Novel Ringdown Amplitude-Phase Consistency Test

Xisco Jiménez Forteza,[1,2] Swetha Bhagwat,[3,4] Sumit Kumar,[1,2] and Paolo Pani[3]
[1]Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Callinstraße 38, 30167 Hannover, Germany
[2]Leibniz Universität Hannover, 30167 Hannover, Germany
[3]Dipartimento di Fisica, "Sapienza" Università di Roma e Sezione INFN Roma1, Piazzale Aldo Moro 5, 00185 Roma, Italy
[4]Institute for Gravitational Wave Astronomy and School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom

(Received 8 June 2022; revised 13 October 2022; accepted 15 November 2022; published 11 January 2023)

The ringdown signal emitted during a binary black hole coalescence can be modeled as a linear superposition of the characteristic damped modes of the remnant black hole that get excited during the merger phase. While checking the consistency of the measured frequencies and damping times against the Kerr BH spectrum predicted by general relativity (GR) is a cornerstone of strong-field tests of gravity, the consistency of measured excitation amplitudes and phases have been largely left unexplored. For a nonprecessing, quasicircular binary black hole merger, we find that GR predicts a narrow region in the space of mode amplitude ratio and phase difference, independently of the spin of the binary components. Using this unexpected result, we develop a new null test of strong-field gravity which demands that the measured amplitudes and phases of different ringdown modes should lie within this narrow region predicted by GR. We call this the *amplitude-phase consistency test* and introduce a procedure for performing it using information from the ringdown signal. Lastly, we apply this test to the GW190521 event, using the multimodal ringdown parameters inferred by Capano *et al.* [arXiv:2105.05238]. While ringdown measurements errors for this event are large, we show that GW190521 is consistent with the amplitude-phase consistency test. Our test is particularly well suited for accommodating multiple loud ringdown detections as those expected in the near future, and can be used complementarily to standard black-hole spectroscopy as a proxy for modified gravity, compact objects other than black holes, binary precession and eccentricity.

DOI: 10.1103/PhysRevLett.130.021001

Introduction.—A binary black hole (BBH) ringdown is the gravitational-wave (GW) signal emitted as the remnant black hole (BH) formed during a BBH coalescence relaxes towards a stationary configuration [1–4]. The Kerr metric [5–8] in Einstein's general theory of relativity (GR) uniquely describes this final state. The ringdown phase is modeled as the evolution of perturbations (set up during the pre-merger stage) on the Kerr metric of the remnant BH. The GW signal emitted is well approximated as a linear superposition of countably infinite quasinormal modes (QNMs), i.e., exponentially damped sinusoid modes with discrete characteristic complex frequencies, which are the eigenvalues of the radial and angular Teukolsky's equations [9–12]. Each mode is characterized by its frequency f_{lmn}, damping time τ_{lmn}, excitation amplitude A_{lmn}, and phase ϕ_{lmn}, where the integers (l, m) identify the angular dependence of the mode, whereas $n = 0, 1, 2, \ldots$ is the overtone number [see Eq. (2) for details].

While the frequencies and damping times are solely determined by the remnant's mass and spin, the perturbation condition setup prior to ringdown phase regulate the mode excitation, namely, the amplitudes and phases. For a BBH coalescence, there is an intrinsic relation between the initial binary's parameters and the perturbation condition setup for ringdown, which determine the magnitude of A_{lmn} and ϕ_{lmn}. QNM amplitudes (i.e., excitation factors) account for spacetime's geometry during merger, providing the initial data for ringdown perturbations. GW ringdown models implicitly incorporate this into A_{lmn} and ϕ_{lmn}. We can estimate A_{lmn} and ϕ_{lmn} using numerical relativity (NR) simulations [13–16] corresponding to a set of BBH masses and spins. Thus, ringdown allows us to check two key predictions of GR in the strong-field regime—(a) the consistency of measured QNM spectrum in the ringdown to the expected Kerr spectrum and, (b) the compatibility of measured mode excitation factors with the prediction obtained from NR BBH coalescences in GR, i.e., consistency with (pre-)merger nonlinear dynamics.

While the former is the focus of a traditional BH spectroscopy, here we concentrate on the latter prospect

and devise a novel test of GR called the ringdown amplitude-phase consistency (APC) test. The APC test is based on the observation that, after a suitable normalization, only a narrow region in the mode amplitude-phase space is allowed for a BBH ringdown in GR. However, the BBH ringdown amplitudes and phases in modified GR [17–21] need not be constrained to lie in this region. We also expect a similar situation when the components of binary system are not Kerr BHs [22] (e.g., in neutron-star or more exotic boson-star [23–26] coalescences) within GR. Note that in both these scenarios, the remnant *can* still be a Kerr BH; so while it could pass the standard BH spectroscopy tests, the QNM amplitudes and phases can provide a way to distinguish the event from a GR BBH coalescence based on the nature of the merger.

A key virtue of this test is that it does not require information from the inspiral phase other than the binary extrinsic parameters. It is therefore particularly well suited for massive BBHs, where the inspiral is short and the parameter estimation of the binary intrinsic parameters (e.g., the mass ratio and spins) is uncertain, jeopardizing the accuracy of inspiral-merger-ringdown (IMR) consistency tests [27]. GW190521 is one such event [28]. It is also the only event that has been reported to show presence of measurable subdominant angular-mode parameters [29]. Furthermore, the parameter estimation for this event yields a primary component mass lying in the pair-instability supernova BH mass gap. This could be an indication for exotic alternatives [30–32] and the data do not exclude some of these possibilities. Thus, while our main scope is to devise the APC test in general, we find GW190521 a particularly interesting test bed because if such exotic scenarios are possible, they can affect the QNM amplitudes and phases and manifest as a violation of this null test.

Amplitude-phase space of a BBH ringdown.—Our aim here is to show that the mode amplitude-phase space corresponding to a BBH ringdown in GR is constrained to a narrow region. For this, we must first extract the amplitudes and phases of several modes by fitting the NR simulation as a function of BBH masses and spins. While we mostly focus on nonprecessing, quasicircular BBHs for simplicity, later on we shall also discuss the effect of spin misalignment and eccentricity.

We use the dominant ($l = m = 2,\ n = 0$) mode as a baseline and work with the intrinsic (independent of sky position, distance, polarization ψ) amplitude ratio and phase difference defined as

$$A^R_{lmn} \equiv \frac{A_{lmn}}{A_{220}}, \qquad \delta\phi_{lmn} \equiv \frac{m}{2}\phi_{220} - \phi_{lmn}, \tag{1}$$

respectively [33]. $\delta\phi_{lmn}$ is defined such that it removes the degeneracy between ϕ_{lmn} and the coalescence phase φ (see Supplemental Material [34] and Ref. [35]). The ringdown waveform can be analytically written down as

$$h_+ + ih_\times = A_{220}\sum_{lmn}(e^{-i\frac{m}{2}\phi_{220}}A^R_{lmn}e^{i\delta\phi_{lmn}}S_{lmn}(\iota,\varphi) \times e^{i2\pi f_{lmn}t}e^{-t/\tau_{lmn}}), \tag{2}$$

where $S_{lmn}(\iota,\varphi)$ are the spin two-weighted spheroidal harmonics. For simplicity, we approximate them to spherical harmonic functions [36] and as discussed later, this introduces a systematic error no larger than 1% for spinning remnants with $a_f \lesssim 0.9$ and for the modes considered in this Letter. For quasicircular BBH mergers, A^R_{lmn} and $\delta\phi_{lmn}$ are functions of the binary mass ratio $q = m_1/m_2 \geq 1$ and spins. We fit for A^R_{lmn} and $\delta\phi_{lmn}$ modewise for a set of 142 nonprecessing and noneccentric NR simulations from the SXS catalog [37] (other catalogs [38,39] are considered in the Supplemental Material [34]). Our simulation set is the same used to calibrate the `SEOBNRv4HM` model [40] and spans $q \in [1, 10]$ and the z component of the spins $\chi_{1,2} \in [-0.9, 0.9]$. We fit for the $(l, m, n) \in \{(3, 3, 0), (2, 1, 0), (4, 4, 0)\}$ modes and provide ready-to-use analytical fits as a function of q and a post-Newtonian motivated effective spin $\chi_{\rm pheno}(q, \chi_1, \chi_2)$ whose explicit form depends on the mode under consideration [34] and Ref. [41]. In the nonspinning limit, these fits are consistent with those obtained in [40,42–45], while they correct the results given in [46] for the phase difference. We refer the reader to Supplemental Material [34] for further details.

The shaded bands in Fig. 1 marks off the region in the $(A^R_{lmn}, \delta\phi_{lmn})$ space allowed by GR BBH ringdowns as obtained by fitting the amplitudes and phases to our NR dataset using Eq. (2). The dashed curves correspond to $\chi_{1,2} = 0$, wherein A^R_{lmn} and $\delta\phi_{lmn}$ parametrically depend only on q. The shaded region around the dashed curve quantifies the effects of nonzero $\chi_{1,2}$. Given the parity and polarization conventions of the odd or even m modes used in the SXS waveforms, we plot $\mathrm{mod}(\delta\phi_{lmn}, \pi)$ for the (2, 1, 0) and (3, 3, 0) modes and $\mathrm{mod}(\delta\phi_{lmn}, 2\pi)$ for the (4, 4, 0) mode (see Supplemental Material [34]). For the (3, 3, 0) mode, we find that the GR admissible region in the $(A^R_{lmn}, \delta\phi_{lmn})$ space is remarkably narrow. This happens because the effects of the progenitor spins are small for the (3, 3, 0) mode. Indeed, $\delta\phi_{330} \in [2.68, \pi]$ and $A^R_{330} \in [0, 0.42]$ for our entire dataset. The latter range of values observed for the amplitude ratio A^R_{330} is about a factor of 2 larger than what was observed in the inspiral regime [42]. This shows that higher harmonics are more excited during the highly dynamical merger and ringdown regimes. However, the effect of nonzero $\chi_{1,2}$ on A^R_{lmn} and $\delta\phi_{lmn}$ is substantial for the (4, 4, 0) mode and (2, 1, 0) mode; this leads to a less-constricted GR-permissible region for these modes. The small bump of the (2, 1, 0) shaded area originates from restricting the amplitude A^R_{210} to be positive at small mass ratio and positive spin (see Ref. [47] and the Supplemental Material [34]).

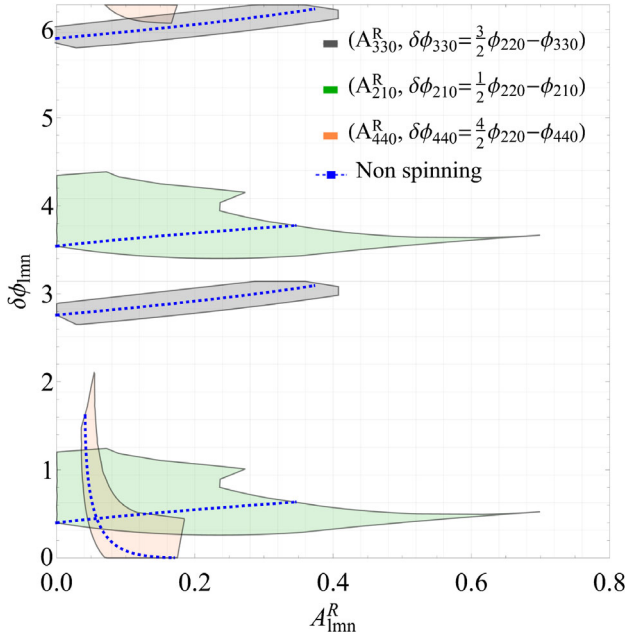

FIG. 1. The regions of the $(A^R_{lmn}, \delta\phi_{lmn})$ plane allowed for BBHs in GR for various modes. They are estimated by fitting the parameters $(A^R_{lmn}, \delta\phi_{lmn})$ in Eq. (2) to 142 nonprecessing waveforms, covering the parameter space $q \in [1, 10]$ and $\chi_{1,2} \in [-0.9, 0.9]$. Given the parity of the odd or even m modes, we plot $\mathrm{mod}(\delta\phi_{lmn}, \pi)$ for the (2, 1, 0) and (3, 3, 0) modes and $\mathrm{mod}(\delta\phi_{lmn}, 2\pi)$ for the (4, 4, 0) mode. The dashed blue curves within each region correspond to the nonspinning limit.

The ringdown APC test.—All nonprecessing, quasicircular BBH ringdown governed by GR must have A^R_{lmn} and $\delta\phi_{lmn}$ within in the GR permissible region on the $(A^R_{lmn}, \delta\phi_{lmn})$ space, i.e., measurements of A^R_{lmn} and $\delta\phi_{lmn}$ must lie within the shaded bands in Fig. 1. We use this to devise our APC null test of strong gravity wherein we check if the measured mode amplitudes and phases in a ringdown signal lie in the narrow GR permissible region. Practically, one can check whether the posterior distributions of the estimated A^R_{lmn} and $\delta\phi_{lmn}$ have significant support in the allowed region. Since all quasicircular BBH mergers must satisfy this constraint, the test is naturally extendable to incorporate a population of observations.

Note that while implementing this test, it is crucial to account for the uncertainty in the ringdown start time with respect to the global peak time t^p, i.e., the time at which the strain amplitude $|h(t) = \Sigma_{lmn} h_{lmn}|$ maximizes [46,48,49]. Estimating the time t^p_{lmn} at which each $|h_{lmn}|$ mode peaks from NR fits, we can shift each mode by $\Delta t = t^p - t^p_{220}$, with $t^p_{220} < t^p_{lm0}$ [46,50]. This induces a correction:

$$A^R_{lmn}(t^{\mathrm{p}}) = A^R_{lmn}(t^{\mathrm{p}}_{220}) e^{\frac{\Delta t}{\tau_{220}} - \frac{\Delta t}{\tau_{lmn}}}. \quad (3)$$

Notice that the τ_{lmn} are comparable for all the modes studied here. For example, for a BH with $a_f \approx 0.85$ (consistent with GW190521 [28,29]) $\tau_{330}/\tau_{220} \sim 0.986$, $\tau_{440}/\tau_{220} \sim 0.974$, and $\tau_{210}/\tau_{220} \sim 0.956$. Here a conservative choice $\Delta t \approx 10M$ translates to a $\approx(1, 2, 3)\%$ correction for the amplitude of the (3,3,0), (4,4,0), (2,1,0) modes, respectively; this is well below the current statistical uncertainties [29]. This conservative choice corresponds to $\Delta t \approx 2(t^p_{330} - t^p_{220}) \approx (t^p_{210} - t^p_{220})$ [46] and so sets a conservative upper bound on Δt.

A similar correction needs to be accounted for the intrinsic phase, $\phi_{lmn}(t^p) = \phi_{lmn}(t^p_{220}) + \omega_{lmn}\Delta t$. Using Eq. (1), this translates to

$$\delta\phi_{lmn}(t^p) = \delta\phi_{lmn}(t^p_{220}) + \left(\frac{m}{2}\omega_{220} - \omega_{lmn}\right)\Delta t. \quad (4)$$

Since $\omega_{lmn} \approx (l/2)\omega_{220}$ for $l = m$ modes (note that this is an exact result in the eikonal $l = m \gg 1$ limit [11,51]), the phase correction induced by Δt is $\approx$10% and $\approx$20% for the (3,3,0) and (4,4,0) mode, respectively, when we assume $\Delta t = 10M$ (see Supplemental Material [34]). This is a conservative choice for the (3,3,0) mode as $\Delta t \lesssim 5M$ [46]; here we expect a systematic uncertainty no larger than 4%. Therefore, for the $l = m$ modes, the NR phase fits can be compared to the measured posteriors inference as

$$\delta\phi_{lmn}(t^p) \approx \delta\phi_{lmn}(t^p_{220}) \quad \forall\, l = m. \quad (5)$$

However note that for the (2,1,0) mode the last term in Eq. (4) yields a non-negligible uncertainty.

Application on GW190521.—We exemplify our test on GW190521, the only GW event with reported subdominant angular mode in the ringdown [29]. The total signal-to-noise ratio (SNR) of this event is $\rho \approx 14$, of which $\rho \approx 12$ comes solely from the ringdown phase. This is a consequence of large total source mass of this binary system ($M_{\mathrm{tot}} = 151^{+29}_{-17} M_\odot$ [28] and the system is a convenient choice to demonstrate a proof-of-concept of the APC test.

In Fig. 2, we first mark the GR-permissible region obtained by the NR fits as a gray band. Note that this region also accounts for the following uncertainties: the 1σ deviations on the best fit results, the $\sim$1% and $\sim$4% uncertainties on A^R_{330} and $\delta\phi_{330}$ caused by ambiguity in ringdown start time, and the errors due to approximating the spheroidal harmonics as spherical harmonics. Next, we project the 67% and 95% credible regions of the measured posterior distribution for A^R_{330} and $\delta\phi_{330}$ from [29] on to the $(A^R_{330}, \delta\phi_{330})$ space. The posteriors obtained by assuming that ringdown has GR predicted Kerr QNM spectrum corresponds to the orange contour whereas the blue contour corresponds to the case where the QNM frequencies and damping times are allowed to vary freely. Interestingly, this more agnostic assumption does not deteriorate the confidence region significantly. The black dot-dashed contour provides the 95% credible region obtained by fixing the polarization angle ψ to the maximum likelihood value

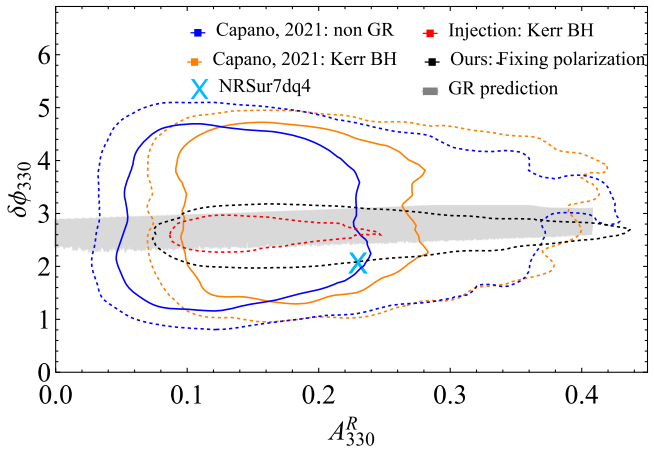

FIG. 2. Ringdown APC test applied to GW190521. The amplitude ratio A^R_{330} and the phase difference $\delta\phi_{330}$ obtained by various parameter estimations done on the data are compared to the GR allowed region in the $(A^R_{330}, \delta\phi_{330})$ space (shaded gray area). The solid and dashed contours, respectively, represent the 67% and 95% credibility contours obtained by Bayesian parameter estimation. The orange and blue contours correspond to the posteriors obtained by fixing the GR QNM spectrum and allowing deviations from the Kerr QNMs, respectively. The black dot-dashed contour provides the 95% credible region obtained as in [29] but fixing the polarization to the maximum likelihood value ($\psi = 0$) given by the IMR analysis [52]. The red dot-dashed contour provides the 95% credible region for the NR injection SXS:0258 consistent with GW190521 and with SNR $\rho = 30$. The cyan cross corresponds to the $(A^R_{330}, \delta\phi_{330})$ fit to the maximum likelihood waveform obtained using the NRSur7dq4 [56] waveform approximant, which includes precession and binary parameters consistent to GW190521 [57].

estimated from the full IMR analysis in [52]. This is similar to fixing right ascension and declination as done in [29] (see Refs. [53–55] for a discussion on fixing these parameters in BH spectroscopy tests). Estimating the polarization angle independently helps to break the degeneracy between ψ and $\delta\phi_{lmn}$. ψ can be estimated from the inspiral-merger regime, while the intrinsic dependence of $\delta\phi_{lmn}$ on q and $\chi_{\rm pheno}$ arises in the ringdown phase (see the Supplemental Material [34]).

Lastly, we inject the NR waveform SXS:0258 into Gaussian noise at SNR $\rho = 30$ for a three-detector (LIGO-Hanford, LIGO-Livingston and Virgo) configuration and perform a parameters estimation using the PyCBC inference library [58]. This numerical waveform has parameters compatible with GW190521 [59]) but has been injected with twice the SNR of GW190521 to estimate the quality of the test achievable in the case of higher SNR events. The red dot-dashed contour in Fig. 2 denotes the 95% credible region obtained for this case. We perform the parameter estimation at $t = t^p + 15$ ms (see the Supplemental Material [34]). We note that, as expected, the confidence region shrinks and the test is significantly more accurate with higher SNR.

We presented the main result of applying APC to GW19021 in Fig. 2. We find that the 1σ credible interval obtained in [29] has a substantial support in the gray GR-permissible region marked on the $(A^R_{lmn}, \delta\phi_{lmn})$ space. Therefore, we conclude that the mode amplitude and phases measured in GW19021's ringdown are compatible with the GR BBH predictions and this event passes the APC test.

Discussion.—The APC test provides a novel strategy for testing GR using the ringdown mode excitations. NR waveforms of BBH mergers give accurate empirical relations between A^R_{lmn} and $\delta\phi_{lmn}$ as functions of the binary's mass ratio and spins. We found that only a narrow strip in the $(A^R_{lmn}, \delta\phi_{lmn})$ space is admissible for ringdown modes of quasi-circular BBHs within GR. We build the APC test based on this feature and present a proof-of-concept implementation of this test on GW190521. We find that the 1σ posterior distributions of A^R_{lmn} and $\delta\phi_{lmn}$ obtained in [29] for this event has substantial support in the GR-permissible region, showing that GW190521 passes the APC test. Furthermore, we verify that the combined modelling uncertainties for the (3,3,0) mode are well below the statistical uncertainties of the current GW observations. Overall, for the (3,3,0) mode the total systematic errors accumulated from the fit, ringdown start time, and spherical-harmonic approximation may reach a value ∼13% for both A^R_{lmn} and $\delta\phi_{lmn}$ (see Supplemental Material [34]). For GW190521, the marginalized 1σ statistical uncertainties on these quantities are ∼100%—much larger than the systematic deviations accumulated from our fit uncertainties. While this holds for any $l = m$ mode, the phase fits of $l \neq m$ modes are non-negligibly affected by the shift of the peak time; this makes $l \neq m$ modes not optimal for the proposed implementation of the test. However, the situation might change for louder detections as those routinely expected in the third-generation era [60,61], in which case systematic errors of the fit might limit the accuracy of the APC test, unless the quality of NR waveforms improves. On the other hand, the measurements of the polarization angle ψ (which is degenerate with the phase difference $\delta\phi_{lmn}$) and of other binary's intrinsic parameters are expected to improve as more interferometers are added to the network, or through an electromagnetic counterpart, and will anyway improve with third-generation detectors, therefore allowing for a more accurate test.

We have focused on quasicircular binaries with aligned spins, although we can extend a similar concept to build a more generic test. In the Supplemental Material [34] we show that the effect of eccentricity on the fits of $\delta\phi_{330}$ (A^R_{330}) is non-negligible only when $e \gtrsim 0.3$ ($e \gtrsim 0.6$). Therefore, the current implementation of the test is robust to mild eccentricities. By comparing the posterior distribution for A^R_{330} shown in Fig. 2 to the eccentricity fits, we have obtained a mild bound of $e \lesssim 0.9$ at the 95% level on the eccentricity of GW19052 [62]. Note also that several

works using the full IMR analysis on GW190521 have reported a moderately high effective precession spin parameter [63], $\chi_p = 0.68^{+0.25}_{-0.37}$ (although waveform systematics and prior choices significantly affect the posterior estimates of this event [57]). Interestingly, when we fit for $\delta\phi_{330}$ and A^R_{330} corresponding to the maximum likelihood waveform including precession [56], we notice that the effect of precession for GW190521 is within the measurement errors for this event. However, the best fit is marginally outside the gray shaded area (corresponding to the nonprecessing scenario) in Fig. 2. Because of the large statistical error, the systematic effect of neglecting the spin precession does not affect GW190521 significantly. Therefore, we can use GW190521 as a proof of concept for the APC test, and future louder events could be used to constrain the binary precession independently from ringdown measurements only. A detailed examination of the effects of precession in the ringdown is an involved problem and requires a dedicated study.

Interpretation and extensions.—As with any null-hypothesis consistency test, its violation suggests a departure from the adopted baseline assumptions, and so a violation of the null-test could have various origins. We spell out the viable interpretation when an event does not pass the APC test—(a) Most conservatively, it might be evidence for mis-modelling the signal, e.g., presence of strong spin precession or large eccentricity in the BBH. (b) It could be because the observed ringdown is not BH coalescence; note this does not preclude the remnant from being a standard Kerr BH. We expect the coalescence of massive neutron stars, boson stars [23–26], and other exotic compact objects [22] to produce QNMs consistent with Kerr BHs in GR. However, the QNM amplitudes and phases can be different from GR as the merger dynamics could be modified, these will therefore fail the APC test while being consistent with a traditional Kerr BH spectroscopy. (c) Finally and most radically, it could be because the underlying coalescence dynamics is not governed by GR. Disentangling these possibilities calls for a generalization of our fits to incorporate features like precession and eccentricity, louder ringdown detections, and detecting a population of them. For instance, if the violation of the test were due to not including eccentricity or precession in our fits, out of a population of ringdowns only a subgroup would be violating it. However, if GR dynamics were under question, there could be a ubiquitous violation of the test. In this context, although measurement errors are large, it is relevant that GW190521 passes the APC test. It would be interesting to assess whether this is in tension with alternative explanations for this event, e.g., a Proca star merger [30], by fitting the ringdown amplitudes and phases for Proca star merger waveforms and performing Bayesian model selection between the two hypotheses [30].

While we focused on the ringdown signal with prior knowledge of the binary's extrinsic parameters (estimated either from the IMR analysis or from another independent sky localization), a variant of this test would be to estimate the initial binaries parameters with ringdown and check for consistency with IMR analysis. In principle one could invert the $A^R_{lmn}(q,\chi_{\rm pheno})$ and $\delta\phi_{lmn}(q,\chi_{\rm pheno})$ relations to infer an estimate of the mass ratio (and spins) from the QNM excitations. However, owing to the mild dependence of $\delta\phi_{lmn}$ on the binary parameters the quality of this test is expected to be rather poor. A more promising avenue is to neglect the phases and use only the amplitude ratios of several subleading QNMs. We discuss this in the Supplemental Material [34]. This is interesting for GW190521-like systems where the short signal duration and low SNR in the premerger part leads to controversial and model-dependent inference on the binary parameters [52,57,64,65]. Also, higher sensitivity at low frequency (as expected for third-generation detectors) will improve this test significantly (see also [66] for a conceptual framework in this direction). Overall, the APC test provides an excellent arena to complement standard BH spectroscopy tests in the strong-gravity regime, especially for the next-generation detectors.

We acknowledge the Max Planck Gesellschaft for support, and we are grateful to the Atlas cluster computing team at AEI Hannover for their help. The authors are specially thankful to Lionel London, Cecilio García-Quiros, and Juan Calderon-Bustillo for the invaluable discussions and further clarifications about the NR phase alignment and phase conventions. X. Jimenez is also thankful to P. Mourier for the useful discussions about the correspondence of the fit and parameter-estimation results. S. B. is supported by the UKRI Stephen Hawking Fellowship, Grant Ref. EP/W005727. P. P. acknowledges financial support provided under the European Union's H2020 ERC, Starting Grant agreement no. DarkGRA–757480. We also acknowledge support under the MIUR PRIN (Grant 2020KR4KN2 "String Theory as a bridge between Gauge Theories and Quantum Gravity") and FARE (GW-NEXT, CUP: B84I20000100001, 2020KR4KN2) programmes, and from the Amaldi Research Center funded by the MIUR program "Dipartimento di Eccellenza" (CUP: B81I18001170001).

[1] S. Chandrasekhar and S. L. Detweiler, Proc. R. Soc. A **344**, 441 (1975).
[2] S. A. Teukolsky, Astrophys. J. **185**, 635 (1973).
[3] W. H. Press and S. A. Teukolsky, Astrophys. J. **185**, 649 (1973).
[4] S. A. Teukolsky and W. H. Press, Astrophys. J. **193**, 443 (1974).
[5] J. D. Bekenstein, Phys. Rev. D **7**, 2333 (1973).
[6] B. Carter, Phys. Rev. Lett. **26**, 331 (1971).
[7] S. W. Hawking, Phys. Rev. Lett. **26**, 1344 (1971).

[8] S. W. Hawking and J. B. Hartle, Commun. Math. Phys. **27**, 283 (1972).
[9] E. Berti, V. Cardoso, and A. O. Starinets, Classical Quantum Gravity **26**, 163001 (2009).
[10] S. Chandrasekhar, *The Mathematical Theory of Black Holes*, Oxford Classic Texts in the Physical Sciences (Oxford University Press, New York, 1985).
[11] V. Ferrari and B. Mashhoon, Phys. Rev. D **30**, 295 (1984).
[12] K. D. Kokkotas and B. G. Schmidt, Living Rev. Relativity **2**, 2 (1999).
[13] X. Jiménez-Forteza, D. Keitel, S. Husa, M. Hannam, S. Khan, and M. Pürrer, Phys. Rev. D **95**, 064024 (2017).
[14] C. O. Lousto and J. Healy, Phys. Rev. D **93**, 124074 (2016).
[15] F. Hofmann, E. Barausse, and L. Rezzolla, Astrophys. J. **825**, L19 (2016).
[16] D. Keitel, X. Jimenez-Forteza *et al.*, Phys. Rev. D **96**, 024006 (2017).
[17] M. Okounkova, Phys. Rev. D **102**, 084046 (2020).
[18] M. Okounkova, L. C. Stein, J. Moxon, M. A. Scheel, and S. A. Teukolsky, Phys. Rev. D **101**, 104016 (2020).
[19] W. E. East and J. L. Ripley, Phys. Rev. Lett. **127**, 101102 (2021).
[20] M. Elley, H. O. Silva, H. Witek, and N. Yunes, Phys. Rev. D **106**, 044018 (2022).
[21] H. Lim, G. Khanna, and S. A. Hughes, Phys. Rev. D **105**, 124030 (2022).
[22] V. Cardoso and P. Pani, Living Rev. Relativity **22**, 4 (2019).
[23] C. Palenzuela, P. Pani, M. Bezares, V. Cardoso, L. Lehner, and S. Liebling, Phys. Rev. D **96**, 104058 (2017).
[24] T. Helfer, E. A. Lim, M. A. G. Garcia, and M. A. Amin, Phys. Rev. D **99**, 044046 (2019).
[25] M. Bezares and C. Palenzuela, Classical Quantum Gravity **35**, 234002 (2018).
[26] M. Bezares, M. Bošković, S. Liebling, C. Palenzuela, P. Pani, and E. Barausse, Phys. Rev. D **105**, 064067 (2022).
[27] R. Abbott *et al.* (LIGO Scientific, Virgo, KAGRA Collaborations), arXiv:2112.06861.
[28] R. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Phys. Rev. Lett. **125**, 101102 (2020).
[29] C. D. Capano, M. Cabero, J. Westerweck, J. Abedi, S. Kastha, A. H. Nitz, A. B. Nielsen, and B. Krishnan, arXiv:2105.05238.
[30] J. C. Bustillo, N. Sanchis-Gual, A. Torres-Forné, J. A. Font, A. Vajpeyi, R. Smith, C. Herdeiro, E. Radu, and S. H. W. Leong, Phys. Rev. Lett. **126**, 081101 (2021).
[31] M. Fishbach and D. E. Holz, Astrophys. J. Lett. **904**, L26 (2020).
[32] J. C. Bustillo, S. H. W. Leong, K. Chandra, B. McKernan, and K. E. S. Ford, arXiv:2112.12481.
[33] J. Calderón Bustillo, A. Bohé, S. Husa, A. M. Sintes, M. Hannam, and M. Pürrer, arXiv:1501.00918.
[34] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.130.021001 for a discussion on the NR fits and the phase and starting time conventions.
[35] V. Baibhav, E. Berti, and V. Cardoso, Phys. Rev. D **101**, 084053 (2020).
[36] E. Berti and A. Klein, Phys. Rev. D **90**, 064012 (2014).
[37] The SXS Collaboration, SXS Gravitational Waveform Database (2019), http://www.black-holes.org/waveforms/.
[38] M. Campanelli, J. Healy, C. Lousto, and Y. Zlochower, CCRG@RIT Catalog of Numerical Simulations (2022), https://ccrgpages.rit.edu/~RITCatalog/.
[39] K. Jani, J. Healy, J. A. Clark, L. London, P. Laguna, and D. Shoemaker, Georgia Tech catalog of binary black hole simulations (2016).
[40] R. Cotesta, A. Buonanno, A. Bohé, A. Taracchini, I. Hinder, and S. Ossokine, Phys. Rev. D **98**, 084028 (2018).
[41] E. Barausse, A. Buonanno, S. A. Hughes, G. Khanna, S. O'Sullivan, and Y. Pan, Phys. Rev. D **85**, 024046 (2012).
[42] S. Borhanian, K. G. Arun, H. P. Pfeiffer, and B. S. Sathyaprakash, Classical Quantum Gravity **37**, 065006 (2020).
[43] L. T. London, Phys. Rev. D **102**, 084052 (2020).
[44] I. Ota and C. Chirenti, Phys. Rev. D **101**, 104005 (2020).
[45] L. London, D. Shoemaker, and J. Healy, Phys. Rev. D **90**, 124032 (2014); **94**, 069902(E) (2016).
[46] X. Jiménez Forteza, S. Bhagwat, P. Pani, and V. Ferrari, Phys. Rev. D **102**, 044053 (2020).
[47] C. García-Quirós, M. Colleoni, S. Husa, H. Estellés, G. Pratten, A. Ramos-Buades, M. Mateu-Lucena, and R. Jaume, Phys. Rev. D **102**, 064002 (2020).
[48] S. Bhagwat, X. J. Forteza, P. Pani, and V. Ferrari, Phys. Rev. D **101**, 044033 (2020).
[49] X. J. Forteza and P. Mourier, arXiv:2107.11829.
[50] H. Estellés, S. Husa, M. Colleoni, D. Keitel, M. Mateu-Lucena, C. García-Quirós, A. Ramos-Buades, and A. Borchers, Phys. Rev. D **105**, 084039 (2022).
[51] V. Cardoso, A. S. Miranda, E. Berti, H. Witek, and V. T. Zanchin, Phys. Rev. D **79**, 064016 (2009).
[52] A. H. Nitz and C. D. Capano, Astrophys. J. Lett. **907**, L9 (2021).
[53] R. Cotesta, G. Carullo, E. Berti, and V. Cardoso, Phys. Rev. Lett. **129**, 111102 (2022).
[54] M. Isi and W. M. Farr, arXiv:2202.02941.
[55] E. Finch and C. J. Moore, Phys. Rev. D **106**, 043005 (2022).
[56] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, D. Gerosa, L. C. Stein, L. E. Kidder, and H. P. Pfeiffer, Phys. Rev. Res. **1**, 033015 (2019).
[57] H. Estellés *et al.*, Astrophys. J. **924**, 79 (2022).
[58] C. M. Biwer, C. D. Capano, S. De, M. Cabero, D. A. Brown, A. H. Nitz, and V. Raymond, Publ. Astron. Soc. Pac. **131**, 024503 (2019).
[59] The progenitor parameters of the simulation are $q = 2$ and $\chi_{\rm pheno} = 0.48$, with dimensionless final spin $a_f = 0.84$ and we have scaled the waveform to the total detector-frame mass consistent with the ringdown. The injected waveform includes the following modes $(2, \pm 2), (2, \pm 1), (3, \pm 3), (3, \pm 2), (4, \pm 4), (4, \pm 3)$. We fix the polarization to the maximum likelihood value $\psi = 0$, consistent with [52].
[60] M. Maggiore *et al.*, J. Cosmol. Astropart. Phys. 03 (2020) 050.
[61] S. Bhagwat, C. Pacilio, E. Barausse, and P. Pani, Phys. Rev. D **105**, 124063 (2022).
[62] B. P. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), Astrophys. J. **883**, 149 (2019).
[63] P. Schmidt, F. Ohme, and M. Hannam, Phys. Rev. D **91**, 024043 (2015).
[64] S. Kastha, C. D. Capano, J. Westerweck, M. Cabero, B. Krishnan, and A. B. Nielsen, Phys. Rev. D **105**, 064042 (2022).
[65] R. Abbott *et al.* (LIGO Scientific, Virgo, KAGRA Collaborations), arXiv:2111.03606.
[66] S. Bhagwat, C. Pacilio, E. Barausse, and P. Pani, arXiv:2201.00023.