Black-hole ringdown with templates capturing spin precession: A reanalysis of GW190521

Chiara Anselmo,[1,2,*] Costantino Pacilio,[1,2,3,4] and Davide Gerosa[1,2]

[1]*Dipartimento di Fisica "G. Occhialini," Universitá degli Studi di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy*

[2]*INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy*

[3]*Institute of Astrophysics, FORTH, GR-71110, Heraklion, Greece*

[4]*INFN, Sezione di Roma, Piazzale Aldo Moro 2, 00185, Roma, Italy*

(Received 4 December 2025; accepted 13 May 2026; published 15 June 2026)

The ringdown stage of a binary black-hole merger provides a clean probe of strong-field gravity, as it can be modeled with minimal assumptions. The quasi–normal-mode frequencies encode the mass and spin of the Kerr black hole remnant, while the mode excitation depends on the progenitor binary. In this paper, we implement a recently developed amplitude model that captures spin precession in a simulation-based inference pipeline that specifically targets ringdown signals. We present a reanalysis of GW190521—a short-duration, merger-dominated event with conflicting interpretations. Spin-aligned and precessing analyses at two ringdown start times show that precession induces modest but systematic shifts in inferred parameters and subdominant mode amplitudes, although such ringdown-only analyses provide no strong evidence for precession. Our results demonstrate the feasibility of physics-informed precessing ringdown modeling, paving the way for the identification of spin precession in gravitational-wave events using solely their ringdown stages, where waveform systematics are expected to be substantially less prominent.

DOI: 10.1103/52cd-81m5

I. INTRODUCTION

The ringdown is the final stage of a binary black-hole (BH) coalescence, during which the merged remnant emits gravitational waves (GWs) as it settles into a stationary state. In general relativity (GR), this equilibrium configuration is uniquely described by a Kerr BH [1,2]. Small perturbations of the Kerr spacetime are described by Teukolsky's equations [3,4], whose solutions correspond to the BH's quasinormal modes (QNMs)—exponentially damped sinusoids characterized by a discrete complex frequency spectrum [5–9]. More generally, the full linear response of a perturbed BH is composed of three contributions: a prompt response determined by the initial data, a superposition of QNMs, and a late-time power-law tail arising from backscattering off the spacetime curvature [10]. In the context of BH mergers, the QNM frequencies depend solely on the remnant's mass and spin, making them a powerful probe of GR through BH spectroscopy [10–13].

Since the first GW detection [14], the LIGO/Virgo/KAGRA (LVK) network has observed about two hundred BH coalescences [15–19]. Accurate characterization of the ringdown is especially important for events dominated by the merger and ringdown stages, such as GW190521 [20] and GW231123 [21], where the postmerger signal enables a robust study of the source properties. Ringdown-focused analyses are also motivated by waveform systematics: small discrepancies between waveform models can mimic apparent deviations from GR [22,23], and analyses based on different models have sometimes produced inconsistent results [21,24–29]. Restricting attention to the ringdown, which can be described analytically and depends on fewer modeling assumptions, provides a complementary and often cleaner approach.

While the QNM frequencies are determined entirely by the remnant's mass and spin, the excitation of the modes, namely their amplitudes and phases, depends on the progenitor binary configuration [30]. Phenomenological fits to numerical-relativity (NR) simulations map binary parameters, such as mass ratio and spins, to QNM amplitudes, effectively linking premerger dynamics to postmerger observables. Simple models employ post-Newtonian–inspired formulas or results from the extreme-mass-ratio limit, while more sophisticated approaches, including Gaussian process regression (GPR) surrogate models, capture complex dependencies (e.g., spin precession and eccentricity) and provide

*Contact author: c.anselmo@campus.unimib.it

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

uncertainty estimates. Several NR-calibrated fits are now available for spin-aligned binaries [30–37], and recent GPR models predict QNM amplitudes and phases in that regime [38,39]. The first NR-calibrated approximant for precessing systems was introduced shortly after by some of us [40].

We report the first integration of precessing ringdown templates—constructed from the fits of Ref. [40]—into a GW parameter estimation framework, and demonstrate their application to GW190521. This event, characterized by a short inspiral and a merger-ringdown–dominated signal, has been extensively studied in the context of overtones, higher harmonics, eccentricity, and precession, both within full inspiral-merger-ringdown (IMR) [20,41–45] and ringdown-only [28,29,45] frameworks. Our model includes higher harmonics and precession effects, enabling a consistent and physically motivated comparison with previous analyses.

Parameter estimation is performed with simulation-based inference (SBI), also known as likelihood-free Bayesian inference [46,47]. Its application to GW data analysis has been demonstrated both in the frequency [48] and in the time [49] domain. In particular, our SBI framework extends the methodology of Ref. [49], developed for time-domain ringdown analyses, to precessing templates with physical amplitude models.

We compare a spin-aligned and a precessing ringdown model, assess their performance for two ringdown start times, and quantify the evidence for precession using Bayesian model comparison. Incorporating precession leads to modest but systematic shifts in the inferred binary parameters and QNM amplitudes. The ringdown alone does not provide strong evidence in favor of precession in GW190521. Overall, our findings demonstrate the viability of physics-informed precessing ringdown models and establish a foundation for future high signal-to-noise ratio (SNR) studies of BH spectroscopy.

The remainder of this paper is organized as follows. Section II introduces the ringdown model. Section III describes the analysis methods. Results are presented in Sec. IV and discussed in Sec. V. We conclude in Sec. VI.

II. RINGDOWN MODEL

A. Waveform definition and symmetries

We model the ringdown waveform as a superposition of damped sinusoids corresponding to the QNMs of the remnant BH. This approach isolates the QNM component of the linear response, which is expected to dominate the signal at times sufficiently after the strain peak. The prompt response and the late-time tail are therefore not included, as they are not expected to be statistically significant at currently accessible SNRs in this regime [7,50–52].

Each (ℓ, m, n) mode contributes a strain component $h_{\ell mn}$, given by

$$h_{\ell mn} = \frac{M}{d_L} A_{\ell mn} e^{-i(2\pi f_{\ell mn} t + \phi_{\ell mn})} e^{-t/\tau_{\ell mn}} \times {}_{-2}S_{\ell mn}(\iota, \varphi, M_f \chi_f \omega_{\ell mn}). \quad (1)$$

Here, $M = m_1 + m_2$ is the redshifted total mass of the initial binary BH system, d_L is the luminosity distance, and $A_{\ell mn}$ are the amplitudes of excitation of each mode. The frequencies $f_{\ell mn}$ and damping times $\tau_{\ell mn}$ correspond to the real and imaginary part of the complex mode frequencies, respectively,

$$\omega_{\ell mn} = \omega_R - i\omega_I = 2\pi f_{\ell mn} - \frac{i}{\tau_{\ell mn}}. \quad (2)$$

These are solutions of Teukolsky's equations [2] and only depend on the mass M_f and spin χ_f of the remnant. Finally, ${}_{-2}S_{\ell mn}(\iota, \varphi, M_f \chi_f \omega_{\ell mn})$ denotes the spin-weighted spheroidal harmonic of that mode, where ι and φ are the inclination and azimuthal angles.

The total waveform is obtained by summing over all relevant modes,

$$\tilde{h} = h_+ - i h_\times = \sum_{\ell mn} h_{\ell mn}, \quad (3)$$

where h_+ and $h_\times$ denote the plus and cross GW polarizations, respectively. For each (ℓ, m, n), Teukolsky's equations admit two sets of solutions, the prograde and retrograde modes [36]. In the above expressions and throughout this work, we neglect the contribution from the retrograde modes. We discuss the motivations and limits of this choice in Sec. V.

The measured strain h is the projection of each of the two GW polarizations $h_{+,\times}$ onto the detector through the antenna pattern functions $F_{+,\times}$,

$$h = F_+ h_+ + F_\times h_\times. \quad (4)$$

The pattern functions depend on the sky location (α, δ) and on the polarization angle ψ. They also vary in time, within the timescale of the Earth's rotation; since, for stellar-mass BHs, the typical ringdown signal duration is much shorter ($\sim$0.1 s), we treat $F_{+,\times}$ as constant.

We further simplify the waveform in two ways. First, we replace the spin-weighted spheroidal harmonics ${}_{-2}S_{\ell mn}$ with spin-weighted spherical harmonics ${}_{-2}Y_{\ell m}$, thus neglecting spherical-spheroidal mode mixing [53]. Since we restrict ourselves to the set of modes $\{(2, \pm 2, 0), (3, \pm 3, 0), (2, \pm 1, 0)\}$, which are only mildly affected by mode mixing [40], we do not expect our analysis to be significantly biased.

Second, we observe that some parameters are degenerate and can be set to zero without loss of generality. Since the angular dependence of the spherical harmonics factorizes

as a phase,

$$ {}_{-2}Y_{\ell m}(\iota,\varphi) \equiv {}_{-2}Y_{\ell m}(\iota)e^{im\varphi}, \tag{5}$$

we can reabsorb φ in a redefinition of the mode phases, $\phi_{\ell mn} \to \phi_{\ell mn} - m\varphi$. Similarly, a change in the polarization angle ψ corresponds to a rotation of h by 2ψ [54,55]. Therefore, we will also reabsorb ψ in a redefinition of the mode phases. Mind that these degeneracies hold as long as $\phi_{\ell mn}$ are treated as free parameters: if $\phi_{\ell mn}$ were instead modeled as functions of the binary parameters, (φ,ψ) would become free parameters of the model.

Finally, only in the case of spin-aligned models, we exploit the following reflection symmetry between $+|m|$ and $-|m|$ modes:

$$ A_{\ell-|m|n} = (-1)^{\ell} A_{\ell+|m|n}. \tag{6}$$

B. Amplitude and remnant fits

We express the QNMs as functions of the binary parameters. We employ two sets of fits, for the mode amplitudes and for the remnant mass and spin, respectively.

1. Mode amplitudes

For the QNM amplitudes, we adopt the seven-dimensional model `Prec7dq10_20M`, first presented in Ref. [40] by some of the authors and implemented in the POSTMERGER PYTHON package [56]. This model was obtained by training a GPR [57] on a subset of precessing, quasicircular NR simulations from the 2019 Simulating eXtreme Spacetimes (SXS) catalog [58]. GPR is a nonparametric Bayesian method that represents data as a distribution over functions, providing both mean predictions and associated uncertainties.

The model parameter space is defined by $\{\delta,\boldsymbol{\chi}_1,\boldsymbol{\chi}_2\}$, where

$$ \delta = \frac{q-1}{q+1}, \qquad q = \frac{m_1}{m_2} \geq 1 \tag{7}$$

and the spins are defined in Cartesian coordinates at the innermost stable circular orbit (ISCO). The fit is calibrated up to mass ratio $q = 10$.[1] In the spin-aligned limit, the parameter space reduces to three dimensions, $\{\delta,\chi_{1z},\chi_{2z}\}$, by setting the remaining spin components to zero.

The initial binary parameters are provided to the model in the *physical frame*, whose z axis is aligned with the orbital angular momentum. Instead, the resulting amplitudes are expressed in a coordinate system whose z axis is aligned with the spin direction of the remnant BH, called the *ringdown frame* [59]. This distinction is relevant only in the precessing case and is handled as described in Sec. II C.

Fitted amplitudes are reported at the reference time

$$ t_{\rm start} = 20M + t_{\rm EMOP}, \tag{8}$$

where $t_{\rm EMOP}$ is the time of maximum energy-maximized orthogonal projection (EMOP) [60] of the simulation, extended here to all $(2,m)$ modes (see Appendix B of Ref. [40]). For inference purposes, it is instead convenient to define the ringdown starting time relative to the waveform peak, $t_{\rm peak}$, and shift the amplitudes accordingly. The model provides both an estimate of $\Delta t_{\rm EMOP} = t_{\rm EMOP} - t_{\rm peak}$ and the possibility to shift the amplitudes with respect to $t_{\rm EMOP}$. Our procedure is described in Sec. III C.

2. Final mass and spin

To enforce consistency with the amplitude fits, we train a GPR model for final masses and spins in terms of the binary parameters at the ISCO frequency. We use the same training set of NR simulations as in Ref. [40]. This implementation has not been presented elsewhere; details and performances are provided in Appendix A.

In particular, going to the ringdown frame only requires knowing the final spin magnitude $\chi_f = |\boldsymbol{\chi}_f|$ and the tilt angle

$$ \theta_f = \arccos\left(\frac{\chi_{fz}}{\chi_f}\right), \tag{9}$$

where χ_{fz} is the spin component parallel to the orbital angular momentum. Therefore, we do not need to provide a surrogate for all the three components of $\boldsymbol{\chi}_f$. We find it convenient to fit the components χ_{fz} and $\chi_{f\perp}$, where $\chi_{f\perp}$ is the projection of the final spin in the orbital plane, defined by

$$ \chi_f = \sqrt{(\chi_{fz})^2 + (\chi_{f\perp})^2}. \tag{10}$$

This choice guarantees that both quantities remain within physical ranges and avoids unphysical extrapolations. An initial attempt to fit (χ_f,χ_{fz}) separately resulted in occasional unphysical extrapolations with $\chi_{fz} > \chi_f$. Fitting $(\chi_{fz},\chi_{f\perp})$ removes this issue by construction.

C. Wigner rotation of spherical harmonics

Since the mode amplitudes are expressed in the ringdown frame, we write the total strain as [61]

$$ \tilde{h} = \sum_{\ell mn} h^{(\rm RD)}_{\ell mn} Y^{(\rm RD)}_{\ell m}(\iota',\varphi'), \tag{11}$$

where the spherical harmonics in the ringdown frame are related to those in the physical frame through a Wigner

[1] Only a few simulations exist at $q = 10$, so we consider the calibration robust up to $q = 8$.

rotation,

$$Y^{(\mathrm{RD})}_{\ell m}(\iota',\varphi') = \sum_{\mu} D^{\ell}_{m\mu}(\alpha,\beta,\gamma) Y^{(\mathrm{phys})}_{\ell\mu}(\iota,\varphi). \quad (12)$$

Here, (α,β,γ) are the Euler angles parametrizing the rotation of the coordinate axes, with $\alpha,\gamma\in[0,2\pi]$ and $\beta\in[0,\pi]$. Following Wigner and Fano [62], the component rotations are performed in the sequence: a rotation through γ about the z axis, then through β about the y axis, and finally through α about the z axis.[2]

The corresponding Wigner D matrix is given by

$$D^{\ell}_{m\mu}(\alpha,\beta,\gamma) = e^{im\gamma} d^{\ell}_{m\mu}(\beta) e^{i\mu\alpha}, \quad (13)$$

where $d^{\ell}_{m\mu}(\beta)$ is the Wigner's (small) d matrix, see Eq. (15.27) of Ref. [62],

$$d^{\ell}_{m\mu}(\beta) = [(\ell+m)!(\ell-m)!(\ell+\mu)!(\ell-\mu)!]^{\frac{1}{2}} \times \sum_{s=s_{\min}}^{s_{\max}} \left[\frac{(-1)^s (\cos\frac{\beta}{2})^{2\ell+\mu-m-2s} (\sin\frac{\beta}{2})^{m-\mu+2s}}{(\ell+\mu-s)!s!(m-\mu+s)!(\ell-m-s)!}\right], \quad (14)$$

with summation limits $s_{\min}=\max(0,\mu-m)$ and $s_{\max}=\min(\ell+\mu,\ell-m)$. Note that our definition of $d^{\ell}_{m\mu}(\beta)$ maps into $d^{\ell}_{m\mu}(-\beta)\equiv d^{\ell}_{\mu m}(\beta)$ when following the alternative conventions of Refs. [59,64].

The geometric interpretation of the Euler angles is as follows: the second angle, β, corresponds to the inclination θ_f between the z axis in the physical frame, parallel to the orbital angular momentum, and the final spin vector $\boldsymbol{\chi}_f$; the first angle, γ, can be reabsorbed into the QNM phases $\phi_{\ell mn}$; the third angle, α, is degenerate with the azimuthal angle φ [see Eq. (5)]. Moreover, the condition that φ' is reabsorbed in a redefinition of the QNM phases implies $\varphi=-\gamma$ in Eq. (12). As a result, we can simplify the rotation matrices so that only $\beta=\theta_f$ appears explicitly,

$$\tilde{h} = \sum_{\ell mn}\sum_{\mu} h^{(\mathrm{RD})}_{\ell mn} d^{\ell}_{m\mu}(\theta_f) Y^{(\mathrm{phys})}_{\ell\mu}(\iota,0). \quad (15)$$

III. METHODS

A. TSNPE

In addition to being an established Bayesian inference framework, SBI offers practical advantages for the present analysis compared to standard likelihood-based sampling approaches. In the current implementation of the POSTMERGER package, waveform evaluations are sufficiently expensive ($t_{\mathrm{eval}}\sim 1$ s) that repeated likelihood estimations within a traditional Markov chain Monte Carlo (MCMC) sampler would lead to a substantial computational cost at the scale required for this study. By contrast, SBI allows waveform simulations to be generated in large batches, amortizing the cost of forward-model evaluations and making the inference computationally tractable.

Among the broad class of SBI algorithms, we adopt truncated sequential neural posterior estimation (TSNPE) [65] as our inference method. In *neural posterior estimation*, parameters are sampled from the prior $\pi(\theta)$ to generate simulated data $x(\theta)$, and a neural density estimator $q_\phi(\theta|x)$ with learnable parameters ϕ is trained to directly approximate the posterior. *Sequential* neural posterior estimation [66,67] specializes this to the case of a particular observation x_{obs} and proceeds in multiple adaptive rounds: in the first round, training samples are drawn from the prior $\pi(\theta)$; in subsequent rounds, they are drawn from a proposal distribution $\tilde{\pi}(\theta)$ that is updated to be denser in regions of high posterior density $q_\phi(\theta|x_{\mathrm{obs}})$, the latter approximated from the previous round. In particular, TSNPE updates the prior with a sequential truncation scheme: at each round, the original training set is augmented with new simulations, drawn from the high-density region of the approximate posterior—see Refs. [38,65] for further details.

Our implementation of TSNPE closely follows that of Ref. [49], which uses the PyTorch-based [68] SBI package [69]. The neural density estimator is modeled as a neural spline flow [70–72]. For this paper, we have made minor updates to the neural-network architecture as detailed in Appendix D.

B. Setup

When analyzing ringdown, one needs to truncate the signal in order to isolate the postmerger part of the waveform. This truncation requires fixing the ringdown starting time. Since different sky locations correspond to different delays in the arrival time across the detector network, it is also necessary to fix the source sky location, ensuring that the same stretch of data is analyzed throughout the inference [73].[3] We adopt from Ref. [75] the waveform peak time t_{peak} at the Hanford detector (expressed in the GPS time scale) and the sky location of GW190521,

[2]Alternative conventions, such as active rotations or different rotation orderings, lead to different sign conventions; see, e.g., Chapter 4 of Ref. [63] for a comprehensive discussion. Our choice follows the passive Wigner convention of Ref. [62], which fixes the signs used below.

[3]While this issue commonly arises in time-domain analyses—where truncation can be implemented without spectral leakage—it would similarly affect frequency-domain analyses involving truncated signals. An alternative frequency-domain ringdown analysis that alleviates these difficulties has been proposed by Finch and Moore [74].

$$t_{\rm peak} = 1242442967.4287114,$$
$$\alpha = 0.16416072898840461 \text{ rad},$$
$$\delta = -1.1434284511661967 \text{ rad}. \quad (16)$$

These values were obtained from the IMR LVK analysis for GW190521 using the posterior samples corresponding to the NRSur7dq4 model. In particular, the starting time corresponds to the time that maximizes $h_+^2 + h_\times^2$, while the sky position is estimated as the median of the posterior samples.

The time delays at the other detectors are computed using the `detector` module from PyCBC [76].

We analyze the ringdown of GW190521 starting at 6 ms and 12 ms after $t_{\rm peak}$. For a binary with redshifted mass $\sim 250 M_\odot$, these times correspond to $\sim 5M$ and $\sim 10M$, respectively. While a starting time of 12 ms is more conservative with respect to the physical model, 6 ms is less conservative but has a higher SNR. Both times constitute an extrapolation with respect to the nominal calibration time of the amplitude fits, which are reported at $20M$ after $t_{\rm EMOP}$. This extrapolation serves to amplify the SNR and is not expected to lead to appreciable biases at low-to-moderate SNRs [77–79]. We find consistent recoveries of the parameters (total mass, mass ratio, spins, and luminosity distance) at both starting times. We interpret this consistency as an *a posteriori* internal validation of the robustness of our analysis. Moreover, at the SNRs considered here (see Sec. IV D), we expect the waveform template to be robust for mismatches $\mathcal{M} \lesssim 0.1$ [80], which is at least an order of magnitude higher than the typical mismatches reported in Ref. [40]. However, note that this assumes that there are no systematics between the template and the true waveforms; the role of systematics is discussed in Sec. V.

The duration of the analysis segment after each start time is set to 0.1 s for all detectors, as in Ref. [45]. We load the strain data using the `catalog` module from PyCBC and downsample them to 2048 Hz with the `decimate` method from `scipy.signal`. See Ref. [81] for recent developments on data conditioning techniques, and the Supplement to Ref. [82] (Appendix II B 2) for small differences that can arise when analyzing different signal lengths. We whiten the time-domain strain following the procedure described in Ref. [49] (see their Appendix A 1).

Our ringdown template includes the six prograde fundamental modes

$$(\ell, m, n) = \{(2, \pm 2, 0), (3, \pm 3, 0), (2, \pm 1, 0)\}. \quad (17)$$

The initial BH spins $\boldsymbol{\chi_i}$ are expressed in spherical coordinates, $\{\chi_i, \gamma_i, \phi_i\}$, where $\chi_i = |\boldsymbol{\chi_i}|$ is the spin magnitude and γ_i and ϕ_i denote the polar and azimuthal angles, respectively. The QNM frequencies, assuming GR, are determined from (M_f, χ_f) using the PYKERR package [83]. Altogether, the model depends on 16 parameters,

$$\theta = \{d_L, \cos\iota, M, q, \chi_i, \cos\gamma_i, \phi_i, \phi_{\ell mn}\}, \quad (18)$$

where $i = 1, 2$ and (ℓ, m, n) run over the included modes. In the spin-aligned limit, the spin vectors $\boldsymbol{\chi}_i$ reduce to their z components, χ_{iz}, lowering the number of dimensions to 12.

C. Priors

We infer the posterior distribution using priors that are uniform in the parameters and ranges listed in Table I. In particular, in the precessing model, we adopt priors that are uniform in the spin magnitudes χ_i and isotropic in orientation, as commonly done in GW astronomy; this corresponds to uniform priors in $\cos\gamma_i$. In the spin-aligned model, we instead use a uniform prior on the aligned spin components χ_{iz}. Note that, in the precessing model, uniform priors in χ_i and $\cos\gamma_i$ induce highly nonuniform priors on χ_{iz}. Therefore, to ensure consistency when comparing results with the posterior samples from the spin-aligned model, we reweight the latter so that their implied prior on χ_{iz} matches that of the isotropic case, as described in Appendix B 1.

For the masses, we adopt uniform priors in total mass M and mass ratio q, chosen for sampling efficiency. To consistently compare our results with IMR analyses, which assume priors uniform in the component masses (m_1, m_2), we reweight our posterior samples following the procedure detailed in Appendix B 2.

Importantly at inference time, for each proposed total mass M, the amplitude fits are consistently evaluated at a time

TABLE I. Prior ranges adopted for the parameter estimation. All parameters are assigned uniform priors over the intervals listed. The *Shared* block (top) lists parameters common to both the spin-aligned and precessing models, while the treatment of the spins differs between the two cases (center and bottom).

Parameter	Prior range
Shared	
$d_L^3 (\mathrm{Gpc}^3)$	$[10^{-3}, 10^3]$
$\cos\iota$	$[-1, 1]$
$M (M_\odot)$	[100, 500]
q	[1, 10]
$\phi_{\ell mn}$	$[0, 2\pi]$
Precessing model	
χ_i	[0, 0.99]
$\cos\gamma_i$	$[-1, 1]$
ϕ_i	$[0, 2\pi]$
Spin-aligned model	
χ_{iz}	$[-0.99, 0.99]$

$$t_{\rm eval} = t_{\rm peak} + \frac{t_{\rm start}}{M} = t_{\rm EMOP} - \Delta t_{\rm EMOP} + \frac{t_{\rm start}}{M} \quad (19)$$

when expressed in NR simulation units. Since the amplitude fits are calibrated at $t_{\rm EMOP} + 20$, this amounts to rescale the amplitudes by a factor $\exp\{t_{\rm start}/M - 20 - \Delta t_{\rm EMOP}\}$. In this way, the start time adapts consistently to the sampled value of M, while the physical onset of the ringdown remains identical across all realizations, ensuring internal consistency in the parameter estimation.

IV. RESULTS

We begin by presenting the posterior distributions for the binary parameters inferred with the spin-aligned and precessing ringdown models, followed by the corresponding QNM amplitudes derived from these posteriors. For completeness, we also report the posteriors for the remnant mass and spin in Appendix C, including results from this work and from other analyses. Throughout this Section, and unless otherwise specified, binary properties and QNM amplitudes are computed using the mean predictions of the GPR fits of Ref. [40]; the impact of including the GPR variances is examined separately in Appendix E. Finally, we compare the two models by computing their Bayes factor and evaluate the ringdown SNRs associated with the chosen start times. Modeling assumptions and potential limitations relevant to the interpretation of these results are discussed in Sec. V.

A. Binary parameters

Our parameter estimation yields a 16- (12-)dimensional posterior distribution in the precessing (spin-aligned) case. Figure 1 shows the joint posterior distributions for ringdown start times of 6 ms (lower triangle) and 12 ms (upper triangle), over the following parameters:

(i) $\{d_L, \cos\iota, M, q\}$, common to both models;
(ii) $\{\chi_{1z}, \chi_{2z}\}$, which are reweighted spin components in the spin-aligned case, and computed as $\chi_{iz} = \chi_i \cos\gamma_i$ in the precessing case; and
(iii) $\{|\chi_{1\perp}|, |\chi_{2\perp}|\}$, present only in the precessing model, and obtained from (χ_i, χ_{iz}) as $|\chi_{i\perp}| = \sqrt{(\chi_i)^2 - (\chi_{iz})^2}$.

We compare our results with the IMR parameter estimation reported by the LVK Collaboration [20,45], which makes use of the full signal and includes precession effects. We use the IMR samples produced with the NRSur7dq4 model [84], which was the preferred waveform approximant in the GW190521 LVK analyses [20,45]. We also include the results of Ref. [85], where the data from the LVK detector network were truncated at different times from the peak and analyzed in the time domain using the NRSur7dq4 model. Posterior samples are taken from the dataset at Ref. [86], using start times of $5M_{\rm ref}$ ($10M_{\rm ref}$) to match our 6 ms (12 ms) analysis ($M_{\rm ref} \simeq 1.27$ ms, see Sec. II of Ref. [85] for details). Since those samples were obtained with uniform priors on total mass M, mass ratio q, and luminosity distance d_L, they have been reweighted following the procedure described in Appendix B.

Across all analyses and start times, we find consistent trends in several key parameters. The inclination ι is similarly constrained in all models, although the IMR posterior for $\cos\iota$ is mildly bimodal. This feature does not appear in our runs, as we fix the sky location (α, δ) to values compatible with $\cos\iota \simeq 1$ in the IMR run, breaking the degeneracy. A similar behavior is observed in the ι posterior of Ref. [85], where the sky location is also fixed. The mass ratio q favors relatively symmetric binaries ($q \lesssim 2$), and the secondary spin χ_{2z} remains largely unconstrained, reflecting the isotropic prior.[4]

The most notable differences between the spin-aligned and precessing analyses arise in the luminosity distance d_L, total mass M, and primary aligned spin χ_{1z}. Including precession systematically shifts d_L to smaller values—smaller even than the IMR result. The d_L posterior from Ref. [85] is in good agreement with our precessing results at both start times. The shift in M follows the same trend, though in this case the precessing posterior moves closer to the IMR estimate. For this parameter, the results of Ref. [85] are consistent with the IMR analysis at 6 ms, but show a similar shift at 12 ms. The effect on χ_{1z} becomes significant in the 12 ms analysis, where precession reduces the inferred value relative to the spin-aligned model. Similarly, while consistent with the IMR analysis at 6 ms, the posterior from Ref. [85] shows better agreement with our precessing results at 12 ms. Finally, the in-plane spin components $(|\chi_{1\perp}|, |\chi_{2\perp}|)$ remain largely unconstrained, with the posterior favoring very small values more strongly than the IMR analysis. This is consistent with the results of Ref. [85] at both start times. The discrepancy between our analysis and those performed with NRSur7dq4 at 6 ms may be related to the absence of the (2, 2, 1) mode in our model [88,89]. More generally, we observe that, as the start time moves further away from the merger, the agreement between the two ringdown analyses (ours and that of Ref. [85]) improves.

B. Mode amplitudes

Starting from the posterior samples, we compute the QNM amplitudes $A_{\ell+|m|n}$ using the same phenomenological fits as in the main analysis. Figure 2 shows the resulting distributions of A_{220} as well as the relative amplitudes A_{330}/A_{220} and A_{210}/A_{220}. The lower (upper) triangle corresponds to the 6 ms (12 ms) case. For the 6 ms analysis, we also include the results of Ref. [28], obtained

[4] In comparing spins, we note that the IMR values are defined at $t = -\infty$, while ours correspond to ISCO. The resulting differences are subdominant compared to statistical uncertainties at current SNRs [87].

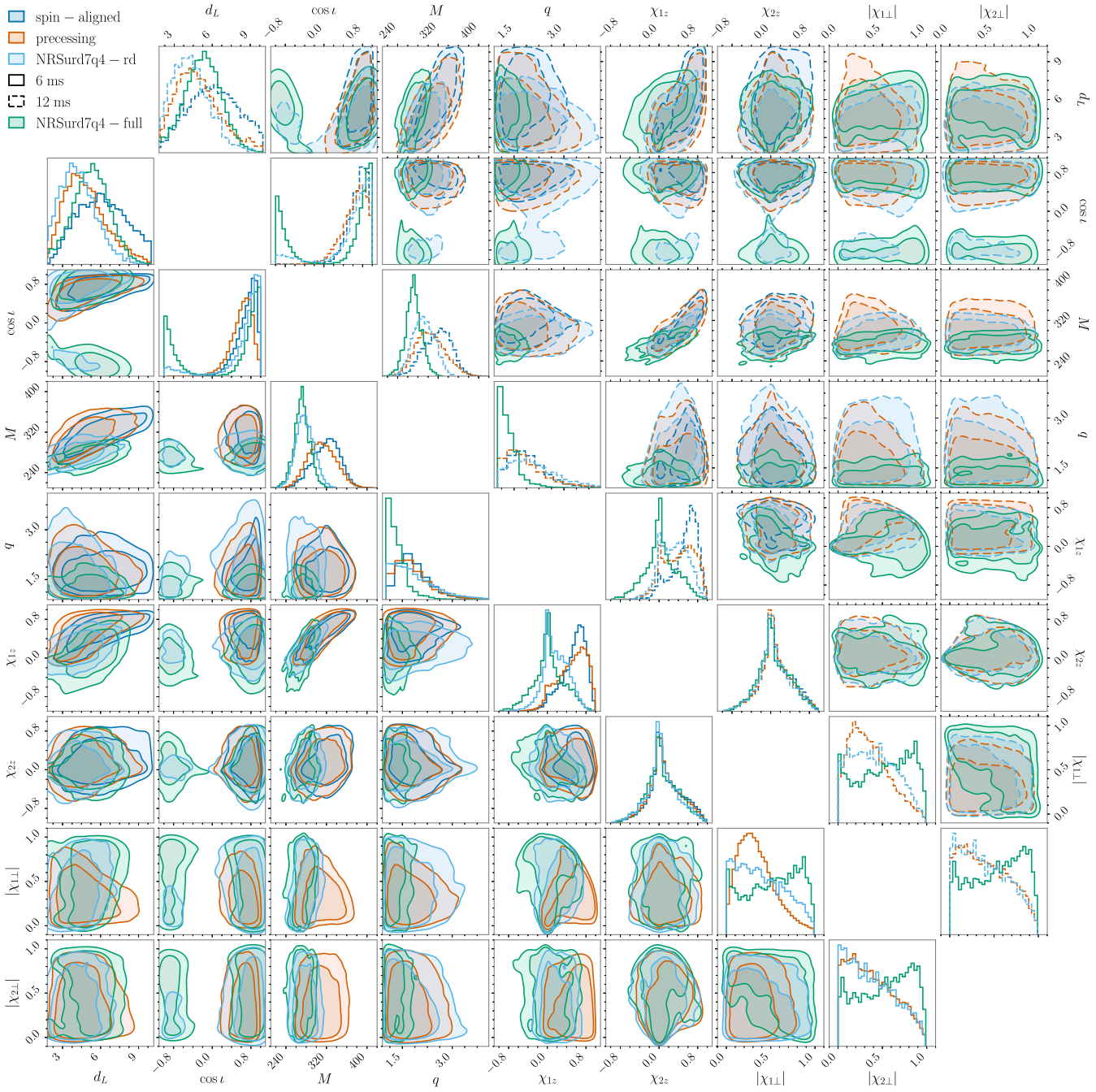

FIG. 1. Joint posterior distributions for the binary parameters inferred with the spin-aligned (blue) and precessing (orange) ringdown models, using the GPR mean predictions for the remnant mass and spin and for the QNM amplitudes. The lower (upper) triangular panels correspond to a ringdown start time of 6 ms (12 ms) after the strain peak, shown with solid (dashed) lines. The IMR posterior from the LVK analysis [20,45] (green), and the results from Ref. [85] (light blue), both obtained with the NRSur7dq4 model [84], are overlaid for comparison. All contours represent the 68% and 90% credible regions.

with a $(2,2,0)+(3,3,0)$ Kerr model and publicly available at Ref. [90].

The spin-aligned posteriors for A_{220} and A_{210}/A_{220} are narrower compared to the case where precession is included, reflecting the smaller parameter space. Both models yield comparable results for A_{220}, but the precessing posteriors favor lower A_{330} and higher A_{210}, indicating a modest enhancement of subdominant mode excitation due to precession.

The spin-aligned relative amplitude A_{330}/A_{220} agrees with the results of Ref. [28]. While that analysis is agnostic in the amplitudes, our results provide an independent, physics-informed indication of the relevance of the (3, 3, 0) mode in the ringdown of GW190521. Finally, for both start

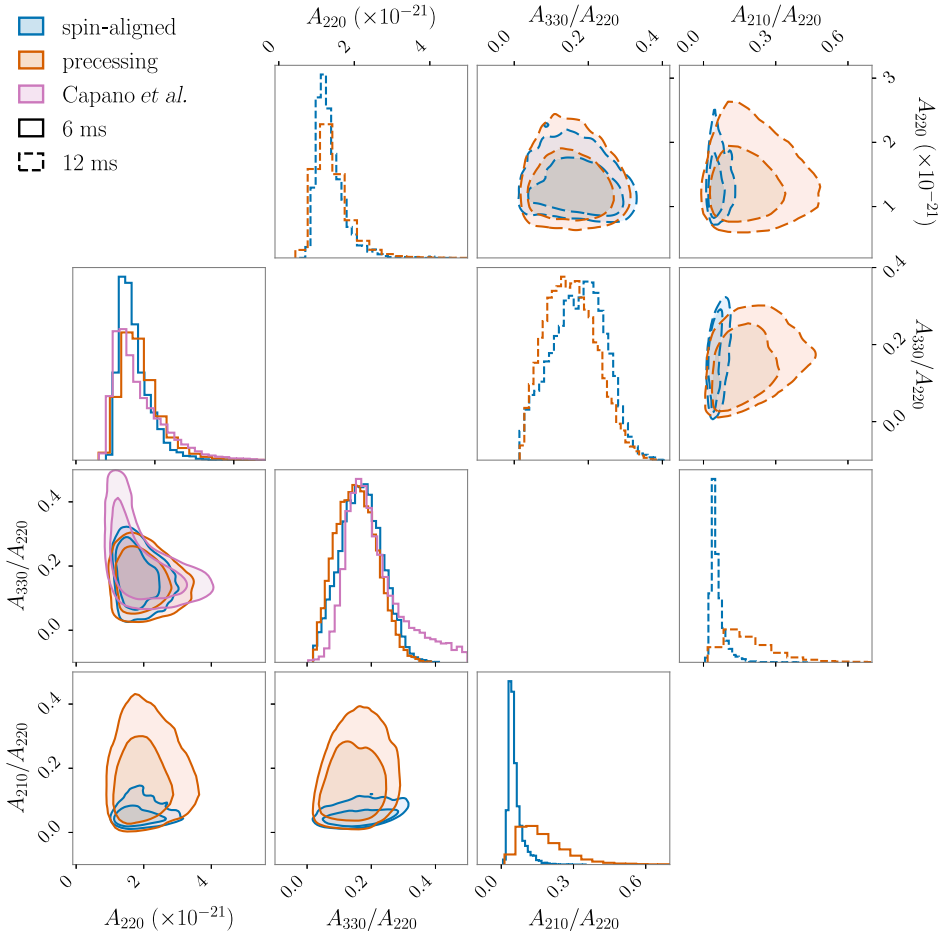

FIG. 2. Joint posterior distributions for the QNM amplitudes inferred from the spin-aligned (blue) and precessing (orange) ringdown models, using the GPR mean predictions for the remnant mass and spin and for the QNM amplitudes. We show the absolute amplitude of the dominant (2, 2, 0) mode, A_{220}, together with the relative amplitudes A_{330}/A_{220} and A_{210}/A_{220}. The lower (upper) triangular panels correspond to a ringdown start time of 6 ms (12 ms) after the strain peak, shown with solid (dashed) lines. For the 6 ms case, we overlay the results of Ref. [28] (pink), obtained using a $(2,2,0)+(3,3,0)$ Kerr model. Contours indicate the 68% and 90% credible regions.

times, the (2, 2, 0) mode remains dominant ($A_{210} < A_{220}$). This differs from the highly precessional interpretation of Ref. [29], which reported support for $A_{210} > A_{220}$ at early start times. In the time window considered here, our analysis favors the standard mode hierarchy, while still indicating that precession can leave measurable imprints on subdominant QNM amplitudes.

A direct comparison with other ringdown analyses in the literature is not always straightforward, due to differences in modeling choices and time windows. In particular, Ref. [29] does not provide posteriors obtained with templates that simultaneously include both the (3, 3, 0) and (2, 1, 0) modes, while Ref. [75] finds the strongest support for the (3, 3, 0) mode at premerger start times. These differences limit the extent to which quantitative comparisons can be drawn. The comparison with Ref. [28] is more direct. Although our template includes the (2, 1, 0) mode, the posterior of the corresponding reconstructed amplitude has non-negligible support at zero, effectively reducing its impact to the overall inference. In this sense, the two analyses are expected to be broadly consistent in the reconstruction of the (3, 3, 0) mode amplitude.

C. Bayes factors

To quantify the evidence for precession, we compute the Bayes factor $\mathcal{B}^{\rm SA}_{\rm P}$ between the spin-aligned model $M_{\rm SA}$ and the precessing model $M_{\rm P}$. The two models are nested,

$$M_{\rm SA} \equiv M_{\rm P} \wedge (\cos^2\gamma_1 = 1, \cos^2\gamma_2 = 1). \tag{20}$$

Therefore, we can evaluate the Bayes factor through the Savage-Dickey ratio

$$\mathcal{B}^{\rm SA}_{\rm P} = \sum_{a,b=\pm 1} \frac{p(\cos\gamma_1 = a, \cos\gamma_2 = b | x_{\rm obs}, M_{\rm P})}{\pi(\cos\gamma_1 = a, \cos\gamma_2 = b | M_{\rm P})} \frac{1}{4|a||b|}. \tag{21}$$

Since we sample isotropically in the component spins, the prior density at the denominator is exactly 1/4.

To evaluate the posterior density at the numerator, we use a two-dimensional kernel density estimator (KDE) on the posterior samples for $(\cos\gamma_1, \cos\gamma_2)$, with Gaussian kernels and a bandwidth of 0.05. To prevent density leakage, we reflect the samples about the upper boundaries $\cos\gamma_i = 1$ and about the upper corner $\cos\gamma_1 = \cos\gamma_2 = 1$ of the domain, where the two-dimensional posterior peaks. We then apply a correction factor of 4 to the KDE, to compensate for the broadening of the domain.

We find $\log_{10}\mathcal{B}_{\rm P}^{\rm SA} \approx 0.28$ at $t_{\rm start} = t_{\rm peak} + 12$ ms, and $\log_{10}\mathcal{B}_{\rm P}^{\rm SA} \approx 0.63$ at $t_{\rm start} = t_{\rm peak} + 6$ ms. These numbers correspond to weak evidence against precession according to Jeffreys's scale [91]. This is consistent with previous results by Miller *et al.* [85], which show that the evidence for precession found in IMR analyses [20,45] mainly comes from the last inspiral cycles before merger (see also Ref. [92]).

D. Signal-to-noise ratios

Let us recall the noise-weighted inner product between two time series x and y,

$$\langle x|y\rangle = \sum_{i=1}^{N} (x_{\rm white})_i (y_{\rm white})_i, \tag{22}$$

where N is the number of data points, and $x_{\rm white}$ denotes the original time series x after whitening it with the noise power spectral density—and similarly for $y_{\rm white}$. Whitening is performed directly in the time domain, see Refs. [49,73]. We report the matched-filter SNR [73,75]

$$\rho_{\rm mf} = \frac{\langle x_{\rm obs}|h(\boldsymbol{\theta})\rangle}{\sqrt{\langle h(\boldsymbol{\theta})|h(\boldsymbol{\theta})\rangle}}. \tag{23}$$

For the spin-aligned model, we find $\rho_{\rm mf} = 10.7^{+0.2}_{-0.4}$ at $t_{\rm start} = t_{\rm peak} + 6$ ms and $\rho_{\rm mf} = 9.1^{+0.5}_{-0.2}$ at $t_{\rm start} = t_{\rm peak} + 12$ ms, where the lower and upper limits denote 90% confidence interval over 10^4 posterior samples. For the precessing model, we find $\rho_{\rm mf} = 10.4^{+0.3}_{-0.9}$ at $t_{\rm start} = t_{\rm peak} + 6$ ms and $\rho_{\rm mf} = 8.7^{+0.4}_{-1.0}$ at $t_{\rm start} = t_{\rm peak} + 12$ ms. The low signal strength at 12 ms past $t_{\rm peak}$ motivated us to extrapolate the model at times earlier than its nominal calibration time $t_{\rm EMOP} + 20M$.

V. DISCUSSION

In this Section, we outline some caveats that should be kept in mind when interpreting the results presented above.

Modeling assumptions. The amplitude model adopted in this work [40] does not yet include all potentially relevant physical ingredients.

First, retrograde modes can contribute non-negligibly to the ringdown signal for certain configurations of mass ratio and spin orientations [61,93,94]. Their inclusion is part of planned future developments. For the event considered here, neglecting retrograde modes is expected to be a reasonable approximation if the final spin tilt angle is small. Using posterior samples and our surrogate model (Sec. II B 2), we find that the tilt angle satisfies $\theta_f \lesssim \pi/8$ at 90% credible probability, supporting this approximation for the present analysis.

Second, the current implementation does not include the (3, 2, 0) mode. Reference [29] found this mode to be consistent with the NRSur7dq4 posteriors for GW190521. While we recover the dominant qualitative features of the ringdown without this mode, its omission limits the extent to which our framework can reproduce analyses that explicitly include it. Incorporating the (3, 2, 0) mode—and, more generally, extending the QNM content—is a natural direction for future improvements of the model.

Reliability of SBI. Simulation-based inference introduces potential sources of uncertainty beyond standard sampling error, including biases arising from the neural posterior approximation and training limitations [95,96]. In the present work, modeling uncertainties dominate over sampling-related uncertainties. Once the amplitude model is improved, a more thorough study of SBI-related uncertainties will be needed; we leave this for future work.

GPR uncertainties. The main results in Sec. IV are obtained using the mean predictions of the GPR fits. A fully consistent treatment would exploit the full GPR predictive distribution. As discussed in Appendix E, sampling from the GPR posterior currently leads to very broad uncertainties in some derived quantities. For this reason, we present the mean-based results as our primary findings and defer a more complete treatment of GPR-induced uncertainties to future work.

VI. CONCLUSIONS

In this work, we presented the first implementation of ringdown templates that incorporate spin precession within a GW parameter estimation framework, and applied them to GW190521. Using physics-informed fits for the remnant and the QNM excitation, we carried out spin-aligned and precessing ringdown analyses of GW190521 at two start times (6 and 12 ms after the strain peak).

Including precession leads to consistent shifts toward smaller luminosity distances, slightly smaller total masses, and reduced primary aligned spins, with the effect most pronounced in the 12 ms analysis. The in-plane spin components remain weakly constrained, as expected at the available SNR. The corresponding QNM amplitude posteriors show that both models yield compatible estimates for the dominant (2, 2, 0) amplitude, while precession modestly enhances subdominant excitation, lowering A_{330}/A_{220} and increasing A_{210}/A_{220}. In all cases, the (2, 2, 0) mode remains dominant, consistent with standard expectations. Bayesian model selection yields $\log_{10}\mathcal{B}_{\rm P}^{\rm SA} \simeq 0.63$ (6 ms) and $\simeq 0.28$ (12 ms), indicating that ringdown data are not

informative enough to draw firm conclusions about precession.

While full IMR analyses consider the entirety of the signal and are therefore more informative, the benefit of a ringdown-only investigation is that waveform systematics are expected to be much less severe. Provided one is sufficiently away from the peak of the strain (and how far is the subject of active investigation [88,97–99]), perturbation theory unambiguously prescribes the strain reported in Eq. (1). On the other hand, IMR analyses for high-mass, merger-ringdown–dominated signals are subject to strong systematics. The most emblematic case is GW231123 [21], whose parameters present order-unity differences when analyzed with different state-of-the-art waveform models.

Although the ringdown waveform is prescribed by perturbation theory, our analysis still inherits potential systematics from the amplitude fits, thus shifting the signpost toward the need for increased accuracy in those surrogate models. To this end, the phenomenological fits for the ringdown parameters can be retrained using the recently released SXS catalog [100], as well as public results from other groups. Future versions of the model could also include additional QNMs beyond the fundamental $\{(2,\pm2,0),(3,\pm3,0),(2,\pm1,0)\}$, or incorporate corrections for orbital eccentricity [37]. The latter effect is often neglected under the assumption that BH binaries circularize efficiently through GW emission before merger, but even such residual eccentricity can introduce systematics in the interpretation of GW signals [101]. Carullo [37] showed that noncircular orbits can substantially affect QNM excitation, with relative changes in mode amplitudes of up to 50%. A more systematic and uniform comparison with existing ringdown analyses, accounting for differences in waveform content and analysis choices, will be important to further assess the robustness and consistency of current interpretations.

Most recently, Dyer and Moore [102,103] presented Bayesian extractions of ringdown modes from NR simulations, using a likelihood that encodes information on the numerical resolution of the simulations themselves. Propagating such uncertainties upstream to analyses like ours would effectively turn potentially undetected systematic errors into statistical ones.

ACKNOWLEDGMENTS

We thank Swetha Baghwat, Francesco Nobili, Gregorio Carullo, and Riccardo Buscicchio for discussions. C. A., C. P., and D. G. are supported by ERC Starting Grant No. 945155–GWmining, Cariplo Foundation Grant No. 2021-0555, MUR PRIN Grant No. 2022-Z9X4XS, Italian-French University (UIF/UFI) Grant No. 2025-C3-386, MUR Grant "Progetto Dipartimenti di Eccellenza 2023-2027" (BiCoQ), the ICSC National Research Centre funded by NextGenerationEU, and the INFN TEONGRAV initiative. C. P. is supported by ERC Starting Grant No. 101117624–MMMonsters, MUR FIS2 Advanced Grant ET-NOW (CUP: B53C25001080001). D. G. is supported by MSCA Fellowship No. 101149270–ProtoBH and MUR Young Researchers Grant No. SOE2024-0000125. Computational work was performed at CINECA with allocations through INFN and the University of Milano-Bicocca, and at NVIDIA with allocations through the Academic Grant program.

DATA AVAILABILITY

The ringdown amplitude model is available at Ref. [56]. Posterior samples for the inference runs presented in this paper are available at Ref. [104].

APPENDIX A: PERFORMANCE OF THE MASS-SPIN FIT

We describe the construction of our surrogate model for predicting the final mass and spin of the remnant BH, as a function of the mass ratio and the component spins at ISCO. Unlike the ringdown amplitude model, these fits for the remnant properties have not been presented elsewhere. The input parameters are $\{\delta,\chi_{+z},\chi_{-z},\chi_{1x},\chi_{1y},\chi_{2x},\chi_{2y}\}$, where

$$\delta=\frac{(q-1)}{(q+1)},\qquad \chi_{\pm z}=\frac{q\chi_{1z}\pm\chi_{2z}}{1+q} \tag{A1}$$

and we order the component masses such that $q=m_1/m_2\geq 1$.

To predict the final mass M_f, we build a GPR for the quantity $1-M_f$. To predict the magnitude χ_f of the final spin, we first build a GPR for the component χ_{fz} of the final spin along the z direction in the physical frame. Then, we build a second GPR for the quantity $\delta\chi_\perp=\chi_{f\perp}-\chi_\perp$, where

$$\chi_{f\perp}=\sqrt{(\chi_f)^2-(\chi_{fz})^2} \tag{A2}$$

is the magnitude of the final spin on the orbital plane,

$$\chi_\perp=\frac{\sqrt{||\mathbf{S}_1+\mathbf{S}_2||^2-(S_{1z}+S_{2z})^2}}{(m_1+m_2)^2} \tag{A3}$$

is the total spin momentum on the orbital plane at ISCO, normalized by the component masses, and $\mathbf{S}_i=m_i^2\boldsymbol{\chi}_i$ is the physical spin of the ith progenitor. Then, the magnitude of the final spin is reconstructed as

$$\chi_f=\sqrt{(\chi_{fz})^2+(\chi_{f\perp})^2}. \tag{A4}$$

This construction ensures that the geometric condition $\chi_f\geq\chi_{fz}$ is always satisfied.

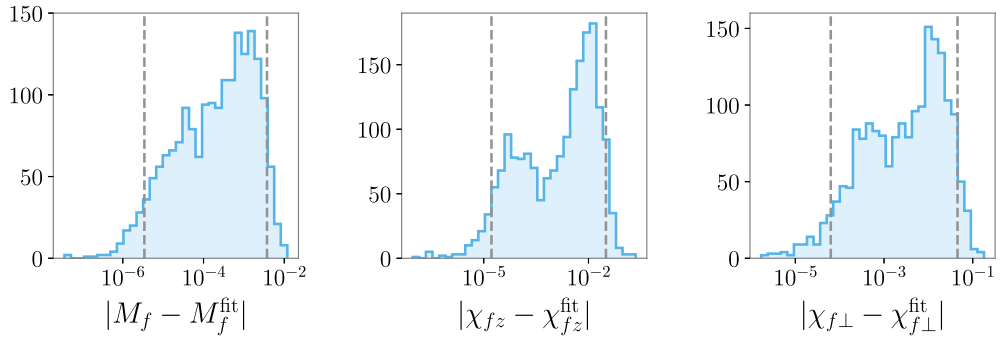

FIG. 3. Residual errors on our remnant GPR fits for $(M_f, \chi_{fz}, \chi_{f\perp})$, respectively, computed via 10-fold cross validation. Dashed vertical lines delimit the 90% density regions.

TABLE II. Root-mean-squared errors (RMSEs) for the fitted quantities $(M_f, \chi_{fz}, \delta\chi_{f\perp})$, reported as mean and standard deviation.

Quantity	RMSE (mean)	RMSE (std. dev.)
M_f	0.002	0.002
χ_{fz}	0.014	0.003
$\delta\chi_{f\perp}$	0.018	0.004

All surrogates are built by first subtracting a linear fit to the quantity of interest, then training a GPR on the corresponding residuals. Similarly to Ref. [40], we use a combination of a radial basis function kernel times a constant kernel, summed to a white kernel.

Table II shows the RMSEs for the three fitted quantities $(M_f, \chi_{fz}, \chi_{f\perp})$, computed via tenfold cross-validation [105]. Figure 3 shows the corresponding residual errors: to avoid overfitting, the residuals are evaluated at validation time, then the ten validation batches are merged together. Overall, test errors lie well within the measurement uncertainties of our parameter estimates, making the model suitable for this work.

APPENDIX B: REWEIGHTING PROCEDURE

As discussed in Sec. III C, the precessing and spin-aligned models adopt different prior choices for the BH spins, and our sampling of the mass parameters employs a different parametrization from that used in IMR analyses. To enable consistent comparisons between models, we reweight the posterior samples to match the corresponding target priors, following a similar procedure to Appendix A of [106]. For each posterior sample, the weight is given by the ratio between the target and initial priors,

$$w(x) = \frac{\pi_{\rm target}(x)}{\pi(x)}. \tag{B1}$$

1. Spins

In the precessing model, the priors are uniform in spin magnitude χ and in $\cos\gamma$ (corresponding to isotropic orientations), whereas in the spin-aligned case the prior is uniform in the aligned component χ_z, with the in-plane components set to zero. To consistently compare the two cases, we reweight the spin-aligned posterior samples so that their prior on χ_z matches the prior induced by isotropy.

We first compute the target prior as a function of $\chi_z = \chi\cos\gamma$,

$$\pi_{\rm target}(\chi_z) = \int_0^1 d\chi \int_{-1}^1 d(\cos\gamma)\pi(\chi)\pi(\cos\gamma) \times \delta(\chi_z - \chi\cos\gamma). \tag{B2}$$

Since $\chi \sim \mathcal{U}(0,1)$ and $\cos\gamma \sim \mathcal{U}(-1,1)$, we have

$$\pi_{\rm target}(\chi_z) = \frac{1}{2}\int_0^1 d\chi \int_{-1}^1 d(\cos\gamma)\delta(\chi_z - \chi\cos\gamma). \tag{B3}$$

Using the identity $\delta(f(x)) = \delta(x - x_0)/|f'(x_0)|$ with $x = \cos\gamma$, we find

$$\delta(\chi_z - \chi\cos\gamma) = \frac{\delta(\cos\gamma - \frac{\chi_z}{\chi})}{\chi}. \tag{B4}$$

The δ function restricts the integration to $|\chi_z| \le \chi$, leading to

$$\pi_{\rm target}(\chi_z) = \frac{1}{2}\int_{|\chi_z|}^1 \frac{d\chi}{\chi} = -\frac{1}{2}\ln|\chi_z|. \tag{B5}$$

The initial prior is uniform, $\pi(\chi_z) = 1/2$, giving

$$w(\chi_z) = \frac{\pi_{\rm target}(\chi_z)}{\pi(\chi_z)} = -\ln|\chi_z|. \tag{B6}$$

For both spins (χ_{1z}, χ_{2z}), the total weight is $w = w_1 w_2$, with each w_i given by Eq. (B6). The resulting weights are then normalized.

2. Masses

Our sampling employs uniform priors in total mass M and mass ratio q for numerical efficiency, while IMR analyses typically adopt priors uniform in the component masses (m_1, m_2). To reweight our samples, we derive the corresponding Jacobian between (m_1, m_2) and (M, q).

For $q \geq 1$, one has $m_1 = Mq/(1+q)$ and $m_2 = M/(1+q)$. The Jacobian determinant of the transformation is thus

$$|\det J| = M/(1+q)^2. \tag{B7}$$

We write the induced density in (M, q) as

$$\pi_{\rm target}(M, q) \propto \frac{M}{(1+q)^2}. \tag{B8}$$

In addition, we must enforce the constraint that the component masses lie within the physical prior bounds, which is tantamount to

$$m_{\rm low}(1+q) \leq M \leq m_{\rm high}(1+q)/q, \tag{B9}$$

where $m_{\rm low}$ and $m_{\rm high}$ are the lower and upper bounds on the component masses. In particular, to reweight our initial posteriors, we use $m_{\rm low} = 25$ and $m_{\rm high} = 250$, consistent with the prior ranges on M and q (see Table I). Therefore, the reweighting factors are given by Eq. (B8), subject to the constraint of Eq. (B9).

APPENDIX C: REMNANT MASS AND SPIN POSTERIORS

Figure 4 shows the posterior distributions for the remnant mass and spin obtained with our spin-aligned and precessing models, for ringdown start times of 6 ms (top panels) and 12 ms (bottom panels). As in the main text, we compare our results with those of Refs. [20,45,85], based on the NRSur7dq4 model. For the latter, remnant quantities are evaluated consistently using the associated remnant fits [84]. In the 6 ms case, we also include the posteriors from Ref. [28], obtained with a $(2,2,0)+(3,3,0)$ ringdown model. Finally, we show the remnant mass and spin reported in Ref. [29], obtained with the $(2,2,0)+(2,1,0)+(3,2,0)$ template, which provides their best agreement with NRSur7dq4. We use their results corresponding to start times of 6.35 ms and 12.7 ms, approximately matching our 6 ms and 12 ms cases, respectively.

APPENDIX D: NEURAL-NETWORK ARCHITECTURE AND TRAINING

The hyperparameters of the neural-network architecture and training are summarized in Table III. With respect to

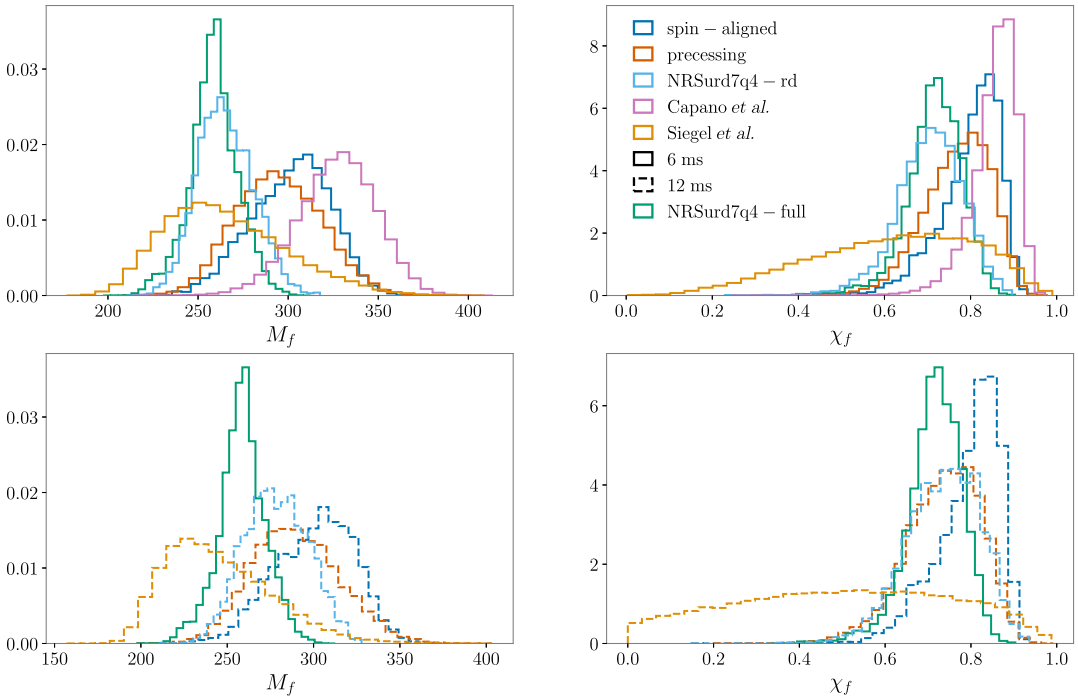

FIG. 4. Posteriors for the remnant mass M_f and spin χ_f, inferred with the spin-aligned (blue) and precessing (orange) ringdown models using the GPR mean predictions. For comparison, we also show the IMR posterior from the LVK analysis [20,45] (green), and the results of Ref. [85] (light blue), both obtained with the NRSur7dq4 model [84]. Additional comparisons include Ref. [28] (pink), based on a $(2,2,0)+(3,3,0)$ Kerr model, and Ref. [29] (dark yellow), using a $(2,2,0)+(2,1,0)+(3,2,0)$ template. The top (bottom) panels correspond to a ringdown start time of 6 ms (12 ms) after the strain peak, shown with solid (dashed) lines.

TABLE III. Summary of the neural-network architecture and training hyperparameters used in our SBI framework. Compared to Ref. [38], this setup employs a simpler embedding network and a larger training set.

Embedding network: Fully connected	
input_dim	612
output_dim	192
nun_hidden_layers	1
hidden_dim	50
Density estimator: Neural spline flow	
num_blocks	2
hidden_features	150
num_transforms	5
num_bins	10
batch_normalization	True
Training hyperparameters	
num_simulations	$300k$
batch_size	512
batch_norm	True
learning_rate	10^{-3}
truncation_quantile	10^{-4}
stopping_volume_ratio	0.8
validation_fraction	0.1
varying_noise	True

Ref. [38] (see their Table III), we adopt a simpler embedding network, consisting of a single hidden layer with 50 neurons (instead of two hidden layers with 150 neurons each), but with a slightly larger output dimension (192 versus 128). We also train the network using a larger number of simulations (3×10^5 fixed, compared to at most 1×10^5). These changes resulted in a more robust inference in our case.

APPENDIX E: ANALYSIS WITH GPR UNCERTAINTIES

In this appendix, we assess the impact of the GPR uncertainties on the inferred binary parameters and QNM amplitudes by sampling predictions from Gaussian distributions using the predicted means and standard deviations. The rest of the analysis is identical to that of the main text.

Note that our SBI implementation is set to optimize the loss function [72]

$$L = \mathbb{E}[-\log q_\phi(\theta|x_{\rm obs})], \tag{E1}$$

where the expectation value is a Monte Carlo sum over $\theta \sim \pi(\theta)$ and $x_{\rm obs} \sim p(d|\theta)$. When we sample the GPR values randomly, we are in fact taking also the expectation value over the latent functional space $\boldsymbol{\lambda}$ of the GPR model. To see this, we rewrite the loss function as

$$L = \mathbb{E}\left[-\int d\boldsymbol{\lambda}\pi_{\rm GPR}(\boldsymbol{\lambda}) \log q_\phi(\theta, |x_{\rm obs}, \boldsymbol{\lambda})\right], \tag{E2}$$

and replace the integral over the measure $d\boldsymbol{\lambda}\pi_{\rm GPR}(\boldsymbol{\lambda})$ with a Monte Carlo integral over random GPR samples. This shows that we are effectively marginalizing the likelihood over the GPR uncertainties. SBI allows us to perform this marginalization by simply extending the scope of the Monte Carlo expectation values in Eq. (E1). On the contrary, in a typical Markov chain Monte Carlo sampler, one would need to marginalize the likelihood via a Monte Carlo integral at each step, thus introducing an additional computational cost.

Figure 5 shows the analog of Fig. 1, now obtained using randomly sampled GPR realizations rather than the GPR mean predictions. As expected, including the GPR uncertainties broadens several posteriors. This effect is stronger for the luminosity distance d_L and total mass M, for which the differences between the spin-aligned and precessing models largely disappear. For χ_{1z}, the precessing posterior at 6 ms becomes consistent with the (lower-SNR) 12 ms result in Fig. 1. The inclination $\cos\iota$ and mass ratio q also show mild widening, but remain broadly consistent with the mean-prediction results. The secondary spin component χ_{2z} and the in-plane spins $(|\chi_{1\perp}|, |\chi_{2\perp}|)$ remain essentially unconstrained, so sampling the GPRs produces negligible changes relative to Fig. 1.

An analogous comparison can be made for the QNM amplitudes. Figure 6 shows the analog of Fig. 2, now using GPR samples. Including the GPR uncertainties tends to accentuate the differences between the spin-aligned and precessing models. In particular, the A_{220} posterior becomes less consistent between the two models, and the precessing posterior for A_{330}/A_{220} shifts to lower values even at 6 ms (whereas in Fig. 2 this occurred only at 12 ms).

Repeating the calculation of the Bayes factor for these GPR-sampled runs, we obtain $\log_{10}\mathcal{B}^{\rm SA}_{\rm P} \approx 0.11$ at $t_{\rm start} = t_{\rm peak} + 12$ ms, and $\log_{10}\mathcal{B}^{\rm SA}_{\rm P} \approx 0.41$ at $t_{\rm start} = t_{\rm peak} + 6$ ms. Similarly, the matched-filter SNRs for the spin-aligned model are $\rho_{\rm mf} = 10.7^{+0.2}_{-0.4}$ at $t_{\rm start} = t_{\rm peak} + 6$ ms and $\rho_{\rm mf} = 9.1^{+0.2}_{-0.4}$ at $t_{\rm start} = t_{\rm peak} + 12$ ms. For the precessing model, we find lower SNRs but with much broader uncertainties: $\rho_{\rm mf} = 8.1^{+2.4}_{-8.2}$ at $t_{\rm start} = t_{\rm peak} + 6$ ms and $\rho_{\rm mf} = 7.3^{+1.6}_{-5.9}$ at $t_{\rm start} = t_{\rm peak} + 12$ ms. These large uncertainties arise from sampling the full GPR posterior; using the mean predictions, as in the main analysis, yields more stable SNR estimates.

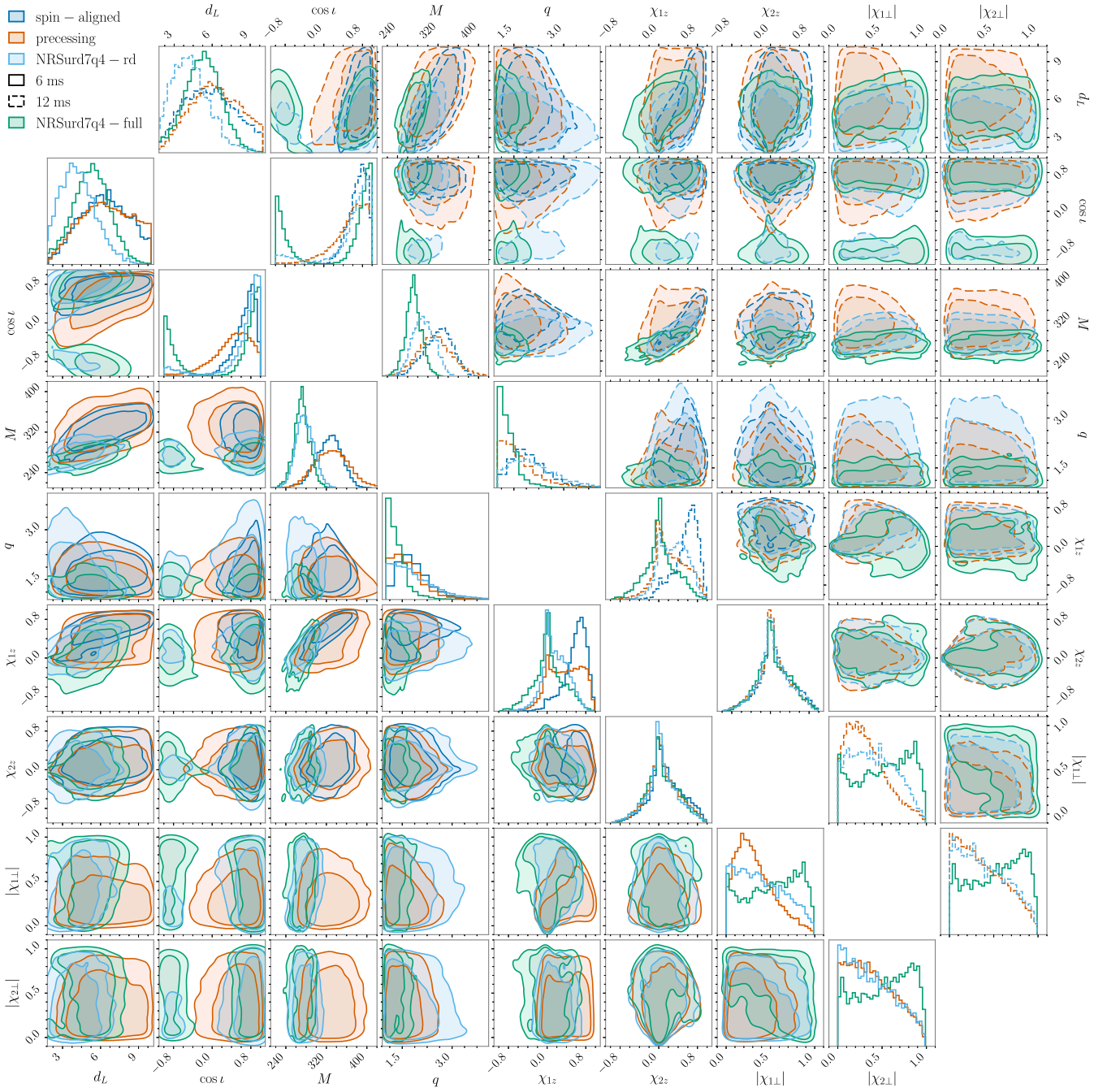

FIG. 5. Joint posterior distributions for the binary parameters inferred with the spin-aligned (blue) and precessing (orange) ringdown models, using randomly sampled GPR realizations of the remnant mass and spin and of the QNM amplitudes. The lower (upper) triangular panels correspond to a ringdown start time of 6 ms (12 ms) after the strain peak, shown with solid (dashed) lines. The IMR posterior from the LVK analysis [20,45] (green), and the results from Ref. [85] (light blue), both obtained with the NRSur7dq4 model [84], are overlaid for comparison. All contours represent the 68% and 90% credible regions.

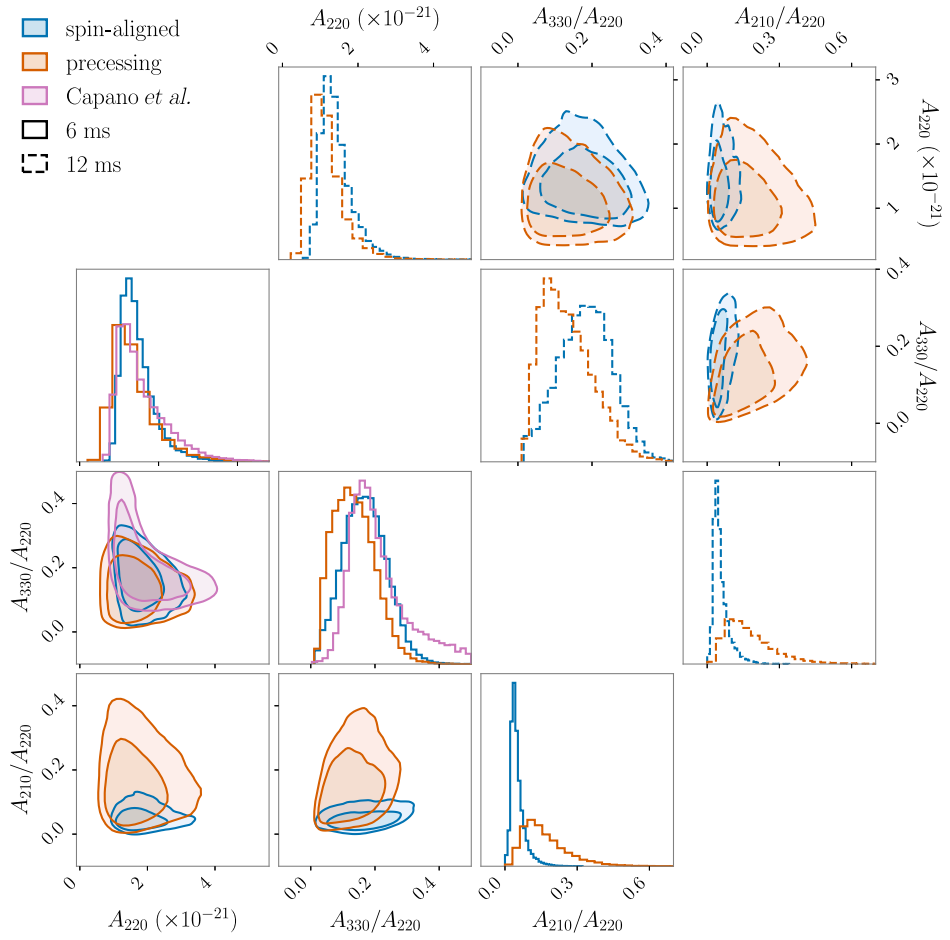

FIG. 6. Joint posterior distributions for the QNM amplitudes inferred from the spin-aligned (blue) and precessing (orange) ringdown models, using randomly sampled GPR realizations of the remnant mass and spin and of the QNM amplitudes. We show the absolute amplitude of the dominant (2, 2, 0) mode, A_{220}, together with the relative amplitudes A_{330}/A_{220} and A_{210}/A_{220}. The lower (upper) triangular panels correspond to a ringdown start time of 6 ms (12 ms) after the strain peak, shown with solid (dashed) lines. For the 6 ms case, we overlay the results of Ref. [28] (pink), obtained using a $(2,2,0)+(3,3,0)$ Kerr model. Contours indicate the 68% and 90% credible regions.

[1] R. P. Kerr, Phys. Rev. Lett. **11**, 237 (1963).
[2] S. A. Teukolsky, Classical Quantum Gravity **32**, 124006 (2015).
[3] S. A. Teukolsky, Phys. Rev. Lett. **29**, 1114 (1972).
[4] S. A. Teukolsky, Astrophys. J. **185**, 635 (1973).
[5] E. Berti, V. Cardoso, and A. O. Starinets, Classical Quantum Gravity **26**, 163001 (2009).
[6] K. D. Kokkotas and B. G. Schmidt, Living Rev. Relativity **2**, 2 (1999).
[7] C. V. Vishveshwara, Nature (London) **227**, 936 (1970).
[8] W. H. Press, Astrophys. J. Lett. **170**, L105 (1971).
[9] W. H. Press and S. A. Teukolsky, Astrophys. J. **185**, 649 (1973).
[10] E. Berti, V. Cardoso, G. Carullo *et al.*, arXiv:2505.23895.
[11] S. Detweiler, Astrophys. J. **239**, 292 (1980).
[12] O. Dreyer, B. Kelly, B. Krishnan, L. S. Finn, D. Garrison, and R. Lopez-Aleman, Classical Quantum Gravity **21**, 787 (2004).
[13] E. Berti, K. Yagi, H. Yang, and N. Yunes, Gen. Relativ. Gravit. **50**, 49 (2018).
[14] B. P. Abbott *et al.*, Phys. Rev. Lett. **116**, 061102 (2016).
[15] B. P. Abbott *et al.*, Phys. Rev. X **9**, 031040 (2019).
[16] R. Abbott *et al.*, Phys. Rev. X **11**, 021053 (2021).
[17] R. Abbott *et al.*, Phys. Rev. D **109**, 022001 (2024).
[18] R. Abbott *et al.*, Phys. Rev. X **13**, 041039 (2023).
[19] A. G. Abac *et al.*, arXiv:2508.18082.
[20] R. Abbott *et al.*, Phys. Rev. Lett. **125**, 101102 (2020).
[21] A. G. Abac *et al.*, Astrophys. J. Lett. **993**, L25 (2025).
[22] C. J. Moore, E. Finch, R. Buscicchio, and D. Gerosa, iScience **24**, 102577 (2021).

[23] A. Gupta *et al.*, SciPost Phys. Commun. Rep. **5**, 11 (2025).
[24] Q. Hu and J. Veitch, Phys. Rev. D **106**, 044042 (2022).
[25] A. Borchers and F. Ohme, Classical Quantum Gravity **40**, 095008 (2023).
[26] D. Fernando, R. O'Shaughnessy, and D. Williams, Phys. Rev. D **112**, 084004 (2025).
[27] H. Siegel, N. M. Khusid, M. Isi, and W. M. Farr, arXiv:2511.02691.
[28] C. D. Capano, M. Cabero, J. Westerweck, J. Abedi, S. Kastha, A. H. Nitz, Y.-F. Wang, A. B. Nielsen, and B. Krishnan, Phys. Rev. Lett. **131**, 221402 (2023).
[29] H. Siegel, M. Isi, and W. M. Farr, Phys. Rev. D **108**, 064008 (2023).
[30] I. Kamaretsos, M. Hannam, and B. S. Sathyaprakash, Phys. Rev. Lett. **109**, 141102 (2012).
[31] L. London, D. Shoemaker, and J. Healy, Phys. Rev. D **90**, 124032 (2014).
[32] V. Baibhav, E. Berti, V. Cardoso, and G. Khanna, Phys. Rev. D **97**, 044048 (2018).
[33] L. T. London, Phys. Rev. D **102**, 084052 (2020).
[34] S. Borhanian, K. G. Arun, H. P. Pfeiffer, and B. S. Sathyaprakash, Classical Quantum Gravity **37**, 065006 (2020).
[35] X. J. Forteza, S. Bhagwat, S. Kumar, and P. Pani, Phys. Rev. Lett. **130**, 021001 (2023).
[36] M. H.-Y. Cheung, E. Berti, V. Baibhav, and R. Cotesta, Phys. Rev. D **109**, 044069 (2024).
[37] G. Carullo, J. Cosmology Astropart. Phys. 10 (**2024**) 061.
[38] C. Pacilio, S. Bhagwat, F. Nobili, and D. Gerosa, Phys. Rev. D **110**, 103037 (2024).
[39] L. Magaña Zertuche, L. C. Stein, K. Mitman *et al.*, Phys. Rev. D **112**, 024077 (2025).
[40] F. Nobili, S. Bhagwat, C. Pacilio, and D. Gerosa, Phys. Rev. D **112**, 044058 (2025).
[41] H. Estellés, S. Husa, M. Colleoni *et al.*, Astrophys. J. **924**, 79 (2022).
[42] I. Romero-Shaw, P. D. Lasky, E. Thrane, and J. Calderón Bustillo, Astrophys. J. Lett. **903**, L5 (2020).
[43] R. Gamba, M. Breschi, G. Carullo *et al.*, Nat. Astron. **7**, 11 (2023).
[44] V. Gayathri, J. Healy, J. Lange, B. O'Brien, M. Szczepańczyk, I. Bartos, M. Campanelli, S. Klimenko, C. O. Lousto, and R. O'Shaughnessy, Nat. Astron. **6**, 344 (2022).
[45] R. Abbott *et al.*, Astrophys. J. Lett. **900**, L13 (2020).
[46] K. Cranmer, J. Brehmer, and G. Louppe, Proc. Natl. Acad. Sci. U.S.A. **117**, 30055 (2020).
[47] M. Deistler, J. Boelts, P. Steinbach, G. Moss, T. Moreau, M. Gloeckler, P. L. C. Rodrigues, J. Linhart, J. K. Lappalainen, B. K. Miller, P. J. Gonçalves, J.-M. Lueckmann, C. Schröder, and J. H. Macke, arXiv:2508.12939.
[48] M. Crisostomi, K. Dey, E. Barausse, and R. Trotta, Phys. Rev. D **108**, 044029 (2023).
[49] C. Pacilio, S. Bhagwat, and R. Cotesta, Phys. Rev. D **110**, 083010 (2024).
[50] C. T. Cunningham, R. H. Price, and V. Moncrief, Astrophys. J. **224**, 643 (1978).
[51] M. De Amicis, H. R. Rüter, G. Carullo, S. Albanesi, C. Melize Ferrus, K. Mitman, L. C. Stein, V. Cardoso, S. Bernuzzi, M. Boyle, N. Deppe, L. E. Kidder, J. Moxon, A. Nagar, K. C. Nelli, H. P. Pfeiffer, M. A. Scheel, W. Throwe, N. L. Vu, and A. Zenginoğlu, Phys. Rev. Lett. **135**, 171401 (2025).
[52] S. Ma, M. A. Scheel, J. Moxon, K. C. Nelli, N. Deppe, L. E. Kidder, W. Throwe, and N. L. Vu, Phys. Rev. D **112**, 024003 (2025).
[53] E. Berti and A. Klein, Phys. Rev. D **90**, 064012 (2014).
[54] B. S. Sathyaprakash and B. F. Schutz, Living Rev. Relativity **12**, 2 (2009).
[55] M. Isi, Classical Quantum Gravity **40**, 203001 (2023).
[56] C. Pacilio and F. Nobili, github.com/cpacilio/postmerger (2025).
[57] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning* (MIT, Cambridge, MA, 2006).
[58] M. Boyle, D. Hemberger, D. A. B. Iozzo *et al.*, Classical Quantum Gravity **36**, 195006 (2019).
[59] E. Finch and C. J. Moore, Phys. Rev. D **103**, 084048 (2021).
[60] E. Berti, V. Cardoso, J. A. Gonzalez, U. Sperhake, M. Hannam, S. Husa, and B. Brügmann, Phys. Rev. D **76**, 064034 (2007).
[61] H. Zhu, H. Siegel, K. Mitman *et al.*, Phys. Rev. D **111**, 064052 (2025).
[62] E. P. Wigner and U. Fano, Am. J. Phys. **28**, 408 (1960).
[63] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum* (World Scientific, Singapore, 1988).
[64] P. Schmidt, M. Hannam, S. Husa, and P. Ajith, Phys. Rev. D **84**, 024046 (2011).
[65] M. Deistler, P. J. Goncalves, and J. H. Macke, arXiv:2210.04815.
[66] G. Papamakarios and I. Murray, arXiv:1605.06376.
[67] D. S. Greenberg, M. Nonnenmacher, and J. H. Macke, arXiv:1905.07488.
[68] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, arXiv:1912.01703.
[69] J. Boelts, M. Deistler, M. Gloeckler *et al.*, J. Open Source Software **10**, 7754 (2025).
[70] G. Papamakarios, E. Nalisnick, D. Jimenez Rezende, S. Mohamed, and B. Lakshminarayanan, J. Mach. Learn. Res. **22**, 2617 (2021), https://arxiv.org/abs/1912.02762.
[71] C. Durkan, A. Bekasov, I. Murray, and G. Papamakarios, arXiv:1906.04032.
[72] S. R. Green and J. Gair, Mach. Learn. Sci. Technol. **2**, 03LT01 (2021).
[73] M. Isi and W. M. Farr, arXiv:2107.05609.
[74] E. Finch and C. J. Moore, Phys. Rev. D **104**, 123034 (2021).
[75] V. Gennari, G. Carullo, and W. Del Pozzo, Eur. Phys. J. C **84**, 233 (2024).
[76] A. Nitz *et al.*, github.com/gwastro/pycbc (2025).
[77] G. Carullo, L. van der Schaaf, L. London, P. T. H. Pang, K. W. Tsang, O. A. Hannuksela, J. Meidam, M. Agathos, A. Samajdar, A. Ghosh, T. G. F. Li, W. Del Pozzo, and C. Van Den Broeck, Phys. Rev. D **98**, 104020 (2018).
[78] R. Abbott *et al.*, Phys. Rev. D **103**, 122002 (2021).

[79] R. Abbott *et al.*, Phys. Rev. D **112**, 084080 (2025).
[80] A. Toubiana and J. R. Gair, arXiv:2401.06845.
[81] H. Siegel, M. Isi, and W. M. Farr, Phys. Rev. D **111**, 044070 (2025).
[82] A. G. Abac *et al.*, Phys. Rev. Lett. **135**, 111403 (2025).
[83] C. Capano, github.com/cdcapano/pykerr (2023).
[84] 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).
[85] S. J. Miller, M. Isi, K. Chatziioannou, V. Varma, and I. Mandel, Phys. Rev. D **109**, 024024 (2024).
[86] S. J. Miller, M. Isi, V. Varma, and K. Chatziioannou, 10.5281/zenodo.8349282 (2023).
[87] M. Mould and D. Gerosa, Phys. Rev. D **105**, 024076 (2022).
[88] V. Baibhav, M. H.-Y. Cheung, E. Berti, V. Cardoso, G. Carullo, R. Cotesta, W. Del Pozzo, and F. Duque, Phys. Rev. D **108**, 104020 (2023).
[89] A. G. Abac *et al.*, Phys. Rev. Lett. **136**, 041403 (2026).
[90] S. Kastha, J. Westerweck, Y.-F. Wang, and C. Capano, github.com/gwastro/BH-Spectroscopy-GW190521 (2023).
[91] H. Jeffreys, *The Theory of Probability* (Oxford University Press, Oxford, 1939).
[92] S. Biscoveanu, M. Isi, V. Varma, and S. Vitale, Phys. Rev. D **104**, 103018 (2021).
[93] E. Hamilton, L. London, and M. Hannam, Phys. Rev. D **107**, 104035 (2023).
[94] A. Dhani, Phys. Rev. D **103**, 104048 (2021).
[95] J. Hermans, A. Delaunoy, F. Rozet, A. Wehenkel, V. Begy, and G. Louppe, arXiv:2110.06581.
[96] J.-M. Lueckmann, J. Boelts, D. S. Greenberg, P. J. Gonçalves, and J. H. Macke, arXiv:2101.04653.
[97] S. Bhagwat, M. Okounkova, S. W. Ballmer, D. A. Brown, M. Giesler, M. A. Scheel, and S. A. Teukolsky, Phys. Rev. D **97**, 104065 (2018).
[98] S. H. Völkel and A. Dhani, Phys. Rev. D **112**, 084076 (2025).
[99] F. Crescimbeni, G. Carullo, E. Berti, G. Caneva Santoro, M. Ho-Yeuk Cheung, and P. Pani, arXiv:2511.02915.
[100] M. A. Scheel, M. Boyle, K. Mitman *et al.*, Classical Quantum Gravity **42**, 195017 (2025).
[101] G. Fumagalli, I. Romero-Shaw, D. Gerosa, V. De Renzis, K. Kritos, and A. Olejak, Phys. Rev. D **110**, 063012 (2024).
[102] R. Dyer and C. J. Moore, arXiv:2510.13954.
[103] R. Dyer and C. J. Moore, arXiv:2510.11783.
[104] C. Anselmo, github.com/chiaranselmo17/ringdown_GW190521 (2025), 10.5281/zenodo.17857210.
[105] L. Liu and M. T. Özsu, *Encyclopedia of Database Systems* (Springer, New York, 2009).
[106] J. Lange, R. O'Shaughnessy, and M. Rizzo, arXiv:1805.10457.