this leaves a strong imprint on the subdominant modes of the gravitational wave signal, which can potentially be used to distinguish up-down binaries from other sources. Finally, we show that post-Newtonian and effective-one-body approximants are able to reproduce the unstable dynamics of up-down binaries extracted from numerical relativity.

DOI: [10.1103/PhysRevD.103.064003](https://doi.org/10.1103/PhysRevD.103.064003)

I. INTRODUCTION

The detections of gravitational waves (GWs) emitted by the inspiral and merger of stellar-mass black hole (BH) binaries are now regular events for LIGO and Virgo [1,2]. Upgrades to detector sensitivities and waveform models and increasing catalog size will lead to improved inference on the parameters of individual detections as well as those of the underlying source population [3–5]. Of particular importance are the BH spins, which for generic binaries are tilted with respect to the orbital angular momentum and can cause significant modulations in the emitted GW signal due to precession of the orbital plane [6,7]. Spin orientations are powerful observables for determining the astrophysical formation channels of GW events [8–15].

Though in general the BH spins will change direction over the inspiral, configurations in which both spins are aligned with the orbital angular momentum of the binary are equilibrium solutions of the spin precession problem. Due to their regular behavior and simpler dynamics, BH binaries with aligned spins have been used extensively to construct waveform models, implement GW searches, and

perform numerical-relativity (NR) simulations. They are also interesting from an astrophysical standpoint, since stellar-mass BH binary formation via isolated stellar evolution [14,16] or embedment in gaseous disks [17,18] may lead to BH binaries with small spin tilts. On the contrary, large misalignment are expected for binaries formed in cluster environments [19]. For supermassive BHs targeted by the Laser Interferometer Space Antenna (LISA) mission, spin orientations might help distinguishing between gas-rich and gas-poor host galaxies [20–22].

For unequal-mass systems, there are four distinct aligned-spin configurations. Referring to the direction of a component BH spin that is aligned (antialigned) with the orbital angular momentum as “up” (“down”), the four alignments are up-up, down-down, down-up and up-down. In this notation, the direction before (after) the hyphen labels spin of the heavier (lighter) BH. As first pointed out by Gerosa *et al.* [23], only the former three configurations are stable equilibria. On the other hand, up-down binaries, where the spin of the heavier (lighter) BH is aligned (antialigned) with the orbital angular momentum, can become unstable (see also Refs. [24,25]). More specifically, for a binary BH with component masses $m_1 \geq m_2$, total mass $M = m_1 + m_2$, mass ratio $q = m_2/m_1 \leq 1$, and

*vvarma@cornell.edu

dimensionless component spin magnitudes χ_1 and χ_2 [where index 1 (2) corresponds to the heavier (lighter) BH], there exist critical orbital separations (in geometrical units $G = c = 1$)

$$r_{\text{ud}\pm} = \frac{(\sqrt{\chi_1} \pm \sqrt{q\chi_2})^4}{(1-q)^2} M, \quad (1)$$

such that the up-down configuration is unstable for orbital separations r in the range $r_{\text{ud}+} > r > r_{\text{ud}-}$.

An up-down binary BH that forms at a large separation $r > r_{\text{ud}+}$ with (infinitesimally small) perturbations to the spin directions remains near its initial configuration until reaching $r = r_{\text{ud}+}$. Upon inspiraling past this threshold, the binary becomes unstable and begins to precess, leading to large tilts between the BH spins and the orbital angular momentum. Alternatively, a perturbed up-down binary initialized within the range $r_{\text{ud}+} > r > r_{\text{ud}-}$ will be immediately unstable, while one initialized at $r < r_{\text{ud}-}$ will be stable (note, however, that $r_{\text{ud}-} \lesssim M$ for most binary parameters).

This effect was further investigated in Ref. [25], which showed that up-down binaries inspiraling from large separation evolve toward specific, predictable spin configurations after hitting the instability onset. The up-down instability therefore provides the means by which binary BH spins initially (anti)aligned by astrophysical formation can become misaligned and precessing in the sensitivity window of ground- and space-based GW interferometers. This may be the case, e.g., for stellar-mass BHs which are captured by and subsequently merge within the accretion disk of an active supermassive BH [17,18,26–30].

But this is only true if such unstable behavior persists until merger. Both the occurrence [23] and the end point [25] of the up-down instability were derived using the multitime-scale post-Newtonian (PN) framework of Refs. [31,32]. While PN techniques accurately describe the binary dynamics during the earlier inspiral, they inevitably fail to capture strong-field effects near merger. The only accurate solutions to the full general relativistic two-body problem are currently provided by NR simulations.

NR simulations of the up-down instability have so far been elusive because of their high computational cost. The instability is a precessional effect, and its observation thus requires a binary to complete at least one, and ideally several, precession cycles. This results in very long simulations because spin precession happens on a longer timescale, $t_{\text{pre}}/M \sim (r/M)^{5/2}(1+q)^2/q$, compared to the orbital period, $t_{\text{orb}}/M \sim (r/M)^{3/2}$. Sampling N_{pre} precession cycles requires simulating a number of orbits $N_{\text{orb}} = N_{\text{pre}} t_{\text{pre}}/t_{\text{orb}} \sim N_{\text{pre}}(r/M)(1+q)^2/q$. In Ref. [25] it was observed that the precessional instability appreciably develops over a typical decrease $\sim 25M$ in the orbital separation, resulting in simulations with $\mathcal{O}(100)$ orbits. For context, typical NR simulations cover $\lesssim 20$ orbits [33],

while the longest numerical relativity simulation performed to date covers 175 orbits [34]. While the astrophysically relevant scenario, in which an up-down binary is initialized in the stability regime ($r > r_{\text{ud}+}$) and becomes unstable, remains prohibitive, simulations instead initialized within the instability regime ($r_{\text{ud}+} > r > r_{\text{ud}-}$) are still challenging but possible with current capabilities.

In this paper, we present the first NR simulations of unstable up-down binaries and confirm that earlier PN predictions hold in the highly dynamical, strong-field regime of general relativity (GR). The rest of the paper proceeds as follows. In Sec. II, we describe our NR runs. In Sec. III, we present our results. In particular, we (i) observe the precessional instability in the up-down simulations, (ii) compare the stable and unstable configurations, and (iii) compare NR results against PN and effective-one-body (EOB) predictions. In Sec. IV, we present our conclusions.

II. NUMERICAL RELATIVITY SIMULATIONS

NR simulations for this work are performed using the Spectral Einstein Code (SPEC) [33] developed by the Simulating eXtreme Spacetimes (SXS) Collaboration [35]. We perform 12 new NR simulations that have been assigned the identifiers SXS:BBH:2313–2324 and are made publicly available through the SXS catalog. The parameters of our runs are summarized in Table I. For each simulation, we take mass ratio $q = 0.9$ and dimensionless spins $\chi_1 = \chi_2 = 0.8$. This is because the up-down instability effect is most prominent for binaries with close to (but not identically) equal masses and large spin magnitudes.

TABLE I. The parameters of the 12 NR simulations performed in this study. We provide the identifier for each simulation within the SXS catalog [35]. Each simulation has mass ratio $q = 0.9$ and spin magnitudes $\chi_1 = \chi_2 = 0.8$. For each of the four spin configurations (up-up, down-down, down-up and up-down) we perform three simulations with initial spin misalignments $\theta_{\text{pert}} = 1^\circ, 5^\circ, 10^\circ$ with respect to the perfectly aligned-spin configuration. The initial orbital frequency ω_{orb} is chosen such that the initial separation is $r = 30M$. We also report the merger time t_{merger} and number of orbits N_{orb} .

ID	q	$\chi_{1,2}$	θ_{pert}	Configuration	$M\omega_{\text{orb}}$	N_{orb}	t_{merger}/M
2313	0.9	0.8	1°	up-up	0.0058	106	68044
2314	0.9	0.8	1°	down-down	0.0055	94	67681
2315	0.9	0.8	1°	down-up	0.0057	100	67965
2316	0.9	0.8	1°	up-down	0.0057	100	67973
2317	0.9	0.8	5°	up-up	0.0058	106	68005
2318	0.9	0.8	5°	down-down	0.0055	94	67679
2319	0.9	0.8	5°	down-up	0.0057	100	67978
2320	0.9	0.8	5°	up-down	0.0058	96	63924
2321	0.9	0.8	10°	up-up	0.0058	106	67970
2322	0.9	0.8	10°	down-down	0.0055	94	67685
2323	0.9	0.8	10°	down-up	0.0057	100	67943
2324	0.9	0.8	10°	up-down	0.0058	96	63926

The initial binary separation is chosen to be $r = 30M$, which is much smaller than the instability threshold PN prediction $r_{\text{ud}+} \simeq 900M$ (note also that $r_{\text{ud}-} \sim 10^{-4} M''$). Ideally, we would like to initialize the binary at a separation $r > 900M$ so we can observe an initially stable binary develop the instability as it passes through $r = r_{\text{ud}+}$. While this would be more astrophysically realistic, unfortunately, $r > 900M$ would lead to $\sim 5 \times 10^5$ orbits before merger, which is well outside the capability of current NR codes. Choosing $r = 30M$ as the initial separation ensures that the instability has sufficient time to develop: these up-down binaries should already be unstable at their initial separation while at the same time undergoing several precession cycles before merger (cf. Fig. 1). Our simulations include $\simeq 100$ orbits, $\simeq 3\text{--}5$ precession cycles, and cover a time $t_{\text{merger}} \simeq 65 \times 10^3 M$, making them some of the longest simulations in the SXS catalog [33], with each simulation requiring about 10^5 CPU hours.

Besides the up-down cases, we perform control simulations in the other three aligned-spin configurations (up-up, down-down, and down-up), which are predicted to be stable. We introduce an initial perturbation to the perfectly aligned configuration to seed the instability. In particular, we consider three initial values $\theta_{\text{pert}} = 1^\circ, 5^\circ, 10^\circ$ of the angles between the BH spins and the orbital angular momentum direction. The angle between the in-plane spin components and the separation vector going from the lighter to the heavier BH are arbitrarily set to $\Phi_1 = 0$ and $\Phi_2 = \pi/2$. The precise values of Φ_1 and Φ_2 have a negligible impact because these simulations span several precession cycles.

The waveform is extracted at several extraction spheres at varying finite radii from the origin and then extrapolated to future null infinity [33,36]. The extrapolated waveforms are then corrected to account for the initial drift of the center of mass [37]. We denote the waveform modes at future null infinity, scaled to unit mass and unit distance, as $\hat{h}_{\ell m}(t)$. These enter the complex strain

$$\hat{h}(t, \iota, \varphi) = \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{\ell} \hat{h}_{\ell m}(t) {}_{-2}Y_{\ell m}(\iota, \varphi), \quad (2)$$

where ${}_{-2}Y_{\ell m}$ are the spin $= -2$ weighted spherical harmonics and ι and φ are the polar and azimuthal angles, respectively, on the sky in the source frame.

The component BH masses m_1 and m_2 and dimensionless spins $\chi_1(\bar{t})$ and $\chi_2(\bar{t})$ are evaluated on the apparent horizons [33] of the BHs. Here, $\bar{t}$ is the simulation time at which the spins are measured in the near zone [33]. Following previous studies [38,39], we identify time t with $\bar{t}$. While this identification is gauge dependent, the spin directions are already gauge dependent. However, we note that the spin and orbital angular momentum vectors in the damped harmonic gauge used by SPEC are in good agreement with the corresponding PN vectors [40].

Following Refs. [38,39], the orbital frequency ω_{orb} is computed as the time derivative of the orbital phase obtained from the coprecessing-frame waveform at future null infinity. The separation is then defined as a proxy for the orbital frequency by adopting the Newtonian expression

$$r \equiv M(M\omega_{\text{orb}})^{-2/3}. \quad (3)$$

We discard the initial $t < t_0 = 5000M$ of data as this is contaminated by spurious initial transients caused by imperfect initial data, also known as junk radiation [33]. This is more stringent than the typical choice of discarding 200–500M [33] of data, but we find that this is necessary to eliminate transient features in ω_{orb} . In addition, these long simulations adopt a larger outer boundary ($R_{\text{max}} \sim 2000M$) [34] compared to typical NR runs ($R_{\text{max}} \sim 600M$) [33], implying that junk radiation takes longer to exit the simulation domain. We report the spin vectors $\chi_{1,2}$ and the gravitational waveform $\hat{h}_{\ell m}$ in a frame where the z direction lies along the orbital angular momentum direction $\hat{L}$ at t_0 , the x direction is given by the separation vector from the lighter to the heavier BH at t_0 , and the y direction completes the orthonormal triad. Note that both $\hat{L}$ and the separation vector are estimated using the gravitational waveform $\hat{h}_{\ell m}$ as in Refs. [38,39].

We compute the tilt angles θ_i for the BH spins ($i = 1, 2$) as

$$\cos \theta_i(t) = \hat{\chi}_i(t) \cdot \hat{L}(t). \quad (4)$$

Their offsets from alignment are given by

$$\Delta \theta_i(t) \equiv |\theta_i(t) - \theta_i^{\text{align}}|, \quad (5)$$

where $\theta_i^{\text{align}} = 0^\circ$ (180°) for an unperturbed up (down) spin component. Finally, we define the instantaneous tilt angle of the orbital plane, θ_L , as

$$\cos \theta_L(t) \equiv \hat{L}(t) \cdot \hat{z}. \quad (6)$$

Strictly speaking, one has $\Delta \theta_1 = \Delta \theta_2 = \theta_{\text{pert}}$ only at the start of the simulation, but this remains approximately true at $t = t_0$ past the junk-radiation stage. Similarly, $\theta_L \simeq 0$ at t_0 . In the following sections, we use $\Delta \theta_{1,2}$ and θ_L to track, respectively, the spin and orbital plane precession as the binary evolves. The component BH apparent horizons are tracked until a common apparent horizon is formed at merger [33]. The variables $\Delta \theta_{1,2}$ are only available until this point, which typically corresponds to a separation $r \simeq 3M$.

Because our simulations are much longer than the typical simulations performed with SPEC, their accuracy needs to be investigated separately. We repeat each simulation with two resolutions, which we refer to as the low- and high-resolution runs. The grids for these simulations are determined using adaptive mesh refinement (AMR) as described in Ref. [33] and references within. The grid resolution

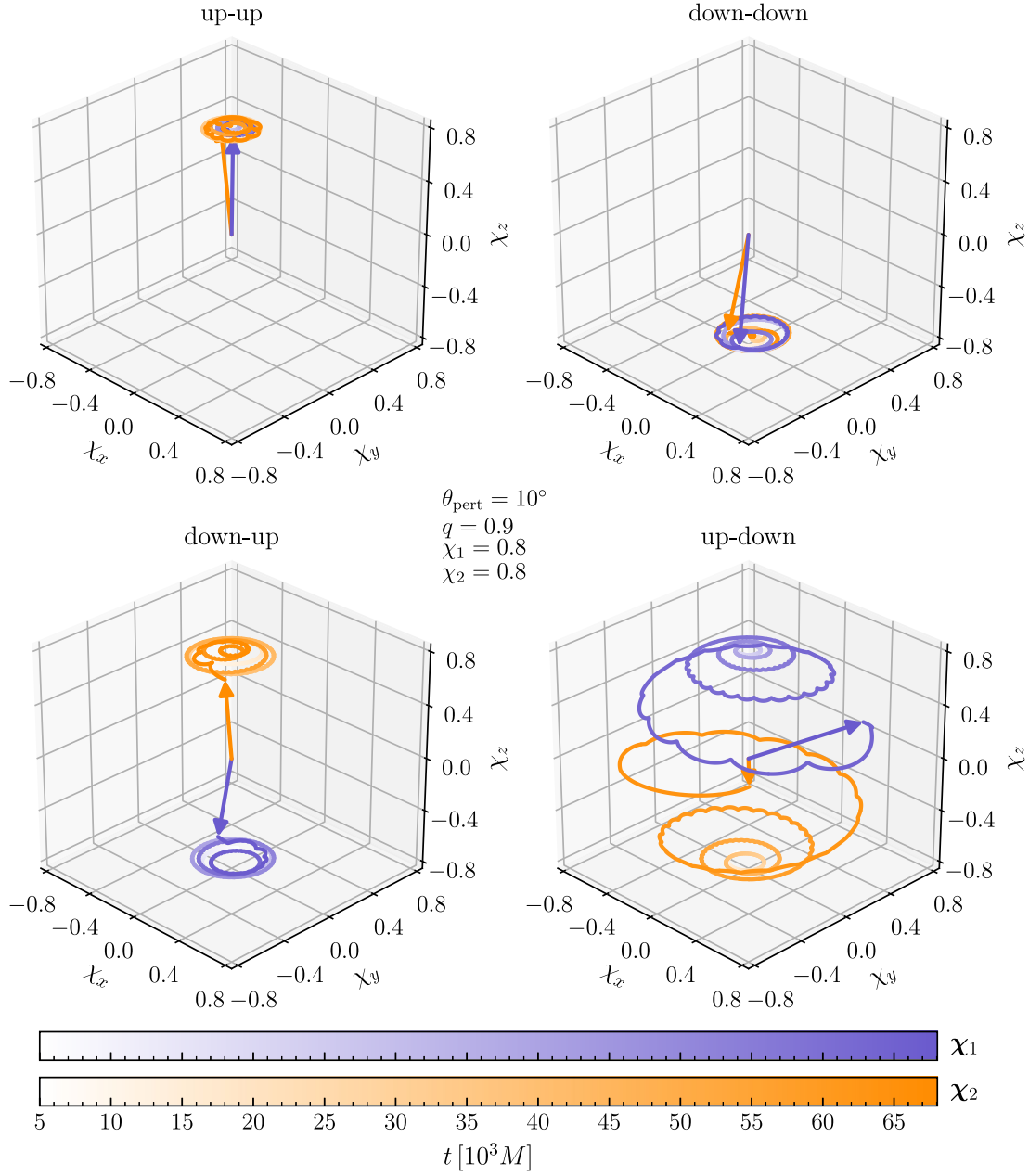


FIG. 1. The up-down instability is demonstrated in NR. We consider binaries with mass ratios $q = 0.9$ and spin magnitudes $\chi_1 = \chi_2 = 0.8$, while the component spin vectors are initially perturbed from a perfectly aligned-spin configuration by an angle $\theta_{\text{pert}} = 10^\circ$. The purple (orange) arrows represent the spin χ_1 (χ_2) of the heavier (lighter) BH near merger. The colored curves trace the evolution of $\chi_{1,2}(t)$ as the binary precesses. Colors darken linearly in time as indicated on the color bars. The spins of the up-up (top-left), down-down (top-right), and down-up (bottom-left) configurations precess stably about $\hat{z}$, while those of the up-down (bottom-right) configuration are unstable and become largely misaligned. Smaller modulations occur on the shorter orbital timescale. An animated version of this figure is available at <http://www.davidegerosa.com/spinprecession>.

varies dynamically during a simulation to satisfy an AMR tolerance parameter for constraint violation [33]. For these simulations, the AMR tolerance is chosen to be a factor of 4 smaller for the high-resolution runs compared to the low-resolution ones. Sections III A and III B present results using the high-resolution runs; the accuracy of the simulations is then evaluated in Sec. III C.

III. RESULTS

A. Unstable precession dynamics in NR

Figure 1 presents our main result: the up-down precessional instability, which is a 2PN prediction [32], is verified with full NR. We show the BH spin evolution for each of the four configurations with initial perturbation $\theta_{\text{pert}} = 10^\circ$.

As the binary inspirals, the BH spins precess around the z direction, as seen by the purple (orange) curves which trace the instantaneous values of χ_1 (χ_2). The spins of the stable configurations (up-up, down-down and down-up) remain close to their initial configuration and trace regular precession cones about alignment. On the other hand, the up-down configuration is very clearly unstable: the BH spins outspiral dramatically, leading to large misalignments with respect to $\hat{L}$.

While the dominant modulations in the spin directions are the polar oscillations due to spin precession, much smaller variations, sometimes referred to as nutations, occur due to the binary orbital motion on the shorter timescale $t_{\text{orb}} \ll t_{\text{pre}}$ [40]. Nutations are negligible near the beginning of the simulations and they become more pronounced near merger.

In Fig. 2 we dissect the unstable dynamics illustrated in Fig. 1 and exhibit their effect on the resulting gravitational waveforms. For all 12 simulations, we plot various quantities characterizing the binary dynamics and GW emission as a function of the binary separation r [cf. Eq. (3)].

The top two rows of Fig. 2 show the spin perturbation angles $\Delta\theta_{1,2}$ [cf. Eq. (5)] for the component BHs, indicating the amount of spin precession. First, considering the stable configurations (up-up, down-down and down-up), as θ_{pert} is increased from 1° (left) to 10° (right), the amplitudes of the oscillations in $\Delta\theta_{1,2}$ increase but remain close to the initial perturbations. In particular, the maximum perturbations across all stable configurations are $\max_i \Delta\theta_{1,2} \simeq 2^\circ, 10^\circ, 21^\circ$ for $\theta_{\text{pert}} = 1^\circ, 5^\circ, 10^\circ$, respectively.

For the up-down binaries, the instability takes a decrease $\sim 10M$ in separation to develop appreciably. Starting from $r \simeq 20M$, we observe the presence of an unstable growth in $\Delta\theta_{1,2}$. This growth becomes more rapid with larger initial perturbations θ_{pert} . The maximum perturbations for the up-down configuration are $\max_i \Delta\theta_{1,2} \simeq 11^\circ, 51^\circ, 94^\circ$ for $\theta_{\text{pert}} = 1^\circ, 5^\circ, 10^\circ$, respectively. In general, the deviations from the initial perturbations increase roughly tenfold by the end of each simulation (compared to a factor of ~ 2 for the stable binaries).

These observations are in qualitative agreement with the results presented in Ref. [25], which showed with orbit-averaged evolutions at 3.5PN order that, over a population of up-down binaries, the precessional instability develops over a typical decrease $\simeq 25M$ in the PN orbital separation. The study also suggested that the end point of the instability is independent of the initial perturbation, in contrast to the findings presented here. This is due to the short nature of NR simulations (though the simulations we performed are very long by NR standards) and the “astrophysically unrealistic” initialization we employed and discussed in Sec. II. The end point derivation of Ref. [25] intrinsically relies on binaries initially close to the up-down configuration at large separations $r > r_{\text{ud+}}$ inspiraling through the critical separation before becoming unstable.

In the third row of Fig. 2, we show the angle θ_L [cf. Eq. (6)], which indicates the amount of orbital-plane precession. The presence of the instability is less apparent in θ_L up to at least $r \sim 10M$. Even then, the maximum value reached by θ_L for the up-down configuration is $\max_i \theta_L \simeq 1^\circ, 7^\circ, 12^\circ$ for $\theta_{\text{pert}} = 1^\circ, 5^\circ, 10^\circ$, respectively [though in each case $\theta_L(t_0) = 0^\circ$]. The smaller deviation in θ_L compared to $\Delta\theta_{1,2}$ is due to the vastly different magnitudes of the spin and orbital angular momenta. The binaries simulated here have an initial orbital frequency $\omega_{\text{orb}} \sim 0.006 \text{ rad}/M$ (cf. Table I) which corresponds to $L = m_1 m_2 (M \omega_{\text{orb}})^{-1/3} \simeq 1.4M^2$. For $q = 0.9$ and $\chi_1 = \chi_2 = 0.8$, the magnitude of the spin angular momenta are given by $S_1 = m_1^2 \chi_1 \simeq 0.22M^2$ and $S_2 = m_2^2 \chi_2 \simeq 0.18M^2$. It is natural to expect that the orbital angular momentum has more “inertia” to modulations from the instability and thus $\theta_L < \Delta\theta_{1,2}$. The angle θ_L completes ~ 3 – 5 cycles compared to only ~ 1 full period for $\theta_{1,2}$. This is in qualitative agreement with previous PN predictions¹ [41].

Notably, orbital precession leads to amplitude modulations of the emitted GW signal [6]. In the bottom two rows of Fig. 2, we investigate the influence of the up-down precessional instability on the gravitational waveforms; we focus in particular on the dominant (2,2) mode and a subdominant mode (2,1) of the GW strain decomposition given in Eq. (2).

The (2,2) mode is largely unaffected by the sensitive details of precession—the simulations of all four configurations and three initial perturbations present qualitatively similar results in this mode. The dominant morphological features of the (2,2) mode waveform are the GW cycles at approximately twice the orbital frequency and the typical “chirp” as the binary merges. However, we find that even the modest growth of θ_L mentioned above leaves a notable imprint on the subdominant modes like the (2,1) mode. This is apparent by comparing the third (θ_L) and fifth (θ_{21}) rows of Fig. 2. Starting from $r \sim 10M$, as θ_L experiences a growth for the up-down binary (indicating precession of the orbital plane), so does the amplitude of the (2,1) mode. This arises because orbital precession induces a transfer of power from the (2,2) mode to the subdominant modes (see, e.g., [38]). This transfer of power to subdominant modes is a feature of generic precessing binaries, not just the peculiar up-down configuration. The growth in θ_L becomes more pronounced as we increase the initial perturbation θ_{pert} from 1° to 10° . For $\theta_{\text{pert}} = 10^\circ$, the amplitude of the (2,1) mode near merger is about twice as large for the up-down configuration compared to the three stable binaries. We note that, for astrophysically realistic binaries, as they undergo the instability at much larger separations, we expect an even larger imprint on the waveform.

¹See Fig. 7 in Ref. [41] for binaries at $r \sim 10M$. In their notation, the ratio between the number of $\hat{L}$ cycles and spin-nutation cycles is given by $\alpha/2\pi$.

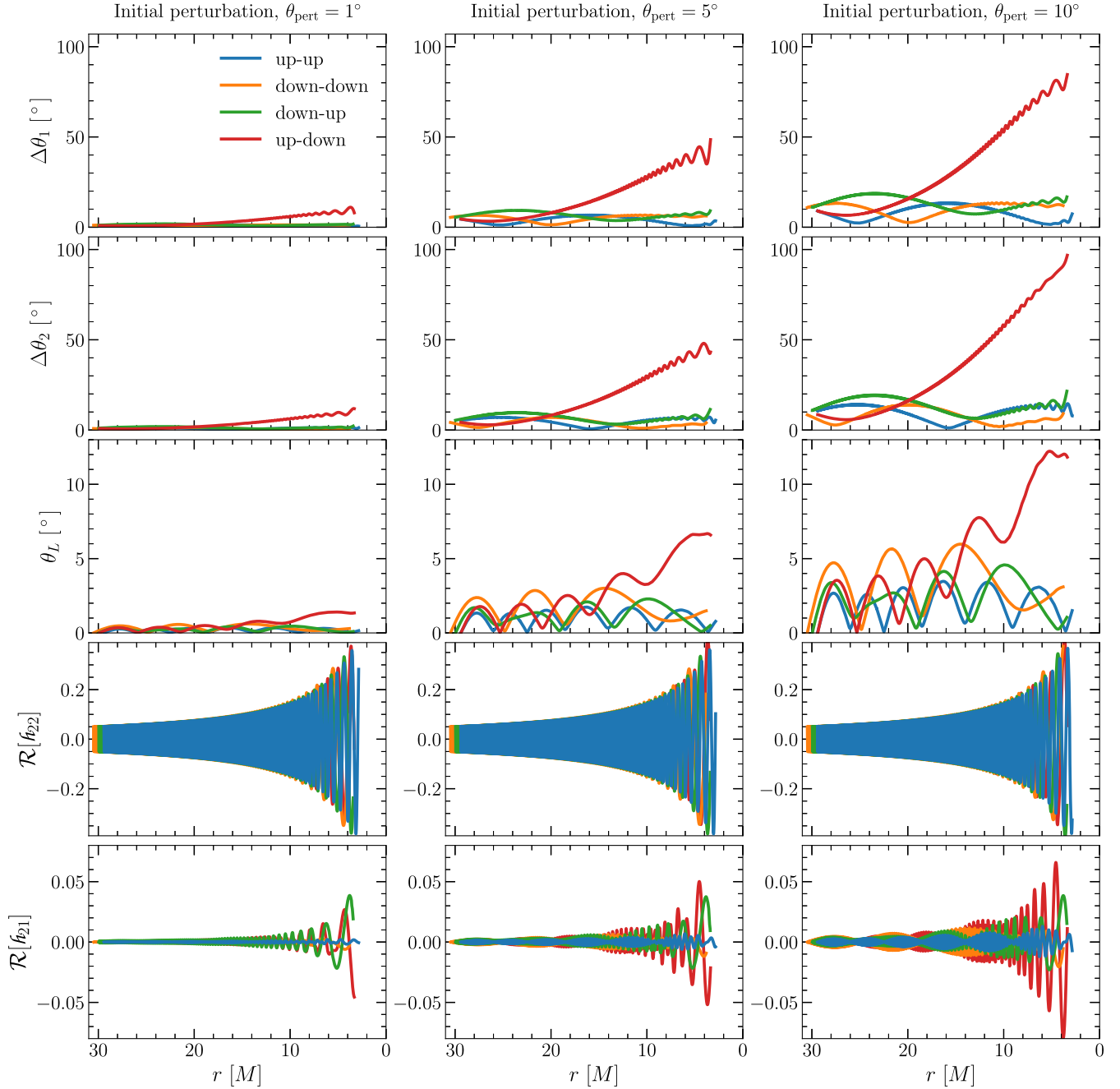


FIG. 2. Precessional dynamics and its effect on the gravitational waveform. Each subplot shows the four configurations we consider: up-up (blue), down-down (orange), down-up (green) and up-down (red). Each column corresponds to a different initial perturbation (θ_{pert} , indicated at the top) of the spin directions with respect to a perfectly aligned-spin system. The first (second) row shows the evolution of the spin perturbation angles [$\Delta\theta_{1,2}$; cf. Eq. (5)] of the heavier (lighter) BH. The third row shows the evolution of the orbital plane tilt angle [θ_L ; cf. Eq. (6)]. The fourth and fifth rows show the real part of the (2,2) and (2,1) GW modes, respectively. All quantities are shown as a function of the Newtonian separation r [cf. Eq. (3)] and are terminated at the separation where the common horizon is found. The instability development is reflected in a rapid growth of the angles $\Delta\theta_{1,2}$ and θ_L for binaries near the up-down configuration. This in turn impacts the observed GW signal, in particular the subdominant modes like (2,1), as the orbital precession causes power leakage from the dominant (2,2) mode.

B. Comparison with approximate evolutions

We now test how well approximate PN and EOB methods capture the full NR dynamics for stable and unstable aligned-spin binaries. The PN dynamics is evaluated using the

SpinTaylorT1 [42,43] approximant at 3.5PN as implemented in the LIGO Algorithm Library (LAL) Suite [44]. We find that SpinTaylorT1 is marginally more accurate than the other available approximants, SpinTaylorT4 [45] and

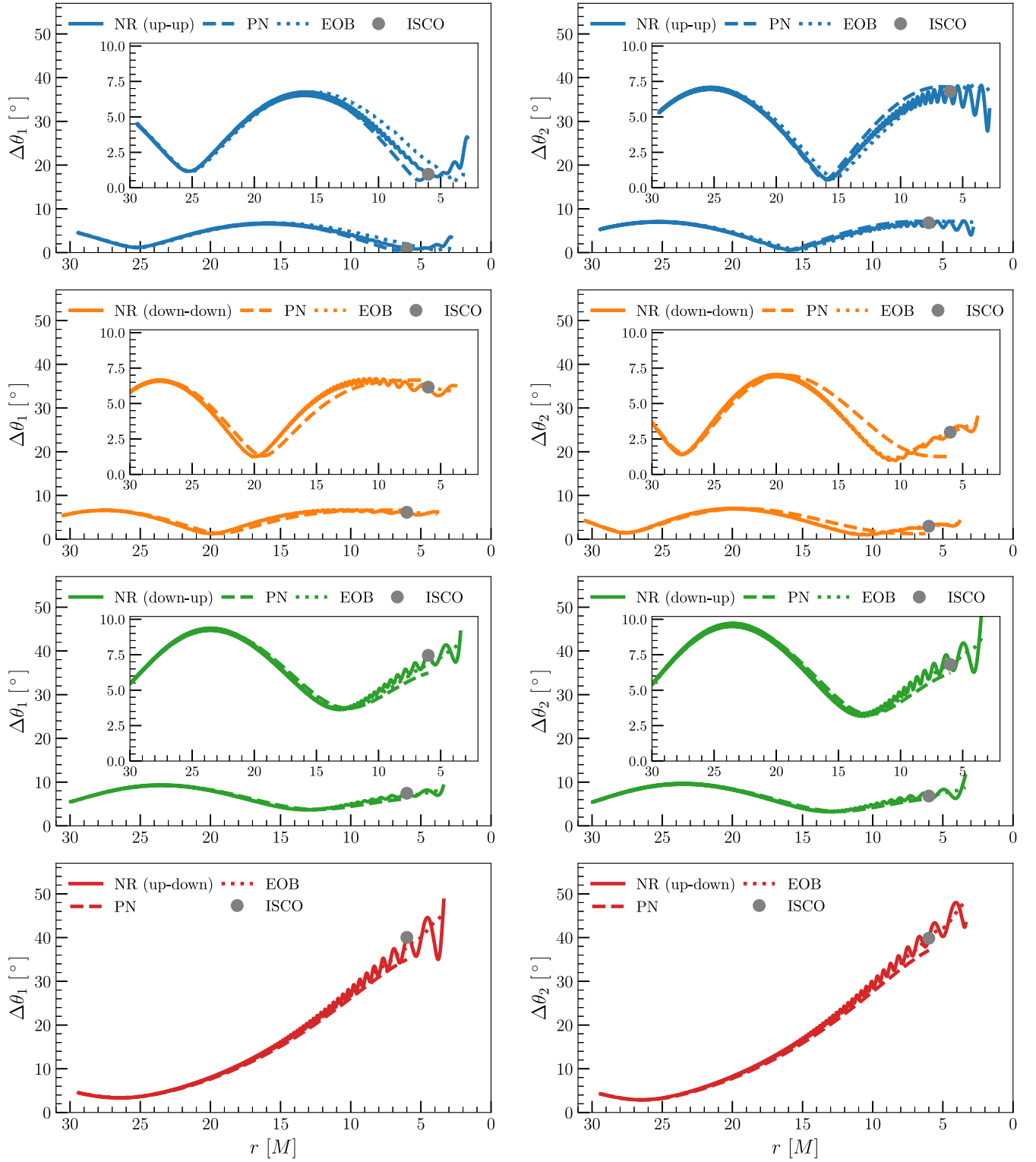


FIG. 3. Comparison of the spin dynamics predicted by PN and EOB approximants against NR. We show the spin perturbation angles $\Delta\theta_{1,2}$ as a function of the Newtonian separation r for each of the four configurations: up-up (top-row), down-down (second-row), down-up (third-row) and up-down (bottom-row), with an initial perturbation $\theta_{\text{pert}} = 5^\circ$. The left (right) column shows $\Delta\theta_1$ ($\Delta\theta_2$). For the stable configurations, the inset shows an enlarged version of the same panel. The Schwarzschild ISCO radius ($r = 6M$) is indicated by a gray marker. Both approximants are able to track the precession modulations in $\Delta\theta_{1,2}$ found in NR, including the unstable growth of the up-down instability, but fail to match the smaller spin oscillations occurring on the orbital timescale.

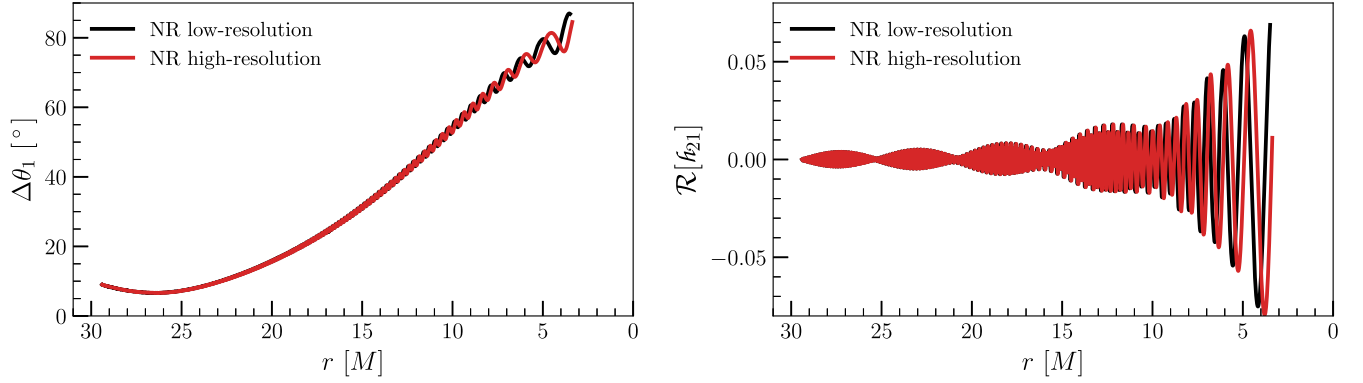


FIG. 4. Comparison of the spin dynamics and gravitational waveform between simulations with two different NR resolutions for the up-down system with an initial perturbation $\theta_{\text{pert}} = 10^\circ$. We show the spin perturbation angle $\Delta\theta_1$ and the real part of the (2,1) mode, as a function of the Newtonian separation r .

SpinTaylorT5 [46]. The EOB dynamics is evaluated using the SEOBNRv4PHM model [47]. The evolutions are initialized using spins and orbital frequency extracted from the NR simulations at t_0 . PN evolutions are terminated near the Schwarzschild innermost stable circular orbit (ISCO), which is located at $r = 6M$. Although the EOB model includes merger and ringdown, we evaluate it only until the NR frequency at which a common horizon forms.

Figure 3 compares the spin perturbation angles $\Delta\theta_{1,2}$ of the PN, EOB, and NR evolutions for each of the four aligned-spin configurations with $\theta_{\text{pert}} = 5^\circ$. Notably, we find that both the PN and EOB evolutions reproduce the unstable growth of up-down binaries (bottom row of Fig. 3). As expected, the faithfulness of approximate evolutions decreases at smaller separations, where the gravitational interaction is strongest and highly nonlinear. The Schwarzschild ISCO provides a simple proxy to characterize this transition (gray markers in Fig. 3).

Overall, we report discrepancies in $\Delta\theta_{1,2}$ between NR and PN/EOB of $\simeq 5^\circ$ for up-down and $\simeq 1^\circ$ for the three stable configurations (up-up, down-down, and down-up). EOB evolutions appear to reproduce the full NR dynamics more accurately. This is perhaps expected since the EOB framework receives NR information, although it is important to point out that its calibration does not make use of simulations with precessing spins. More specifically, in the up-down case, while the EOB curve follows the orbit-averaged value of the NR evolution, the PN curve tracks the minimum of each orbital cycle. Neither method matches the orbital timescale modulations present in the NR simulations (cf. Refs. [40,48] for related work).

C. Impact of NR resolution error

The results presented so far were based on our high-resolution NR simulations. We now investigate the impact of the numerical resolution error on our results by comparing the output of our low- and high-resolution simulations (cf. Sec. II). Figure 4 compares two of the main quantities

of interest in this work: the spin perturbation angle $\Delta\theta_1$ ($\Delta\theta_2$ is qualitatively similar) and the (2,1) GW mode, for the up-down case with $\theta_{\text{pert}} = 10^\circ$. The agreement between the two resolutions slowly degrades as the binary approaches merger, accumulating a dephasing of ~ 0.9 rad over a phase evolution of ~ 290 rad in the (2,1) mode. This suggests that even higher resolution simulations might be necessary to fully capture the fine details of the instability. However, all the key features, such as the growth of the instability in $\Delta\theta_{1,2}$, magnitude of the orbital timescale oscillations in $\Delta\theta_{1,2}$, and the growth of the (2,1) mode, are well captured and qualitatively similar between the two sets of runs.

IV. CONCLUSIONS

We presented the first NR simulations of aligned-spin binary BHs undergoing a precessional instability, verifying that previous PN predictions [23] hold in the strong-field regime of GR. The instability occurs for binary BHs in the up-down configuration, where the spin of the heavier (lighter) BH is aligned (antialigned) with the orbital angular momentum. Initialized with a small spin misalignment, as the binary evolves, the instability causes the spins to tilt dramatically from their initial configuration, achieving misalignments up to $\sim 90^\circ$ at merger. In order to observe this precessional effect, the simulations we perform are necessarily long. Each consists of ~ 100 orbits and lasts for a time $\sim 65 \times 10^3 M$, putting them among the longest simulations in the SXS catalog [33,35], which previously had just 12 simulations with >90 orbits.

We perform three simulations of up-down binaries, each with mass ratio $q=0.9$, dimensionless spins $\chi_1 = \chi_2 = 0.8$, and increasing initial misalignment $\theta_{\text{pert}} = 1^\circ, 5^\circ, 10^\circ$. All three exhibit the unstable precession behavior, while the rate of growth of the instability increases with the initial misalignment. We show that the instability leaves a strong imprint on the subdominant modes of the GW signal in the up-down binaries, which can potentially be used to distinguish them from other sources. We repeat these

simulations for the other three aligned-spin configurations (up-up, down-down, and down-up) with the same parameters and show that they all remain stable, undergoing only small-angle precession oscillations.

We compare the results of our NR simulations against both PN [42] and EOB [47] evolutions. We find that both frameworks capture the occurrence and growth of the up-down instability. While the EOB dynamics is more accurate in predicting the precession-timescale oscillations, both methods fail to match the orbital-timescale modulations seen in NR. Current NR surrogate models [38,39] are unable to reproduce the up-down instability because they are trained on NR simulations that last only ~ 20 orbits, during which the instability does not have time to develop. It would be interesting to see if current surrogate techniques can indeed capture this instability if applied to hybridized EOB-NR waveforms (cf. Ref. [49] for work in this direction).

Through the up-down instability, binary BHs whose astrophysical formation leads to spins that are initially (anti)aligned with the orbital angular momentum can become misaligned and strongly precessing near merger. Whether current LIGO/Virgo parameter-estimation techniques can confidently identify unstable up-down binaries as such remains an open point of investigation.

ACKNOWLEDGMENTS

We thank Daria Gangardt, Serguei Ossokine, Ulrich Sperhake, and Richard O’Shaughnessy for useful discussions. V. V. is supported by a Klarman Fellowship at Cornell and National Science Foundation (NSF) Grants No. PHY-170212 and No. PHY-1708213 at Caltech. D. G. and M. M. are supported by European Union’s H2020 ERC Starting Grant No. 945155-GWmining and Royal Society Grant No. RGS-R2-202004. D. G. is supported by Leverhulme Trust Grant No. RPG-2019-350. V. V., M. A. S., and L. E. K. are supported by the Sherman Fairchild Foundation. M. A. S. is supported by NSF Grants No. PHY-2011961, No. PHY-2011968, and No. OAC-1931266 at Caltech. L. E. K. is supported NSF Grants No. PHY-1912081 and No. OAC-1931280 at Cornell. Simulations for this work were performed on the Frontera cluster [50], which is supported by the Texas Advanced Computing Center (TACC) at the University of Texas at Austin. Additional computational work was performed on the Wheeler cluster at Caltech, which is supported by the Sherman Fairchild Foundation and Caltech, the University of Birmingham BlueBEAR cluster, the Athena cluster at HPC Midlands+ funded by Engineering and Physical Sciences Research Council Grant No. EP/P020232/1, and the Maryland Advanced Research Computing Center (MARCC).

-
- [1] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs, *Phys. Rev. X* **9**, 031040 (2019).
 - [2] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GWTC-2: Compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run, *arXiv:2010.14527*.
 - [3] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Binary black hole population properties inferred from the first and second observing runs of Advanced LIGO and Advanced Virgo, *Astrophys. J. Lett.* **882**, L24 (2019).
 - [4] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Population properties of compact objects from the second LIGO-Virgo gravitational-wave transient catalog, *arXiv:2010.14533*.
 - [5] J. Roulet, T. Venumadhav, B. Zackay, L. Dai, and M. Zaldarriaga, Binary black hole mergers from LIGO/Virgo O1 and O2: Population inference combining confident and marginal events, *Phys. Rev. D* **102**, 123022 (2020).
 - [6] T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, Spin-induced orbital precession and its modulation of the gravitational waveforms from merging binaries, *Phys. Rev. D* **49**, 6274 (1994).
 - [7] L. E. Kidder, Coalescing binary systems of compact objects to post-Newtonian 5/2 order. 5. Spin effects, *Phys. Rev. D* **52**, 821 (1995).
 - [8] D. Gerosa, M. Kesden, E. Berti, R. O’Shaughnessy, and U. Sperhake, Resonant-plane locking and spin alignment in stellar-mass black-hole binaries: A diagnostic of compact-binary formation, *Phys. Rev. D* **87**, 104028 (2013).
 - [9] C. L. Rodriguez, M. Zevin, C. Pankow, V. Kalogera, and F. A. Rasio, Illuminating black hole binary formation channels with spins in Advanced LIGO, *Astrophys. J. Lett.* **832**, L2 (2016).
 - [10] S. Vitale, R. Lynch, R. Sturani, and P. Graff, Use of gravitational waves to probe the formation channels of compact binaries, *Classical Quantum Gravity* **34**, 03LT01 (2017).
 - [11] C. Talbot and E. Thrane, Determining the population properties of spinning black holes, *Phys. Rev. D* **96**, 023012 (2017).
 - [12] W. M. Farr, S. Stevenson, M. Coleman Miller, I. Mandel, B. Farr, and A. Vecchio, Distinguishing spin-aligned and isotropic black hole populations with gravitational waves, *Nature (London)* **548**, 426 (2017).
 - [13] S. Stevenson, C. P. L. Berry, and I. Mandel, Hierarchical analysis of gravitational-wave measurements of binary black hole spin-orbit misalignments, *Mon. Not. R. Astron. Soc.* **471**, 2801 (2017).

- [14] D. Gerosa, E. Berti, R. O’Shaughnessy, K. Belczynski, M. Kesden, D. Wysocki, and W. Gladysz, Spin orientations of merging black holes formed from the evolution of stellar binaries, *Phys. Rev. D* **98**, 084036 (2018).
- [15] K. Belczynski *et al.*, Evolutionary roads leading to low effective spins, high black hole masses, and O1/O2 rates for LIGO/Virgo binary black holes, *Astron. Astrophys.* **636**, A104 (2020).
- [16] V. Kalogera, Spin orbit misalignment in close binaries with two compact objects, *Astrophys. J.* **541**, 319 (2000).
- [17] Y. Yang, I. Bartos, Z. Haiman, B. Kocsis, Z. Marka, N. C. Stone, and S. Marka, AGN Disks harden the mass distribution of stellar-mass binary black hole mergers, *Astrophys. J.* **876**, 122 (2019).
- [18] B. McKernan, K. E. S. Ford, R. O’Shaughnessy, and D. Wysocki, Monte Carlo simulations of black hole mergers in AGN discs: Low χ_{eff} mergers and predictions for LIGO, *Mon. Not. R. Astron. Soc.* **494**, 1203 (2020).
- [19] I. Mandel and R. O’Shaughnessy, Compact binary coalescences in the band of ground-based gravitational-wave detectors, *Classical Quantum Gravity* **27**, 114007 (2010).
- [20] M. C. Miller and J. H. Krolik, Alignment of supermassive black hole binary orbits and spins, *Astrophys. J.* **774**, 43 (2013).
- [21] A. Sesana, E. Barausse, M. Dotti, and E. M. Rossi, Linking the spin evolution of massive black holes to galaxy kinematics, *Astrophys. J.* **794**, 104 (2014).
- [22] D. Gerosa, B. Veronesi, G. Lodato, and G. Rosotti, Spin alignment and differential accretion in merging black hole binaries, *Mon. Not. R. Astron. Soc.* **451**, 3941 (2015).
- [23] D. Gerosa, M. Kesden, R. O’Shaughnessy, A. Klein, E. Berti, U. Sperhake, and D. Trifirò, Precessional Instability in Binary Black Holes with Aligned Spins, *Phys. Rev. Lett.* **115**, 141102 (2015).
- [24] C. O. Lousto and J. Healy, Unstable flip-flopping spinning binary black holes, *Phys. Rev. D* **93**, 124074 (2016).
- [25] M. Mould and D. Gerosa, Endpoint of the up-down instability in precessing binary black holes, *Phys. Rev. D* **101**, 124037 (2020).
- [26] J. M. Bellovary, M.-M. M. Low, B. McKernan, and K. E. S. Ford, Migration traps in disks around supermassive black holes, *Astrophys. J. Lett.* **819**, L17 (2016).
- [27] I. Bartos, B. Kocsis, Z. Haiman, and S. Márka, Rapid and bright stellar-mass binary black hole mergers in active galactic nuclei, *Astrophys. J.* **835**, 165 (2017).
- [28] N. C. Stone, B. D. Metzger, and Z. Haiman, Assisted inspirals of stellar mass black holes embedded in AGN discs: Solving the final au problem, *Mon. Not. R. Astron. Soc.* **464**, 946 (2017).
- [29] B. McKernan *et al.*, Constraining stellar-mass black hole mergers in AGN disks detectable with LIGO, *Astrophys. J.* **866**, 66 (2018).
- [30] Y. Yang *et al.*, Hierarchical Black Hole Mergers in Active Galactic Nuclei, *Phys. Rev. Lett.* **123**, 181101 (2019).
- [31] M. Kesden, D. Gerosa, R. O’Shaughnessy, E. Berti, and U. Sperhake, Effective Potentials and Morphological Transitions for Binary Black-Hole Spin Precession, *Phys. Rev. Lett.* **114**, 081103 (2015).
- [32] D. Gerosa, M. Kesden, U. Sperhake, E. Berti, and R. O’Shaughnessy, Multi-timescale analysis of phase transitions in precessing black-hole binaries, *Phys. Rev. D* **92**, 064016 (2015).
- [33] M. Boyle *et al.*, The SXS Collaboration catalog of binary black hole simulations, *Classical Quantum Gravity* **36**, 195006 (2019).
- [34] B. Szilágyi, J. Blackman, A. Buonanno, A. Taracchini, H. P. Pfeiffer, M. A. Scheel, T. Chu, L. E. Kidder, and Y. Pan, Approaching the Post-Newtonian Regime with Numerical Relativity: A Compact-Object Binary Simulation Spanning 350 Gravitational-Wave Cycles, *Phys. Rev. Lett.* **115**, 031102 (2015).
- [35] Simulating eXtreme Spacetimes, <http://www.black-holes.org/>.
- [36] M. Boyle and A. H. Mroue, Extrapolating gravitational-wave data from numerical simulations, *Phys. Rev. D* **80**, 124045 (2009).
- [37] M. Boyle, Transformations of asymptotic gravitational-wave data, *Phys. Rev. D* **93**, 084031 (2016).
- [38] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, D. Gerosa, L. C. Stein, L. E. Kidder, and H. P. Pfeiffer, Surrogate models for precessing binary black hole simulations with unequal masses, *Phys. Rev. Research* **1**, 033015 (2019).
- [39] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, C. D. Ott, M. Boyle, L. E. Kidder, H. P. Pfeiffer, and B. Szilágyi, Numerical relativity waveform surrogate model for generically precessing binary black hole mergers, *Phys. Rev. D* **96**, 024058 (2017).
- [40] S. Ossokine, M. Boyle, L. E. Kidder, H. P. Pfeiffer, M. A. Scheel, and B. Szilágyi, Comparing post-Newtonian and numerical-relativity precession dynamics, *Phys. Rev. D* **92**, 104028 (2015).
- [41] X. Zhao, M. Kesden, and D. Gerosa, Nutational resonances, transitional precession, and precession-averaged evolution in binary black-hole systems, *Phys. Rev. D* **96**, 024007 (2017).
- [42] A. Buonanno, B. Iyer, E. Ochsner, Y. Pan, and B. S. Sathyaprakash, Comparison of post-Newtonian templates for compact binary inspiral signals in gravitational-wave detectors, *Phys. Rev. D* **80**, 084043 (2009).
- [43] S. Isoyama, R. Sturani, and H. Nakano, Post-Newtonian templates for gravitational waves from compact binary inspirals, [arXiv:2012.01350](https://arxiv.org/abs/2012.01350).
- [44] LIGO and Virgo Collaborations, LIGO Algorithm Library - LALSuite, 2018.
- [45] A. Buonanno, Y.-b. Chen, and M. Vallisneri, Detecting gravitational waves from precessing binaries of spinning compact objects: Adiabatic limit, *Phys. Rev. D* **67**, 104025 (2003); Erratum, *Phys. Rev. D* **74**, 029904 (2006).
- [46] P. Ajith, Addressing the spin question in gravitational-wave searches: Waveform templates for inspiralling compact binaries with nonprecessing spins, *Phys. Rev. D* **84**, 084037 (2011).
- [47] S. Ossokine *et al.*, Multipolar effective-one-body waveforms for precessing binary black holes: Construction and validation, *Phys. Rev. D* **102**, 044055 (2020).
- [48] R. Owen, A. S. Fox, J. A. Freiberg, and T. P. Jacques, Black hole spin axis in numerical relativity, *Phys. Rev. D* **99**, 084031 (2019).

- [49] V. Varma, S.E. Field, M.A. Scheel, J. Blackman, L.E. Kidder, and H.P. Pfeiffer, Surrogate model of hybridized numerical relativity binary black hole waveforms, *Phys. Rev. D* **99**, 064045 (2019).
- [50] D. Stanzione, J. West, R.T. Evans, T. Minyard, O. Ghattas, and D.K. Panda, Frontera: The evolution of leadership computing at the national science foundation, in *Practice and Experience in Advanced Research Computing, PEARC '20* (Association for Computing Machinery, New York, 2020), pp. 106–111.