

Accuracy of ringdown models calibrated to numerical relativity simulations

Francesco Crescimbeni^{1,2,*} Gregorio Carullo^{3,†} Emanuele Berti^{4,‡} Giada Caneva Santoro^{5,6,§}
Mark Ho-Yeuk Cheung^{4,7,||} and Paolo Pani^{1,2,¶}

¹*Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 5, 00185, Roma, Italy*

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

³*School of Physics and Astronomy and Institute for Gravitational Wave Astronomy,
University of Birmingham, Birmingham, B15 2TT, United Kingdom*

⁴*William H. Miller III Department of Physics and Astronomy,
Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland, 21218, USA*

⁵*Institut de Física d'Altes Energies (IFAE), The Barcelona Institute*

⁶*Center of Gravity, Niels Bohr Institute, Blegdamsvej 17, 2100 Copenhagen, Denmark*

⁷*School of Natural Sciences, Institute for Advanced Study,
Einstein Drive, Princeton, New Jersey, 08540, USA*



(Received 13 November 2025; accepted 20 May 2026; published 10 June 2026)

The “ringdown” stage of gravitational-wave signals from binary black hole mergers, mainly consisting of a superposition of quasinormal modes emitted by the merger remnant, is a key tool to test fundamental physics and to probe black hole dynamics. However, ringdown models are known to be accurate only in the late-time, stationary regime. A key open problem in the field is to understand if these models are robust when extrapolated to earlier times, and if they can faithfully recover a larger portion of the signal. We address this question through a systematic time-domain calculation of the mismatch between non-precessing, quasicircular ringdown models parametrized by the progenitor binary’s degrees of freedom and full numerical relativity inspiral-merger-ringdown waveforms from the simulating extreme spacetimes (SXS) simulation catalog. For the best-performing models, the mismatch is typically in the range $[10^{-6}, 10^{-4}]$ for the $(\ell, |m|) = (2, 2)$ harmonic, and $[10^{-4}, 10^{-2}]$ for higher-order modes. Our findings inform ongoing observational searches for quasinormal modes, and underscore the need for improved modeling of higher-order modes to meet the sensitivity requirements of future gravitational-wave detectors.

DOI: [10.1103/wdqf-11xc](https://doi.org/10.1103/wdqf-11xc)

I. INTRODUCTION

The first observation of a gravitational-wave (GW) signal emitted by the coalescence of two black holes (BHs) [1] has revolutionized our understanding of the Universe, with important implications in astrophysics, fundamental physics, and cosmology [2–5]. While the early inspiral signal provides information on the dynamics of the two objects, the merger and ringdown stages offer a unique observational

window into the strong-field dynamical regime of general relativity (GR) and the nature of the remnant object [6–10].

Within black hole perturbation theory [11–14], it is possible to describe part of the signal as a superposition of damped sinusoids, or quasinormal modes (QNMs) [15–17]. As a consequence of BH uniqueness theorems, the QNM frequencies and damping times of uncharged BHs belong to an infinite set of discrete values that solely depend on the final mass and final spin of the remnant [18–21]. The research program aimed at measuring these frequencies and damping times is known as “black hole spectroscopy” [22–26] (see [27] for a recent review). A recent breakthrough was the detection by the LIGO-Virgo-Kagra (LVK) collaboration of GW250114, a loud event for which at least two QNMs (the fundamental mode and the first overtone) were confidently identified [28,29]. The degree of excitation of different QNMs in a merger remnant depends on the parameters of the binary progenitors. The complex amplitudes of the modes grow dynamically in time during the plunge-merger stage [30]. They later saturate to constant values after a timescale of $\sim 10\text{--}20GM/c^3$ past the peak of the radiation, where M denotes the merger remnant mass, G

*Contact author: francesco.crescimbeni@uniroma1.it

†Contact author: g.carullo@bham.ac.uk

‡Contact author: berti@jhu.edu

§Contact author: giada.santoro@nbi.ku.dk

||Contact author: mcheung@ias.edu

¶Contact author: paolo.pani@uniroma1.it

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

the gravitational constant, and c the speed of light (here and in the rest of the paper we use geometrical units, $G = c = 1$). Through numerical relativity (NR) and analytical work, one can model these constant excitation amplitudes, using them to infer the binary properties for massive signals for which the inspiral is unobservable [26,30–37]. Further, accurately modeling the QNM amplitudes and phases reduces systematic biases in parameter estimation and boosts searches of new physics [38–45]. This is particularly important for future, high signal-to-noise ratio (SNR) observations with the space interferometer LISA [46] and with third-generation ground-based detectors, like the Einstein Telescope (ET) [47] and Cosmic Explorer (CE) [48].

Large efforts have been devoted to modeling the complex QNM excitation amplitudes via calibration to numerical simulations [49–55]. In this work, we focus on two such *representative* models: the post-Newtonian inspired closed-form expressions for the QNMs excitation amplitudes of Ref. [49] (henceforth London) and the model of Ref. [51] (henceforth Cheung), which extended those predictions through an efficient and robust fitting algorithm capable of handling a larger number of modes, including quadratic QNMs [56–68]. Both models consider quasicircular, spin-aligned binaries, although these assumptions have recently been relaxed by including either precession [53] or eccentricity [69].

Earlier times, during merger and early postmerger phases, are contaminated by the prompt response and dynamical growth [17,30,36,37,70–73] or nonmodal effects beyond the linear approximation [68,71,74–77]. The precise time at which the transition between the dynamical phase, where the amplitudes evolve over time, and the stationary, constant-amplitude QNM phase takes place depends on the binary parameters, and its analytical determination is still an open problem [30,37]. This implies an uncertainty on the starting time (relative to the waveform peak amplitude) at which it makes sense to start the QNM analysis. Starting at early times can induce systematics in the measurement of QNMs, since a pure damped sinusoid model is invalid; conversely, starting too late leads to an exponential reduction in the ringdown SNR. Ringdown models that go beyond the assumption of purely QNM-based signals have been developed within the effective-one-body framework to address this issue [78–80]. In this approach, the entire postmerger portion is modeled via phenomenological *Ansätze* for the time-dependent QNM amplitudes that retain information about the progenitor’s parameters, and effectively incorporate transient and nonlinear effects in the early postmerger regime [81–88]. These models have the advantage of capturing the entire postmerger signal starting from the peak, but their parametrization does not reflect our first-principles understanding of the underlying physics, and hence they are less useful for the interpretation of possible signatures of new physics. Moreover, the numerical ansätze with multiple free parameters makes it significantly more complicated to easily identify physical quantities of

interest. Out of these models, in this paper we will focus on the nonprecessing, quasicircular TEOBPM template [81,82,85,86].

In this work, we aim to determine the starting time for which these three numerically informed ringdown models (London, Cheung, and TEOBPM) provide an accurate description of the signal by directly comparing them to NR simulations. To this end, we perform time-domain mismatch computations [27,89,90]. We work in the time domain to avoid spectral leakage: see, e.g., Ch. 6 of Ref. [27]. These mismatch comparisons do not suffer from overfitting issues, since there is no “fitting” involved in our work: the physical modal content has already been determined in [51,86,88,91] and is not varied here.

When comparing London and Cheung to numerical relativity simulations, we find that both models fit the waveform well at late times, with mismatches $\sim 10^{-4}$ at times ~ 25 M after the merger peak. For most (ℓ, m) waveform harmonics, the Cheung model performs relatively better at earlier times because it includes overtones, it extracts the QNM amplitudes in a time window where they are most stable, and it is built from the same simulation catalog used in this work to validate the models. When we also include TEOBPM in the comparison, we find that the mismatch between the most accurate ringdown models and numerical simulations typically reaches values between $[10^{-6}, 10^{-4}]$ for the $(\ell, |m|) = (2, 2)$ harmonic, and $[10^{-4}, 10^{-2}]$ for higher-order modes. While this level of mismatch is typically acceptable for current detections, it may be insufficient to prevent systematic biases for the high-SNR events anticipated with future detectors [40,92–94]. Our calculations provide a practical criterion to identify the optimal starting time for a given NR-informed ringdown model, and they can be readily applied to current QNM searches through a public repository implementing ready-to-use starting-time interpolants for a given required accuracy [95].

The remainder of the paper is organized as follows. In Sec. II, we detail the ringdown modal content, the time-domain mismatch calculation, and its application to the ringdown regime. The main results are presented in Sec. III, and our conclusions are summarized in Sec. IV.

II. COMPARING RINGDOWN MODELS AND NUMERICAL RELATIVITY DATA

A. Ringdown modeling

We start by describing how we model ringdown signals. We will limit attention to quasicircular, nonprecessing binaries. Following the conventions of Ref. [27], we can express the ringdown waveform in GR as

$$h(t) = \sum_{\ell, m, n, \pm} \mathcal{A}_{\ell mn}^{\pm} e^{i\phi_{\ell mn}^{\pm}} S_{\ell mn}^{\pm}(\iota, \varphi) e^{i\omega_{\ell mn}^{\pm}(t - t_{\ell mn}^{\pm})} \quad (1)$$

where (ℓ, m, n) are the angular, azimuthal, and overtone indices, respectively; $+$ and $-$ denote the prograde and retrograde modes (see Sec. 5.1.6 of [27]). When discussing contributions of a given QNM index, we use the compact notation $(\ell, m, n)^+$ and $(\ell, m, n)^-$. ${}_{-2}S_{\ell mn}(\iota, \varphi, \omega_{\ell mn}^\pm)$ denotes the spin-2 spheroidal harmonics, where ι and φ denote the polar (relative to the direction of the orbital angular momentum) and azimuthal angles in the source frame; $\omega_{\ell mn}^\pm = 2\pi f_{\ell mn}^\pm - i/\tau_{\ell mn}^\pm$ are the complex QNM frequencies of the remnant BH, with mass M and spin χ ; $t_{\ell mn}^\pm$ is a conventional reference starting time; $\mathcal{A}_{\ell mn}^\pm$ and $\phi_{\ell mn}^\pm$ are the amplitudes and phases corresponding to the $(\ell, m, n)^\pm$ modes, which depend on the progenitor parameters $\theta = (\eta, \chi_1, \chi_2)$, where the symmetric mass ratio is $\eta = \frac{m_1 m_2}{(m_1 + m_2)^2}$ in terms of the binary component masses m_1, m_2 , and χ_1, χ_2 are the dimensionless progenitor spin components parallel to the orbital angular momentum. The GW strain in NR simulations is usually expanded in spin-2 spherical harmonics ${}_{-2}Y_{\ell m}(\iota, \varphi)$

$$h(t) = h_+(t) - ih_\times(t) = \sum_{\ell, m} h_{\ell m}(t) {}_{-2}Y_{\ell m}(\iota, \varphi). \quad (2)$$

In general, Eqs. (1) and (2) are equivalent upon expanding the spheroidal harmonics in terms of spherical ones

$${}_{-2}S_{\ell mn}(\iota, \varphi, \omega_{\ell mn}^\pm) = \sum_{\ell'} \mu_{m\ell'\ell n}^\pm {}_{-2}Y_{\ell' m}(\theta, \varphi), \quad (3)$$

where $\mu_{m\ell'\ell n}^\pm$ are the complex spherical-spheroidal mixing factors [96,97], see Ref. [98]. Hence, we can write the $h_{\ell m}(t)$ components of the ringdown strain as

$$h_{\ell m}(t) = \sum_{\ell'=2}^{\infty} \sum_{n=0}^{\infty} \sum_{\pm} \mu_{m\ell'\ell n}^\pm \mathcal{A}_{\ell' mn}^\pm e^{i\phi_{\ell' mn}^\pm} e^{i\omega_{\ell' mn}^\pm (t - t_{\ell' mn}^\pm)}. \quad (4)$$

Then, each spherical component (ℓ, m) corresponds to a superposition of spheroidal QNMs with the same value of m but different values of ℓ . Since we focus on non-precessing systems, negative- m modes are related to positive- m modes by symmetry (see, e.g., [91,99,100]):

$$h_{\ell, -m} = (-1)^\ell h_{\ell, m}^*. \quad (5)$$

B. Time-domain formalism and construction of the autocorrelation function

Here, we describe the mismatch computation between two signals in the time domain following Refs. [89,90]. The starting point is the detector's noise power spectral density in the frequency domain, which we convert into an acyclic estimate of the autocovariance function (ACF). We model the instrumental noise as a discrete random process, denoted by $n_i = n(t_i)$ for $i = 1, \dots, N$, where the time

samples are $t_i \in \{0, \Delta t, 2\Delta t, \dots, (N-1)\Delta t\}$, and the total signal duration is $T = N\Delta t$. We assume the noise to be Gaussian, and as such fully characterized by its mean and covariance matrix. Without loss of generality, we can set the mean of the noise to zero

$$\mu = E[n_i] = 0, \quad (6)$$

where $E[\cdot]$ denotes the expectation value, and we denote the covariance matrix by $C_{ij} = E[n_i n_j]$.

We also assume the noise to be *stationary*, so the covariance matrix does not depend on the individual times t_j, t_l , but only on their difference $|t_j - t_l|$

$$C_{jl} = \rho(|j - l|), \quad (7)$$

where $\rho(k)$ is known as the ACF. If, in addition to stationarity, we impose periodic boundary conditions, then $\rho(k) = \rho(N - k)$ is said to be cyclic. This function can be estimated by performing the inverse Fourier transform of the power spectral density (PSD) $S(f)$

$$\rho(k) = \frac{1}{2T} \sum_{j=0}^{N-1} S(|f_j|) e^{2\pi i j k / N}, \quad (8)$$

where $f_i \in \{0, \Delta f, \dots, f_{\max}\}$, and

$$\Delta f = \frac{f_s}{N} = \frac{1}{T}, \quad (9)$$

with sampling frequency $f_s = 2f_{\max}$. In this work we use the LIGO design sensitivity curve provided in Ref. [101].

In practice, since the PSD is a discrete frequency series, computing the ACF requires Fourier transform algorithms (e.g., [102]). Artifacts at high frequencies may be present due to aliasing when the sampling frequency f_s is smaller than the Nyquist frequency, i.e., $f_s < 2f_{\max}$; also, care must be taken at lower frequencies to avoid contributions from signal portions below the nominal detector sensitivity. To prevent these effects, we apply a window function to the edges of the PSD prior when computing its inverse Fourier transform (see Appendix A for details).

Given Eq. (7), we can define the scalar product between two signals $x(t)$ and $y(t)$ in the time domain as

$$\langle x|y \rangle = \sum_{i,j=0}^N x_i C_{ij}^{-1} y_j. \quad (10)$$

In this section, the signals are assumed to be aligned in time and phase (something that we will always enforce in our analysis). The formalism described above can be applied to GW data analysis of full inspiral-merger-ringdown (IMR) signals. A “windowing” function can be applied to set the IMR signal to zero at the boundary,

when the emission has fallen below a detectable level. This smooth tapering additionally suppresses spectral leakage when performing the Fourier transform [103]. However, the data for ringdown-only analyses are nonzero on the left boundary (being preceded by the merger). In this case, a standard windowing procedure can still be used, at the price of losing part of the signal or introducing contamination from the merger [89,104].

To avoid these issues, we work in the time domain without assuming cyclic boundary conditions. We restrict our analysis to the time window t_{trunc} , defined as

$$t_{\text{trunc}} = [t_{\text{start}}, \dots, t_{\text{end}}] \quad (11)$$

where t_{start} and t_{end} can be associated to a discretized pair t_s , t_e such that

$$\begin{cases} t_s \equiv \min(t_i | t_{\text{start}} \leq t_i), \\ t_e \equiv \max(t_i | t_{\text{end}} \geq t_i). \end{cases} \quad (12)$$

We associate the index $i = 0$ to $t = 0$, and $i = N_{\text{trunc}}$ to $t_e - t_s$, where $N_{\text{trunc}} = \dim(t_{\text{trunc}})$. We construct the autocovariance matrix using an acyclic estimate of the ACF, taking the ACF constructed from Eq. (8) (of length N_{FFT}) and truncate it at the desired ringdown length N_{trunc} , with $N_{\text{FFT}} \gg N_{\text{trunc}}$. The resulting $\rho^A(k)$ is illustrated in Fig. 1, and the corresponding autocovariance matrix is $C_{ij}^A = \rho^A(|i - j|)$. Given two ringdownlike signals $x(t)$, $y(t)$, we define a scalar product in the range $t \in [t_s, \dots, t_e]$ as follows:

$$\langle x | y \rangle_{\text{tr}} = \sum_{i,j=0}^{N_{\text{trunc}}} x_i (C_{ij}^A)^{-1} y_j. \quad (13)$$

We are interested in comparing a given ringdown model $h_{\ell m}^{\text{rd}}(t)$, where “rd” stands for ringdown, against an NR waveform $h_{\ell m}^{\text{NR}}(t)$. We define the mismatch between the two individual waveform multipoles as

$$\mathcal{M}^{(\ell,m)} = 1 - \max_{\phi_c, t_c} \frac{\langle h_{\ell m}^{\text{rd}} | h_{\ell m}^{\text{NR}} \rangle_{\text{tr}}}{\sqrt{\langle h_{\ell m}^{\text{rd}} | h_{\ell m}^{\text{rd}} \rangle_{\text{tr}} \langle h_{\ell m}^{\text{NR}} | h_{\ell m}^{\text{NR}} \rangle_{\text{tr}}}}, \quad (14)$$

where we recall that the scalar product of Eq. (13) is defined in the time window

$$t \in [t_{\text{start}}^{(\ell,m)}, t_{\text{end}}^{(\ell,m)}], \quad (15)$$

and where $\max_{\phi_c, t_c}$ denotes optimization over the coalescence time and phase.

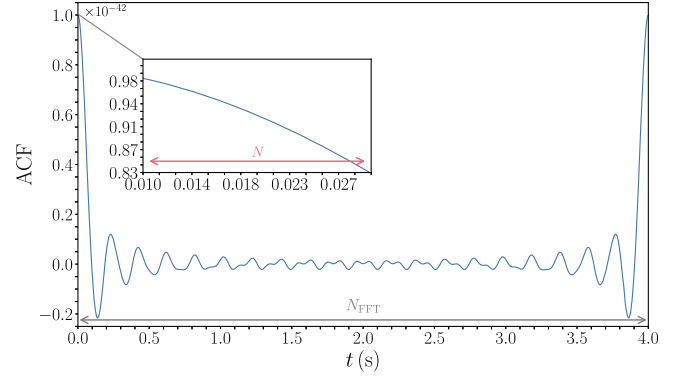


FIG. 1. ACF as a function of time in the range $t = [0, 4]$ s, where T is the observation time. The inset shows the truncated ACF in the time range $[t_s, t_e]$, of length N .

III. MISMATCH FOR QUASICIRCULAR, NONPRECESSING BINARIES

In this section, we present the main results of our work. We use the mismatch, computed as in Eq. (14), to compare ringdown models against the SXS catalog of NR simulations [105,106]. Each of the simulations we considered is labeled SXS:BBH:NNNN, with NNNN denoting a unique simulation ID corresponding to a set of progenitor parameters. We restrict our analysis to nonprecessing quasicircular binaries with spin components in the x and y directions such that $|\chi_{i,x}|, |\chi_{i,y}| \lesssim 10^{-3}$, and with eccentricity $\lesssim 10^{-3}$. Throughout this work, we denote the resulting ensemble of simulations by $\mathcal{I}_{\text{tot}}$. In the following, we will focus on the progenitor parameters $\theta = (\eta, \chi_{1,z}, \chi_{2,z})$, where $\chi_{1,z}$ and $\chi_{2,z}$ denote the dimensionless spin components projected along the z -axis, all considered at their relaxation values.

To access the $(\ell, |m|)$ harmonics of each NR waveform $h_{\ell m}^{\text{NR}}(t, \theta)$, we make use of the SXS package [107]. The latter is interfaced through bayRing, a publicly available PYTHON package [108] that relies on pyRing [109] to access ringdown waveform models, and on the QNM [110] package to compute QNM frequencies.

Within the bayRing software package, both the NR and ringdown model are time-aligned such that the time origin is at $t = t_{\text{peak}}^{(2,2)}$, the peak time of the $(2, 2)$ mode. The package uses the NR strain as the data and the ringdown model as the template, and it estimates the free parameters of the template via nested sampling. In our case, since the NR-informed ringdown models $h_{\ell m}^{\text{rd}}(t, \theta)$ have amplitudes and relative phases fixed by the set of parameters θ associated with the SXS simulation with the given ID, the only model parameter left to infer is a free phase.

We quote results for representative GW150914-like values of the remnant mass, $M = 62M_{\odot}$, and of the luminosity distance, $d_L = 410$ Mpc [1]. The London and Cheung models are late-ringdown models composed of pure QNM superpositions with amplitudes and phases that

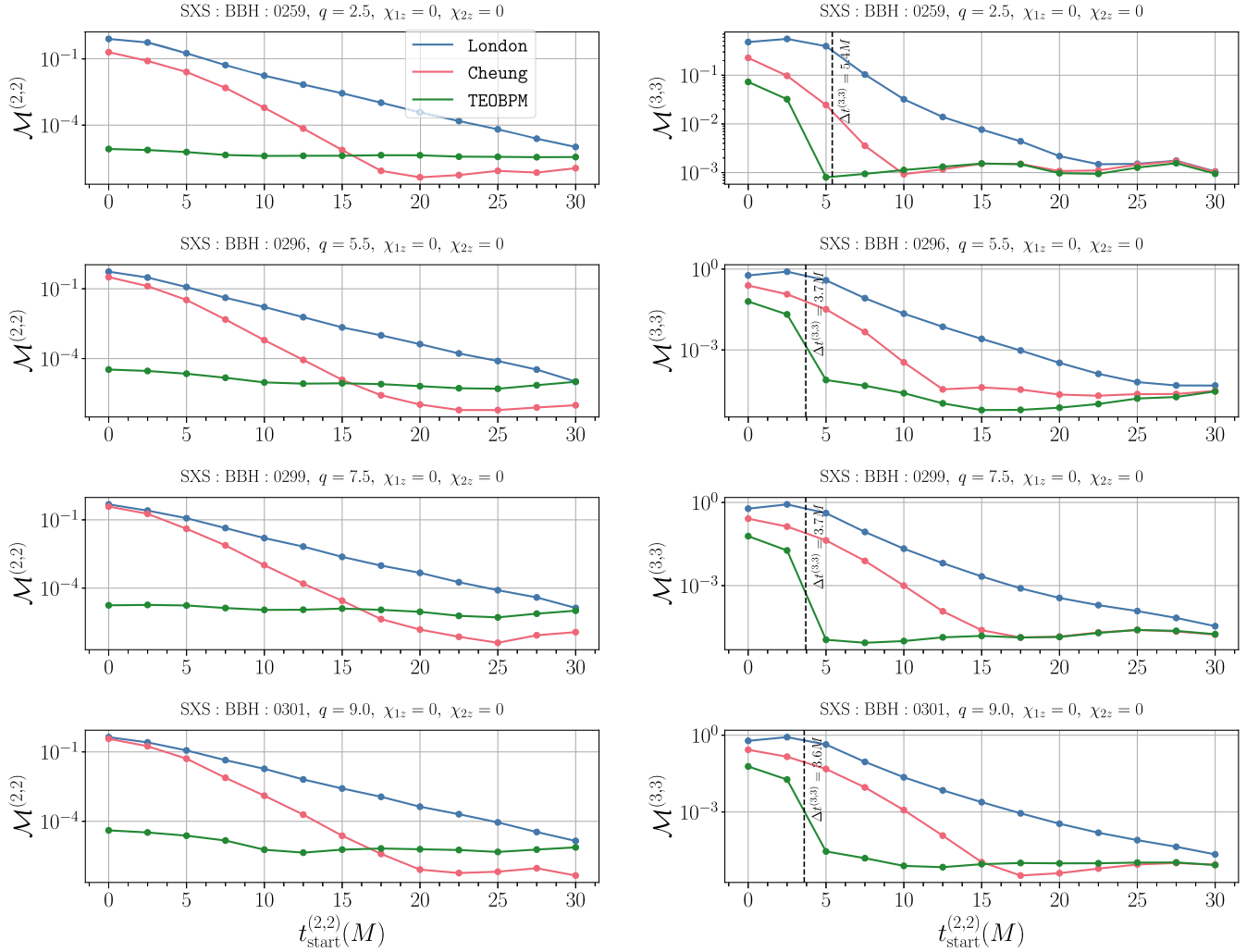


FIG. 2. Mismatches between the NR waveforms and the London (blue), Cheung (red), and TEOBPM (green) models as a function of the starting time $t_{\text{peak}}^{(2,2)}$, in units of the remnant mass. The left column refers to $(\ell, |m|) = (2, 2)$, and the right column to $(\ell, |m|) = (3, 3)$. In both cases, the origin of the starting time is taken to be the peak of the $(2, 2)$ waveform, $t_{\text{peak}}^{(2,2)}$. Different rows refer to representative nonspinning NR simulations with varying mass ratios. Dashed vertical black lines mark the time delay $\Delta t^{(\ell,|m|)}$ between the peak of the $(\ell, |m|)$ multipole and the peak of the $(2, 2)$ multipole, as defined in Eq. (17).

depend on progenitor parameters, while the frequencies and damping times depend on (M, χ) as predicted by Kerr BH perturbation theory [22,24]. The London amplitudes are calibrated in the parameter space (η, χ_+, χ_-) at $t = 20 M$ after the time $t_{\text{peak}}^{(2,2)}$, at which the absolute value of the $(2, 2)$ strain reaches its maximum. The calibration was performed using 101 nonprecessing simulations from the Georgia Tech catalog, dated before 2018 [111]. In the Cheung model, the amplitudes and phases of the ringdown are fitted in the same parameter space using a subset of 188 nonprecessing simulations from the SXS catalog (we denote this subset of simulations as $\mathcal{I}_C$). Amplitudes and phases are extracted at varying starting times, depending both on the specific simulation and on the stability of the mode under investigation. In practice, this means selecting a time region where the mode amplitudes have stabilized—i.e., where the

signal is well described by a QNM superposition. We will identify those models, using the same nomenclature as in Ch. 6 of Ref. [27], as belonging to the KERRBINARY class.

Note that in this work we use simulations from the SXS catalog to quantify the accuracy of the models. This can favor the Cheung model, as it was built using waveforms belonging to the same catalog. To mitigate it, in some sections of this work, we will only test the models on a subset of the SXS simulations that were not used to build the Cheung model.

The TEOBPM model belongs to the class of NR-informed models covering the entire postpeak regime, identified in Chapter 6 of Ref. [27] as the KERRPOSTMERGER class. Beside modeling the $n = 0$ QNM amplitude as a function of the progenitor parameters (just like the KERRBINARY models), the model effectively includes nonmodal and transient

effects that produce time-dependent amplitudes and phases $\{A_{\ell mn}, \phi_{\ell mn}(t)\}$ close to the peak of the waveform amplitude. Further, the TEOBPM model depends on $(m_1, m_2, \chi_1, \chi_2)$.

For each simulation, and for each model, we vary the starting time t_{start} between $t_{\text{peak}}^{(2,2)}$ and $t_{\text{end}} = 80$ M in steps of 2.5M, and compute the mismatches with respect to the plus and cross polarizations of the signal, denoted by $\mathcal{M}_+^{\ell, m}(t_{\text{start}}|t_{\text{end}}, \theta)$ and $\mathcal{M}_\times^{\ell, m}(t_{\text{start}}|t_{\text{end}}, \theta)$, respectively. Notice that, since we have aligned in time the model and the NR strain, and the phase between the two is estimated via parameter inference, we are indeed computing a mismatch, not an overlap. Since the two quantities are simply related by a phase shift, we will plot the *averaged* mismatch, defined as

$$\mathcal{M}^{(\ell, |m|)} = \frac{\mathcal{M}_+^{(\ell, |m|)} + \mathcal{M}_\times^{(\ell, |m|)}}{2}. \quad (16)$$

In Fig. 2 we show some examples of the mismatches $\mathcal{M}^{(2,2)}$ and $\mathcal{M}^{(3,3)}$ as a function of the starting time for nonspinning SXS simulations with different mass ratios. Let us first focus on the (2,2) (left) panels. The mismatches of the ringdown models London and Cheung generally decrease with the starting time, since QNM superpositions are a good description of the signal far from the peak. On the contrary, the mismatch of the TEOBPM model is approximately flat, because the model is designed to also reproduce the merger and early postmerger parts of the signal. The mismatches of the Cheung model are lower than those of the London model at large enough starting times, roughly by a factor $\mathcal{O}(10\text{--}100)$. The general trends for the (3, 3) multipoles for the KERRBINARY models (right panels) are similar to the (2, 2). The TEOBPM mismatch with respect to the numerical (3, 3) multipole shows a rapid decrease at early starting times. This is because higher multipoles with $(\ell, |m|) \neq (2, 2)$ typically peak after $t_{\text{start}} = t_{\text{peak}}^{(2,2)}$, with a time delay

$$\Delta t^{(\ell, |m|)} = t_{\text{peak}}^{(\ell, |m|)} - t_{\text{peak}}^{(2,2)}. \quad (17)$$

The observed drop in the TEOBPM mismatch corresponds to this delay time.

A. $(\ell, |m|)$ mismatch accuracy for a given starting time

We now address the following question: *given a ring-down model, what is the mismatch with respect to the $(\ell, |m|)$ multipole of the corresponding NR simulation catalog at some fixed starting time $t_{\text{start}}^{(\ell, |m|)}$?* In particular, which of the available models yields the lowest mismatch?

Since the Cheung model was calibrated on a large fraction of SXS data, denoted as $\mathcal{I}_C$, we perform

mismatches for all models using the subset of simulations not employed for training the Cheung model,

$$\mathcal{I}_{\text{tot-C}} = \mathcal{I}_{\text{tot}} \setminus \mathcal{I}_C. \quad (18)$$

As we show in Fig. 19 in Appendix B, the results do not change appreciably if we consider also the simulations in the set $\mathcal{I}_C$: the mismatches of the Cheung model computed over the $\mathcal{I}_C$ and $\mathcal{I}_{\text{tot-C}}$ sets have similar distributions.

We apply the procedure explained at the beginning of this section to the subset $\mathcal{I}_{\text{tot-C}}$ for the dominant multipoles $(\ell, |m|) = \{(2, 2), (2, 1), (3, 3), (3, 2), (4, 4)\}$. The mismatch histograms at selected starting times for the three ringdown models are plotted in Fig. 3. Since the TEOBPM model is calibrated from the peak of each $(\ell, |m|)$ multipole, in this plot we measure the starting time $t_{\text{start}}^{(\ell, |m|)}$ with respect to the peak of the $(\ell, |m|)$ multipole, i.e., the starting times presented in Fig. 3 are defined as

$$t_{\text{start}}^{(\ell, |m|)} = t_{\text{start}} - \Delta t^{(\ell, |m|)}. \quad (19)$$

For completeness, in Appendix B we show a version of this plot in which the mismatches are computed with respect to $t_{\text{peak}}^{(2,2)}$ instead. The two plots address different questions: in Fig. 3 we ask how accurate the ringdown waveform is for each mode, while by measuring the mismatches relative to $t_{\text{peak}}^{(2,2)}$ (as in Appendix B) we can understand how the different models perform in a data-analysis setting, when TEOBPM is used to analyze the entire post- $t_{\text{peak}}^{(2,2)}$ signal [112].

For a given $(\ell, |m|)$, the distribution of the London and Cheung mismatches is peaked at progressively lower values for higher starting times. This is particularly evident for the $(\ell, |m|) = (2, 2)$ mode, and it occurs because the amplitude fits in the London model were performed at $t_{\text{start}} = 20$ M after the peak: the larger the starting time, the closer we are to the fitting dataset. By contrast, as shown in Fig. 2, TEOBPM has roughly similar mismatch distributions at different starting times. The London model has a lowest mismatch of $\mathcal{O}(10^{-4})$ for $t_{\text{start}} = 20$ M, while the Cheung model can have mismatches as low as $\lesssim 10^{-7}$, performing better than TEOBPM at ≈ 20 M. Our computed TEOBPM mismatches at $t_{\text{start}} = 0$ M are in good agreement with those computed in Ref. [112], which did not apply a treatment of the PSD boundaries in the mismatch definition.

For the $(\ell, |m|) = (3, 3)$, and (4, 4) multipoles, TEOBPM is more accurate than the London and Cheung models at early times, and comparable to those models at late times. The only exception is the (4, 4) multipole, for which the Cheung model mismatch histograms show a tail at low values. This is likely due to the inclusion of quadratic modes, which are absent in the TEOBPM model.

The trend is reversed in the (3, 2) multipole, for which TEOBPM has the largest mismatches at all times.

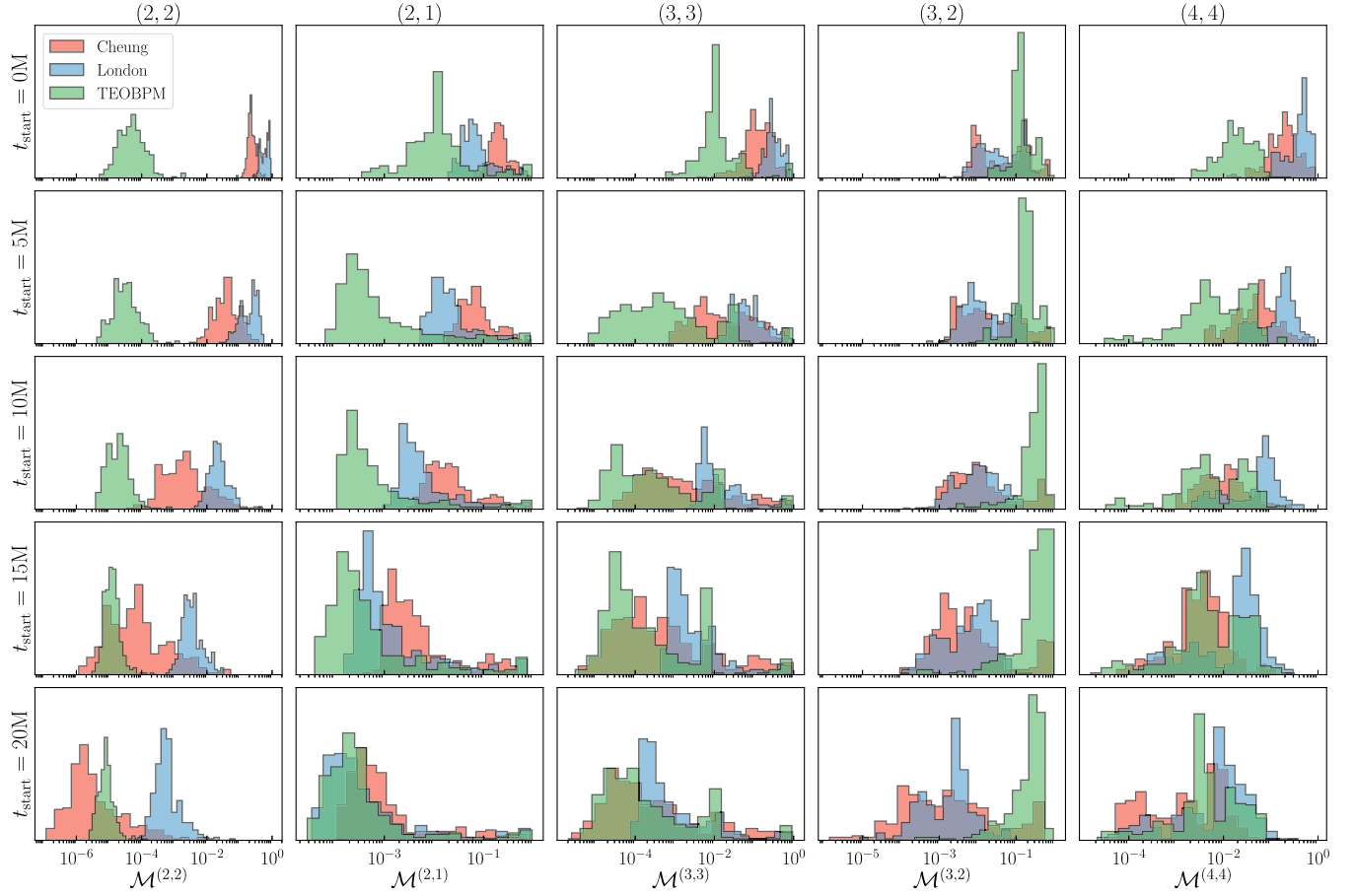


FIG. 3. Histograms of mismatches, computed with the simulation set $\mathcal{I}_{\text{tot-C}}$, of the London (blue), the Cheung (red), and TEOBPM (green) models at different $(\ell, |m|)$, corresponding to different columns. On different rows, moving from top to bottom, we consider different starting times, defined with respect to the given $(\ell, |m|)$ mode, from 0 to 20 M, at steps of 5 M.

As pointed out in Ref. [112], this is due to the lack of mode mixing, which is instead included in the KERRBINARY set of models.

The $(\ell, |m|) = (2, 1)$ multipole has an interesting behavior: the Cheung mismatches are worse than the London mismatches for most simulations. This is because the London model includes the $(2, 1, 0)^+$ mode only, while the Cheung model includes also the $(2, 1, 0)^-$ and $(2, 1, 1)^+$ modes. The $(2, 1, 0)^-$ mode does not contribute much compared to the other two modes. The $(2, 1, 1)^+$ mode does contribute, with an amplitude that can be up to one order of magnitude higher than the $(2, 1, 0)^+$ mode when extrapolated back to the waveform peak time. Therefore, if the $(2, 1, 1)^+$ mode is not as precisely modeled as the $(2, 1, 0)^+$ mode, its inclusion in the model can actually hurt the model's accuracy for two main reasons. First, the $(2, 1, 1)^+$ mode was not confidently found in a significant fraction of simulations in Ref. [51], especially in the regions of the parameter space corresponding to large χ_{eff} or $\eta \sim 1/4$ (see the figure hosted online at https://mhychung.github.io/jaxqualin/mode_md/2.1.1). The fact that there are not enough simulations with high χ_{eff} to infer the $(2, 1, 1)^+$ mode

amplitude means that the polynomial fit model is extrapolated, possibly giving unreliable results. Second, the uncertainty of the extracted amplitudes for the $(2, 1, 1)^+$ mode is relatively high compared to all other modes listed in Ref. [51], since the $(2, 1)$ multipole has low amplitude and is more prone to being contaminated by mode mixing. This reduces the accuracy of the model even when extrapolation is not necessary. At the end of this section, we will discuss what happens to the mismatch distribution when we remove one mode at a time.

B. Starting time of ringdown models from fixed mismatch threshold

In this subsection, we address a complementary question regarding KERRBINARY models: *what is the earliest starting time t_{start} at which a target mismatch can be achieved?* Our analysis will determine the time intervals over which a given model provides sufficient agreement with NR simulations. In practice, we will find that if the chosen mismatch threshold is too stringent, there may be regions of the parameter space (η, χ_+, χ_-) where the model fails to meet the desired accuracy for *any* starting time.

TABLE I. Mean $\pm 1 - \sigma$ values of the starting time t_{start} at which the KERRBINAR models drop below a given mismatch threshold compared to the NR simulations.

$\mathcal{M}_{\text{thr}}$	Model	$t_{\text{start}}^{(2,2)} [M]$	$t_{\text{start}}^{(2,1)} [M]$	$t_{\text{start}}^{(3,3)} [M]$	$t_{\text{start}}^{(3,2)} [M]$	$t_{\text{start}}^{(4,4)} [M]$
10.0%	London	6.7 ± 1.7	-1.0 ± 4.0	4.5 ± 3.1	-1.6 ± 6.4	7.5 ± 4.7
	Cheung	3.5 ± 1.9	4.4 ± 4.4	1.9 ± 4.5	-1.3 ± 6.8	3.3 ± 4.0
3.5%	London	9.2 ± 1.7	3.3 ± 4.0	7.4 ± 3.5	2.2 ± 7.6	12.0 ± 5.3
	Cheung	5.9 ± 2.2	8.5 ± 4.5	4.4 ± 4.9	0.7 ± 8.3	7.4 ± 5.7
0.5%	London	14.4 ± 2.4	10.4 ± 3.8	12.0 ± 2.5	11.1 ± 6.4	17.5 ± 6.0
	Cheung	9.4 ± 2.6	13.9 ± 3.5	6.8 ± 3.4	9.2 ± 5.6	13.9 ± 5.3
0.1%	London	18.9 ± 2.6	14.7 ± 3.4	16.4 ± 2.4	15.1 ± 4.5	16.6 ± 4.8
	Cheung	11.9 ± 2.9	17.7 ± 2.3	10.3 ± 4.4	15.7 ± 3.9	16.9 ± 4.0

1. Starting time statistics for a given mismatch threshold

To address the above question, we generate a histogram of the starting times $t_{\text{start}}^{(\ell, |m|)}$ at which the mismatch reaches a threshold value of $\mathcal{M}_{\text{thr}} = [10.0, 3.5, 0.5, 0.1]\%$ (if at all). The value $0.035 \approx 1 - 0.9^{\frac{1}{3}}$ is associated with the loss of 10% of the match in a cubic volume. The other values were instead chosen (somewhat arbitrarily) to probe the models' accuracy. We only report results for the KERRBINAR models: the need to introduce a starting time arises because these models lack accuracy at early times. To have a fair comparison, we compare the London and Cheung models over the dataset $\mathcal{I}_{\text{tot-C}}$. A summary of our results is presented in Table I.

Let us first focus on the (2, 2) mode. For a given threshold, the Cheung model typically achieves a lower mean starting time with respect to the London model. The reason is clear from Fig. 2: the mismatch as a function of the starting time decays faster for the Cheung model, and therefore $t_{\text{start}}(\mathcal{M} = \mathcal{M}_{\text{thr}})$ is lower.

This is not always true for higher modes. The starting times corresponding to a given threshold mismatch are often comparable, and for $(\ell, |m|) = (2, 1)$ case the London model actually performs better. This is consistent with the second column of Fig. 3: there the distribution of the mismatches for the Cheung model peaks at higher values with respect to the corresponding distribution for the London model.

2. Fraction of simulations that reach a given mismatch threshold

It is also interesting to quantify the fraction of simulations that could reach a given mismatch threshold for a suitable choice of starting time. For each $(\ell, |m|)$ multipole, we compute the quantity

$$N_{\text{cross}} = \frac{\{\# \text{simulations } |\mathcal{M} \leq \mathcal{M}_{\text{thr}}\}}{\{\# \text{total number of simulations}\}} \quad (20)$$

as a function of $\mathcal{M}_{\text{thr}}$, using the minimum starting time at which the threshold is met. The result for the various

models is shown in Fig. 4. The (2, 2) mode has N_{cross} very close to unity when $\mathcal{M}_{\text{thr}} \gtrsim 10^{-5}$ for the TEOBPM and Cheung models, and at $\mathcal{M}_{\text{thr}} \gtrsim 10^{-3}$ for the London model. The implication is that the models are accurate for all simulations, at least at that level. For lower mismatches, N_{cross} drops below unity: the required mismatch is never achieved for some of the simulations, regardless of the starting time. The mismatch drop of the Cheung model is slower than for the TEOBPM model (see the bottom left panel of Fig. 3). For $(\ell, |m|) \neq (2, 2)$, N_{cross} drops below unity at higher mismatch thresholds, confirming that the models are less accurate for higher harmonics.

3. Starting times such that KERRBINAR models have the same mismatch as TEOBPM

As we pointed out when we discussed Fig. 2, there are two clear trends for any given SXS simulation: (1) the TEOBPM mismatches, especially for the (2, 2) mode, are nearly constant as a function of the starting time, since the effect of nonlinearities and transient effects are included in the modeling from the peak; (2) the London and Cheung models mismatches decrease as functions of the starting time, because perturbation theory becomes a better approximation at late times. In general, there will be some critical starting time at which the mismatch of the QNM-based models becomes comparable with the mismatch of the TEOBPM model. Here we quantify the distribution of such starting times, defined as

$$t_{\text{start,eq}}^{(\ell, |m|)} = t_{\text{start}}^{(\ell, |m|)}(\mathcal{M}_{\text{KERRBINAR}}^{(\ell, |m|)} = \mathcal{M}_{\text{TEOBPM}}^{(\ell, |m|)}). \quad (21)$$

At these times, roughly speaking, the QNM part of the signal should become dominant in the TEOBPM model.

In Fig. 5 we plot histograms of this quantity. In Table II we list the fraction of the total set of simulations $\mathcal{I}_{\text{tot}}$ such that $\mathcal{M}_{\text{KERRBINAR}}^{(\ell, |m|)} = \mathcal{M}_{\text{TEOBPM}}^{(\ell, |m|)}$. For the (2, 2) mode, the Cheung model distribution peaks around $(15 \pm 5)M$, with $\approx 99\%$ of the simulations such that $\mathcal{M}_{\text{Cheung}}^{(\ell, |m|)} = \mathcal{M}_{\text{TEOBPM}}^{(\ell, |m|)}$. By contrast, the London model achieves comparable

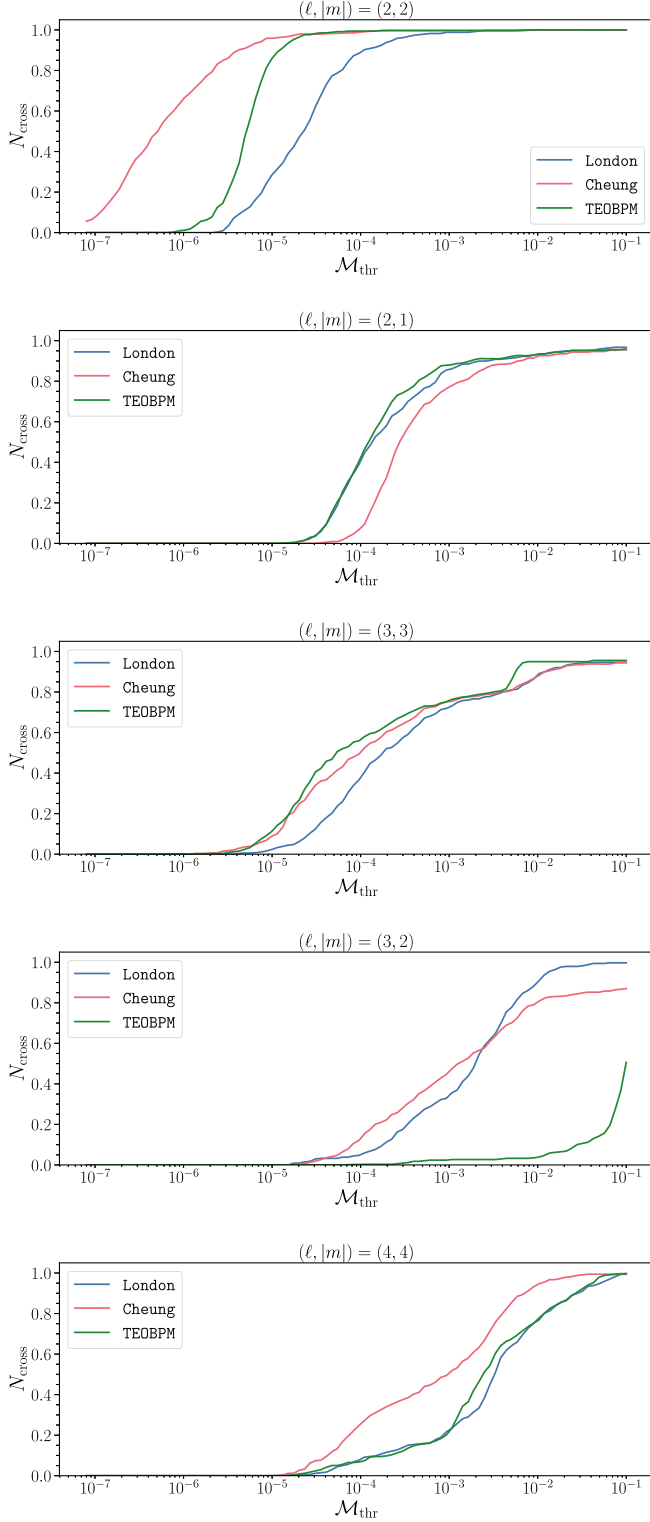


FIG. 4. Number of simulations N_{cross} that can cross a given mismatch threshold for different $(\ell, |m|)$ harmonics, computed at the minimum starting time at which the mismatch threshold is achieved.

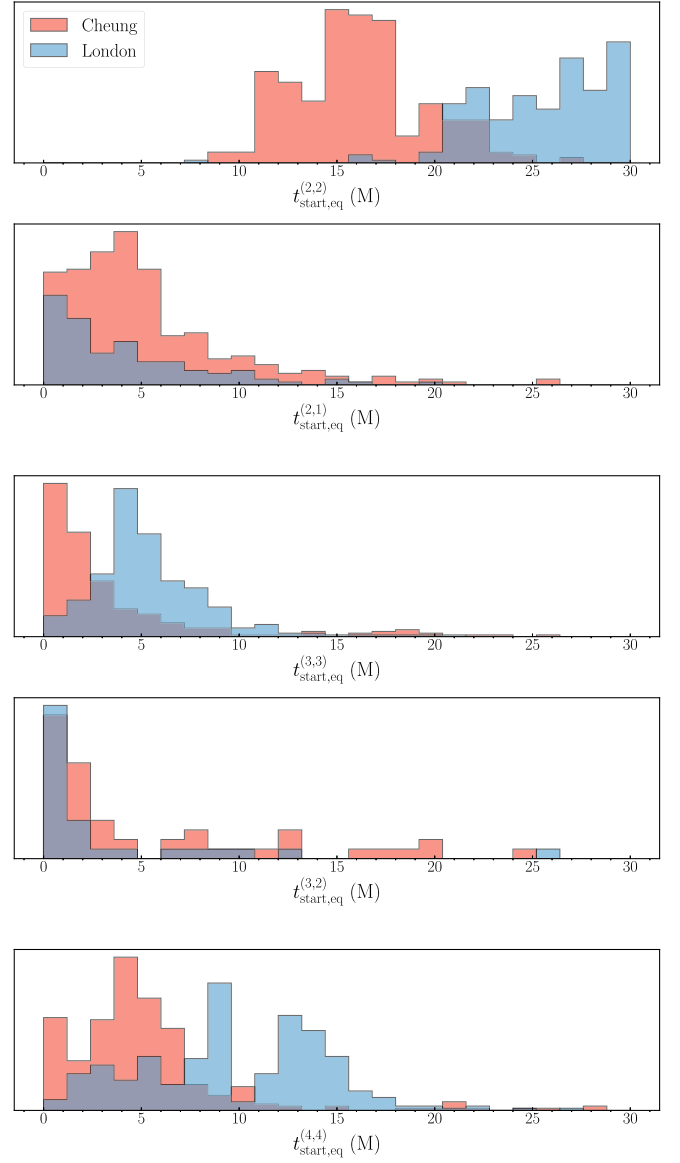


FIG. 5. Histograms of the starting times $t_{\text{start,eq}}^{(\ell, |m|)}$ such that the Cheung (red) and London (blue) model mismatches are equal to the TEOBPM model mismatches, as defined in Eq. (21).

accuracy to TEOBPM at later times, with $\approx 68\%$ of the simulations yielding $\mathcal{M}_{\text{London}}^{(\ell, |m|)} = \mathcal{M}_{\text{TEOBPM}}^{(\ell, |m|)}$. The starting time distributions shift toward earlier times for the (2, 1), (3, 3), and (4, 4) multipoles. This reflects the lower accuracy of both the TEOBPM and the KERRBINARY models at early times for these higher harmonics. For the (3, 2) mode, the mismatches obtained with TEOBPM are systematically worse than those of the London and Cheung models across all times. This is why the starting time distributions found for this multipole in Fig. 5 are tightly concentrated at low values.

TABLE II. Fraction of the simulations such that the Cheung and London models have the same mismatch as TEOBPM at some starting time, for different multipoles.

(ℓ, m)	$\%(\mathcal{M}_{\text{Cheung}}^{(\ell, m)} = \mathcal{M}_{\text{TEOBPM}}^{(\ell, m)})$	$\%(\mathcal{M}_{\text{London}}^{(\ell, m)} = \mathcal{M}_{\text{TEOBPM}}^{(\ell, m)})$
(2, 2)	99%	68%
(2, 1)	97%	98%
(3, 3)	98%	99%
(3, 2)	35%	32%
(4, 4)	98%	98%

C. The role of mode exclusion for the Cheung model

We now address another important question: *given a spherical multipole $(\ell, |m|)$, which QNMs contribute most significantly to the agreement between a given ringdown model and the NR signal?* From Eq. (4) we see that each $(\ell, |m|)$ multipole generally receives contributions from multiple QNMs (ℓ', m, n) , coming from either higher overtones or mode mixing.

For a given $(\ell, |m|)$ multipole, associated with a list of contributing QNMs, we systematically remove individual (ℓ, m, n) modes from the Cheung model waveform. We list the QNMs implemented in `pyRing` for each $(\ell, |m|)$ multipole in Table III. For each $(\ell, |m|)$, we remove the last $(\ell, m, n)^\pm$ QNM of the row: e.g., for the (2, 2) harmonic we first remove $(3, 2, 0)^+$, then $(2, 2, 0)^-$, and so on. We then quantify how much the mismatch distribution of the Cheung model with respect to the NR waveform is affected by each removal. If the mismatch distribution remains essentially unchanged upon removal, the corresponding QNM can be considered globally subdominant; conversely, if its removal leads to a significant degradation in the mismatch, the mode should be thought of as dominant.

We perform this analysis for different starting times. The results are presented in Fig. 6. The findings are strongly dependent on the specific $(\ell, |m|)$ multipole considered, and we will discuss them on a case-by-case basis.

The (2, 2) multipole mismatch is only mildly affected by the removal of the $(3, 2, 0)^+$ QNM, particularly at late times. Removing the counter-rotating mode $(2, 2, 0)^-$ does not produce any appreciable change. When we remove the $(2, 2, 1)^+$ overtone (purple histogram), the mismatch at early times increases; however, as t_{start} approaches ~ 30 M the impact of removing the overtone diminishes, and the mismatch distributions of the $(2, 2, 0)^+$ mode alone and the combined $(2, 2, 0)^+ + (2, 2, 1)^+$ modes tend to overlap. This makes sense: the overtone decays faster than the fundamental mode, and thus its contribution becomes less dominant at late times.

On the contrary, for the (2, 1) mode, removing the first overtone $(2, 1, 1)^+$ improves the accuracy at intermediate times. This is due to the poor accuracy of the $(2, 1, 1)^+$ amplitude fit discussed in Sec. III A. At late times the

TABLE III. Dominant QNMs contributing to the various multipoles $(\ell, |m|)$ for the Cheung model.

(ℓ, m)	Contributing $(\ell, m, n, \pm)$ modes
(2, 2)	$(2, 2, 0)^\pm, (2, 2, 1)^+, (3, 2, 0)^+$
(2, 1)	$(2, 1, 0)^\pm, (2, 1, 1)^+$
(3, 3)	$(3, 3, 0)^+, (3, 3, 1)^+$
(3, 2)	$(3, 2, 0)^+, (2, 2, 0)^+$
(4, 4)	$(4, 4, 0)^+, ((2, 2, 0)^+ \times (2, 2, 0)^+)$

overtone removal has a smaller impact, since the overtone has decayed.

The (3, 3) mode consists of a superposition of the $(3, 3, 0)^+$ and $(3, 3, 1)^+$ QNMs. The trend is similar to the previous cases: the removal of the overtone $(3, 3, 1)^+$ is significant at early times, but not so much at late times.

In the case of the (3, 2) multipole, removing the $(2, 2, 0)^+$ QNM results in a mismatch distribution approaching $\mathcal{O}(1)$. This dramatic loss of accuracy is expected: the Cheung model includes mode mixing between the $(3, 2, 0)^+$ and $(2, 2, 0)^+$ modes, which is the dominant contribution to this mode in nonprecessing configurations.

Finally, the (4, 4) multipole has contributions from the linear mode $(4, 4, 0)^+$ and from the quadratic mode $(2, 2, 0)^+ \times (2, 2, 0)^+$. When we remove the quadratic mode, the mismatch distribution gets somewhat worse at intermediate times (around ≈ 15 M). At later times the quadratic mode has decayed, and its removal has little to no impact on the mismatch.

To quantify the loss of mismatch due to removing the quadratic mode, in Fig. 7 we plot the ratio

$$r_{(4,4)} = \frac{\mathcal{M}_{(4,4,0)^+}}{\mathcal{M}_{((4,4,0)^+, (2,2,0)^+ \times (2,2,0)^+)}} \quad (22)$$

for different selected values of t_{start} , in the subspace (η, χ_+) . As the starting time increases, the ratio $r_{(4,4)}$ gets larger than unity for $\eta \sim 0.25$. This is consistent with Refs. [58,59], where the inclusion of the quadratic mode was found to improve the agreement with the SXS: BBH:0305 simulation. The fact that the quadratic mode matters more at comparable masses is because its amplitude scales as $\sim A_{(2,2,0)^+}^2$, and that $(2, 2, 0)^+$ mode is more dominant for comparable-mass binaries.

D. Application to LIGO-Virgo-KAGRA analyses

Our results can also find applications to observational analyses with LVK data. In this context, one typically considers parametrized deviations in the QNM frequencies or amplitudes, as these parameters have a direct physical interpretation. As starting at earlier times yields higher

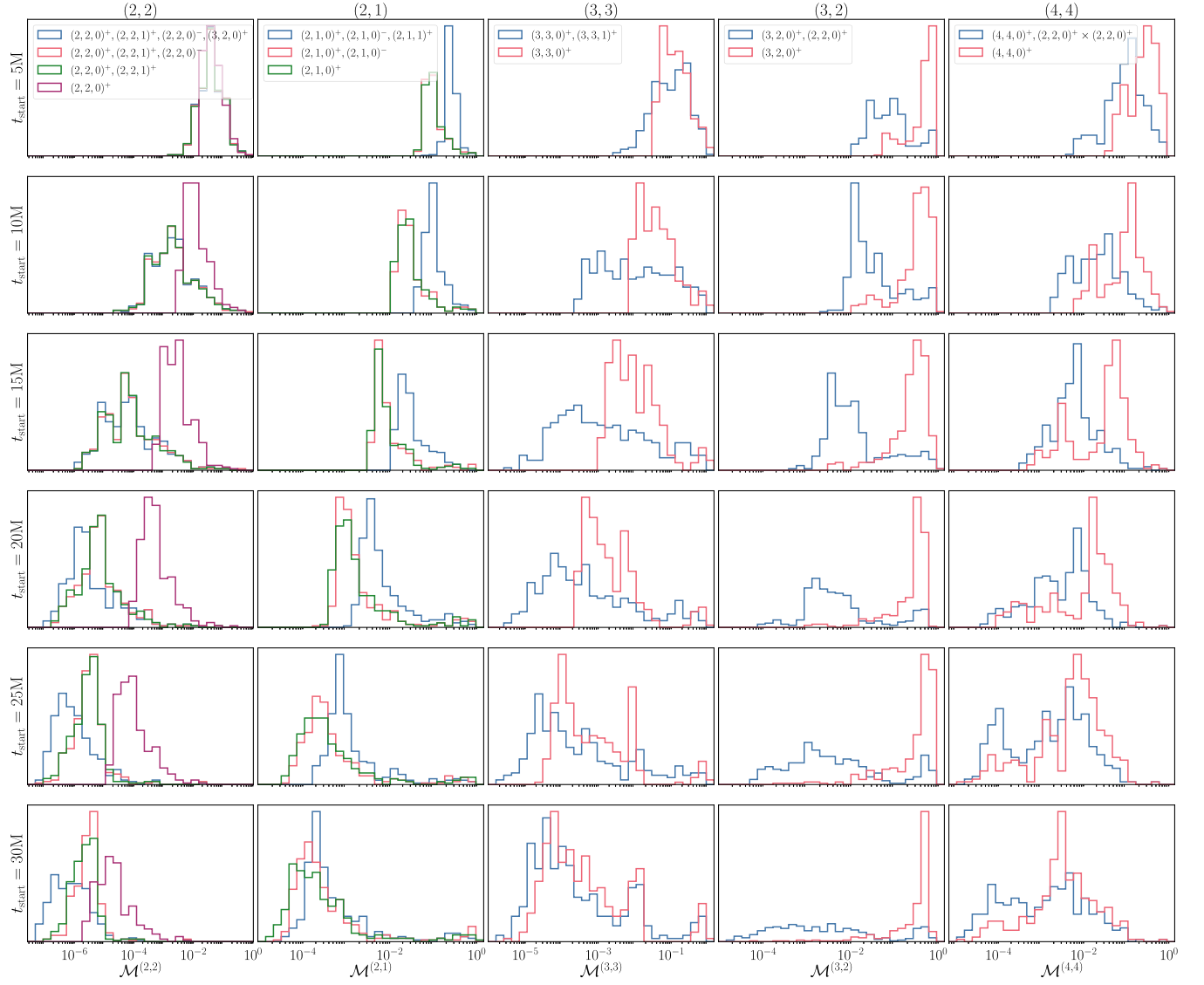


FIG. 6. Cheung mismatch distributions for different $(\ell, |m|)$ multipoles (as indicated at the top of each column), computed at discrete values of t_{start} ranging from 0M to 30M (as indicated to the left of each row). Different colors denote different sets of QNMs (see Table III).

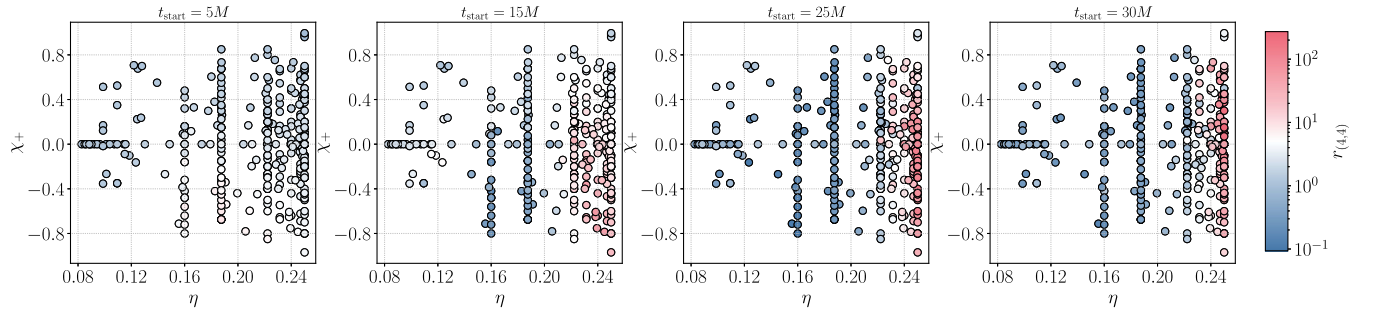


FIG. 7. Ratio $r_{(4,4)}$ in the (η, χ_+) plane, as defined in Eq. (22), for four selected starting times: $t_{\text{start}} = [5, 15, 25, 30]$ M.

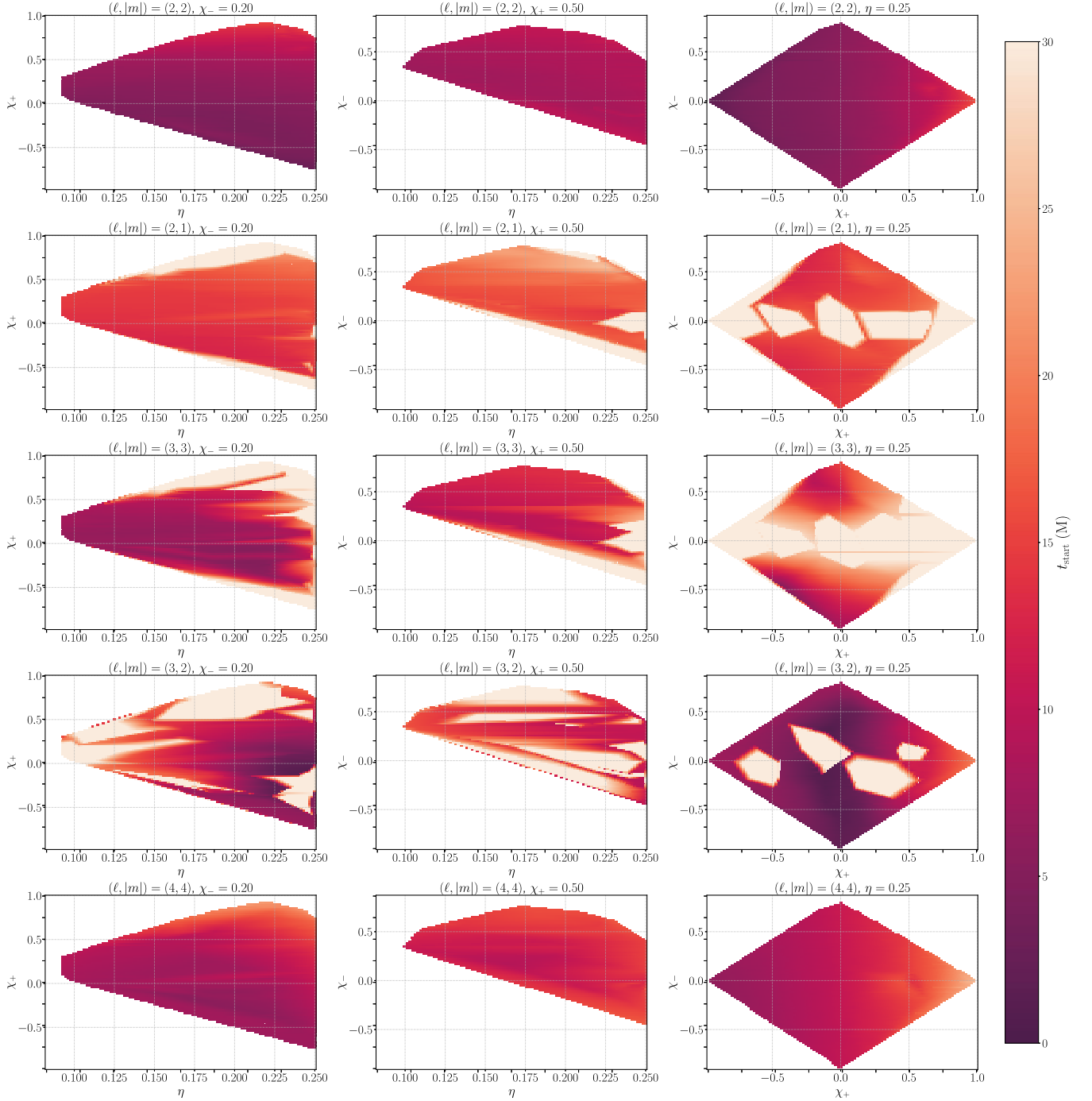


FIG. 8. Colormaps of $t_{\text{start}}(\mathcal{M} = 3.5\%)$ for the Cheung model in different “cuts” of the parameter space: (η, χ_+) (left column), (η, χ_-) (middle column), and (χ_+, χ_-) (right column). Different rows refer to different multipoles. In this case we use the whole set of SXS simulations, $\mathcal{I}_{\text{tot}}$.

SNR, a key parameter is the minimum starting time at which the analysis can be applied, while still maintaining a sufficient level of accuracy against GR numerical solutions to avoid biases.

Here we provide a simple tool to answer this question. We tabulate values of the starting time $t_{\text{start}}(\eta, \chi_+, \chi_- | \mathcal{M}_{\text{th}})$ at which a given mismatch threshold $\mathcal{M}_{\text{th}}$ is reached for the

entire SXS dataset under consideration. We then interpolate across these discrete values, so we can perform this estimate over the entire parameter space.

To the simulations that do not achieve the given mismatch threshold, we assign a “saturation” starting time of 100M, indicating that in those regions we cannot reach the required accuracy. To guarantee interpolation accuracy, we

need as many points as possible, and therefore, we use the entire $\mathcal{I}_{\text{tot}}$ dataset. These results are available online in the repository [95] for the models considered in this study.

An example of interpolation for the Cheung model is shown in Fig. 8: the interpolant is well reconstructed for $(\ell, |m|) = (2, 2)$, but fails for higher multipoles in certain regions of the parameter space and at sufficiently low mismatch, given the sparsity of available simulations at which the required accuracy can be achieved (see Fig. 4).

To test the accuracy of the interpolation across the parameter space, we first divided the SXS dataset $\mathcal{I}_{\text{tot}}$ into a *training* (*validation*) set, containing 80% (20%) simulations randomly distributed in the parameter space. We performed the interpolation on the training set. Then we plotted the distributions of the residuals in t_{start} —defined as the difference between the mismatch threshold’s starting time and the starting time found by interpolation—for the validation set. The results are shown in Fig. 18 of Appendix B.

IV. DISCUSSION AND CONCLUSIONS

We have investigated the accuracy of several state-of-the-art ringdown models by comparing them to binary BH merger simulations from the SXS catalog [110]. As a measure of accuracy, we have computed the time-domain mismatch defined in Refs. [89,90] between these models and the numerical waveforms, thus avoiding the conceptual and practical issues that arise when we analyze ringdown-only signals in the frequency domain.

We focused on the London [49], Cheung [51], and TEOBPM [81,82,85,86] models, all developed for non-precessing, quasicircular progenitor systems. The first two are constructed as QNM superpositions with amplitudes and phases fitted to NR simulations, and are expected to be valid at sufficiently late times after merger. In contrast, TEOBPM includes the merger and early postmerger regime, with a phenomenological time-domain parametrization calibrated to NR simulations.

We first analyzed the distribution of mismatches as a function of the starting time of the ringdown for different $(\ell, |m|)$ multipoles across an ensemble of aligned-spin, quasicircular SXS simulations. We found that the Cheung model gives lower mismatches than the London model for most binaries. This is expected, as the Cheung model was built by fitting waveforms in the same SXS catalog used to perform the mismatch tests and because it includes overtones, which improve the match for most waveform multipoles, and quadratic modes. However, the London model outperforms Cheung for the $(2, 1)$ mode waveform. This is because the Cheung model includes the $(2, 1, 1)$ overtone, which is robustly found only in a low number of simulations, resulting in a poor fitting of its overall amplitude. Consistently with the findings of Ref. [112], the mismatch of the TEOBPM model is excellent for the dominant harmonic even at early times, but its accuracy is lower

for some of the higher modes. This is due to the lack of mode-mixing effects and additional intermediate-time contributions (such as the $n = 1$ overtone). Excluding the $(3, 2)$ multipole, in which the mode mixing is dominant, TEOBPM generally outperforms the other two models, except at late times ($t \gtrsim 15\text{--}20$ M), where it does not fit the waveform as well because it only includes a single QNM.

Next, we studied the fraction of simulations N_{cross} for which the ringdown models cross a given mismatch threshold. For the $(\ell, |m|) = (2, 2)$ multipole and across all starting times considered, the London model achieves a mismatch of $\gtrsim 10^{-3}$ for essentially all simulations; the TEOBPM model has mismatches better than $\gtrsim 10^{-5}$, and the Cheung model does slightly better than TEOBPM due to its extended QNM content (see Table III). For the $(3, 2)$ multipole, the number of simulations that reach a given mismatch threshold for TEOBPM is always lower than for the two KERRBINARY models. We provide an interpolating function to estimate the earliest starting time for which a given QNM-based model meets a given mismatch threshold across the parameter space (η, χ_+, χ_-) [95]. This should be valuable information for ringdown searches in LVK data.

Finally, we have investigated the relative importance of individual QNMs in the Cheung model by progressively removing overtones and evaluating their impact on the mismatch. We have found that excluding overtones degrades the model accuracy at early times, but not so much at late times. This is expected, as overtones decay faster than the fundamental mode. In the future, it will be important to repeat this sort of study by removing the modes in different combinations and assessing which modes are most relevant in other regions of the parameter space. An interesting case is the $(4, 4)$ multipole, that contains a significant quadratic $(2, 2, 0)^+ \times (2, 2, 0)^+$ QNM component in addition to the $(4, 4, 0)^+$ QNM. We have found that removing the quadratic mode from the $(4, 4)$ harmonic leads to higher inaccuracy at $\eta \simeq 0.25$, but not for more asymmetric systems. This is mainly because $A_{(2,2,0)^+ \times (2,2,0)^+} \propto A_{(2,2,0)^+}^2$, and $A_{2,2,0}$ is larger for comparable mass binaries, but it will be important to better understand possible instabilities in the quadratic mode fits.

We restricted attention to quasicircular, nonprecessing binary BHs, but our analysis should be extended to include precession [53,113,114] and eccentricity [69]. For instance, in precessing systems the $(2, 2, 0)^+$ mode may no longer dominate over some of the higher multipoles. Neglecting eccentricity or precession could also bias the determination of the optimal starting times obtained via the interpolating function. We leave these extensions to future work.

As a useful rule of thumb, for a signal with a given SNR, two waveforms are considered indistinguishable for parameter estimation purposes if their mismatch [115,116]

$$\mathcal{M} \lesssim \frac{1}{2 \text{SNR}^2}. \quad (23)$$

Since this criterion is sufficient but not necessary, it is generally too conservative (see, e.g., [117]). Furthermore, its violation does not necessarily imply that differences are measurable, and detectable effects may be confined to a subset of parameters that could be of little interest [118].

Nonetheless, the above indistinguishability criterion suggests that events with ringdown $\text{SNR} \approx 20$, such as GW250114 [28,29], would require $\mathcal{M} \lesssim 10^{-3}$ to avoid waveform systematics. Our results show that the Cheung and TEOBPM models typically have a smaller mismatch with respect to NR waveforms for $(\ell, |m|) = (2, 2)$, when $t_{\text{start}} \gtrsim 15$ M and $t_{\text{start}} \gtrsim 0$ M, respectively, while this is not the case for the London model. We can conclude that using the TEOBPM model for large-SNR, GW250114-like postmerger detections with LIGO-Virgo-KAGRA should not introduce large modeling systematics, at least for the fundamental mode. Instead, given the overall larger mismatches for higher harmonics, care should be applied when performing BH spectroscopy with those.

The outlook changes significantly when considering the landscape of future detectors. Next-generation detectors such as ET and CE can reach ringdown $\text{SNR} \approx 100$ for several events per year [119], whereas space missions like LISA [46] can even achieve ringdown $\text{SNR} \approx 1000$ in certain optimistic scenarios [120,121]. In these cases one would need $\mathcal{M} \lesssim 5 \times 10^{-5}$ and $\mathcal{M} \lesssim 5 \times 10^{-7}$, respectively, to make sure that the adopted waveform model is sufficiently accurate. Our results show that this level of accuracy is hard to achieve even for $(\ell, |m|) = (2, 2)$, and even more so for $(\ell, |m|) = (3, 3)$. This calls for a drastic improvement of ringdown waveforms if one wishes to perform precision BH spectroscopy with future instruments.

Note added. While this work was being completed, the *sxs* collaboration updated the catalog by nearly doubling the total number of binary configurations from 2018 to 3756, covering more densely the region of parameter space involving asymmetric, precessing, and eccentric binaries [106]. Each waveform was also corrected to make sure that it refers to the binary’s center-of-mass frame, and updated to include the memory effect (see [122]). We have analyzed how these improvements affect the mismatch for the Cheung. The results of this comparison are reported in Appendix B, and they show that our main conclusions (obtained from the earlier version of the catalog) do not change appreciably. This is not unexpected: at the level of accuracy we are interested in, the most significant catalog update concerns the inclusion of precessing simulations, which we do not consider in the present study. Finally, since we are using the last catalog version before the new release [106], we also exclude simulation *SXS:BBH:1110* from our analysis due to significant numerical artifacts in the $(2, 2)$ mode waveform (this simulation was removed in the latest release).

ACKNOWLEDGMENTS

We thank Fabrizio Corelli, Konstantinos Kritos, Francesco Iacovelli, Luca Reali, Sophia Yi, and Jay Wadekar for discussions. F. C. acknowledges Johns Hopkins University and the University of Birmingham for the kind hospitality during the finalization of this project. F. C. acknowledges the financial support provided under the “Progetti per Avvio alla Ricerca Tipo 1”, Protocol No. AR12419073C0A82B. E. B. is supported by NSF Grants No. AST-2307146, No. PHY-2513337, No. PHY-090003, and No. PHY-20043, by NASA Grant No. 21-ATP21-0010, by John Templeton Foundation Grant No. 62840, by the Simons Foundation [MPS-SIP-00001698, E. B.], by the Simons Foundation International [SFI-MPS-BH-00012593-02], and by Italian Ministry of Foreign Affairs and International Cooperation Grant No. PGR01167. The Center of Gravity is a Center of Excellence funded by the Danish National Research Foundation under Grant No. DNR184. M. H. Y. C. is a Croucher Fellow supported by the Croucher Foundation. M. H. Y. C. acknowledges support by the Jonathan M. Nelson Center for Collaborative Research at the Institute for Advanced Study. This work was carried out at the Advanced Research Computing at Hopkins (ARCH) core facility ([129]), which is supported by the NSF Grant No. OAC-1920103. P. P. is partially supported by the MUR FIS2 Advanced Grant ET-NOW (CUP: B53C25001080001), the MUR PRIN Grant 2020KR4KN2 “String Theory as a bridge between Gauge Theories and Quantum Gravity,” by the MUR FARE programme (GW-NEXT, CUP: B84I20000100001), and by the INFN TEONGRAV initiative. Some numerical computations were performed at the Vera cluster supported by MUR and Sapienza University of Rome.

DATA AVAILABILITY

The data that support the findings of this article are openly available at [95].

APPENDIX A: EQUIVALENCE BETWEEN TIME DOMAIN AND FREQUENCY DOMAIN AND IMPACT OF THE WINDOW ON THE PSD EDGES

In this appendix, we investigate the impact of boundary conditions on the mismatch.

We start by asking: at what precision does the time-domain formalism of Sec. I give the same results as a frequency-domain approach, given standard choices for the treatment of boundary effects?

Consider two signals $h(t)$ and $s(t)$, assumed to be optimized over time and phase, with time-domain mismatch given by

$$\mathcal{M}_{\text{TD}} = 1 - \max_{\phi, t_c} \frac{\langle h|s \rangle}{\sqrt{\langle h|h \rangle \langle s|s \rangle}}, \quad (\text{A1})$$

where $\langle . \rangle$ denotes the scalar product in the time domain defined in Eq. (10). The Fourier-domain mismatch is instead given by [105,116,123]

$$\mathcal{M}_{\text{FD}} = 1 - \max_{\phi, t_c} \frac{(h|s)}{\sqrt{(h|h)(s|s)}}, \quad (\text{A2})$$

where the scalar product $(. | .)$ is defined as

$$(h|s) = 4\mathcal{R} \int_{f_{\min}}^{f_{\max}} \frac{\tilde{h}(f)\tilde{s}^*(f)}{S_n(f)}. \quad (\text{A3})$$

Here $\mathcal{R}$ denotes the real part, and $\tilde{h}(f)$ is the Fourier transform of $h(t)$. We construct representative signals $h(t)$ and $s(t)$ as follows:

- (i) We generate waveforms using both IMRPhenomD [124] and IMRPhenomXP [125], choosing a detector-frame total mass $M = 62M_{\odot}$, mass ratio $q = 0.7$, $d_L = 410$ Mpc, and $\chi_1 = \chi_2 = 0$. With this prescription, the frequency domain SNRs of the two signals are ~ 33 . The two signals have been chosen to represent two waveforms with small differences, due in this case to the different baseline aligned-spin parametrization.
- (ii) We align the two waveforms at the peak, and isolate the ringdown parts by tapering the left edge of the signals, to zero-out the strain. This prevents artifacts when building the fast Fourier transform of the two signals.
- (iii) The right edge of the ringdown signal is padded with zeros to match the fast Fourier transform length. In other words, the signal will be zero for a stretch of length $N_{\text{pad}} = N_{\text{ACF}} - N_{\text{signal}}$.

The two waveforms are plotted in Fig. 9 in the time domain.

We compute the mismatch on $h \equiv h_+$ both in the time and in the frequency domains. We use the PSD provided in Ref. [101], we set $f_{\min} = 5$ Hz and $f_{\max} = 4096$ Hz, and we compare the results of Eq. (A1) as implemented in bayRing [108] against Eq. (A2), accessed through PYCBC [126,127].

We find:

$$\begin{cases} \mathcal{M}_{\text{FD}} = 0.00025582775, \\ \mathcal{M}_{\text{TD}} = 0.00025582767, \end{cases} \quad (\text{A4})$$

with a relative difference

$$p_{\text{TD-FD}} = \frac{|\mathcal{M}_{\text{TD}} - \mathcal{M}_{\text{FD}}|}{\mathcal{M}_{\text{FD}}} = 3.4 \times 10^{-7}. \quad (\text{A5})$$

Such a difference between the two mismatch values could be relevant at values of the SNR of order $\mathcal{O}(10^3)$ [128] typical of next-generation detectors, such as LISA.

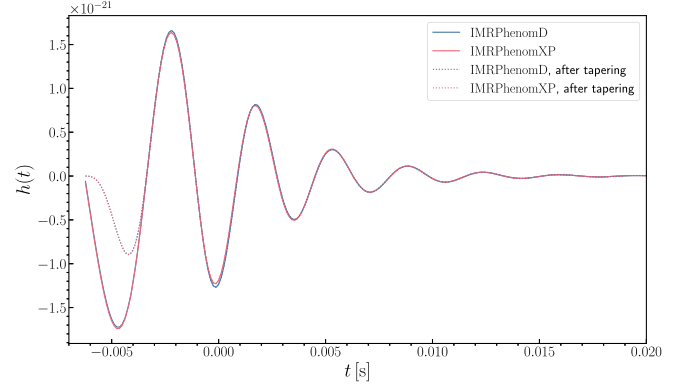


FIG. 9. Waveforms used to compute the mismatches in the time and frequency domains, respectively, either without (continuous lines) or with (dashed lines) tapering at the lower edge.

As discussed in Sec. II, it is important to apply a window function to “smooth out” the PSD edges to obtain the ACF through a fast Fourier transform. This is because simply setting the PSD to zero outside of the (finite) detector frequency band can have undesired consequences for signals whose frequency support extends beyond the sensitivity band, while setting it equal to a constant (see, e.g., Fig. 1 of [90]) can introduce Gibbs-like phenomena in the ACF. Here we demonstrate that these effects are practically negligible in our case. We replace the PSD $S(f)$ with

$$\begin{cases} S(f) \cdot e^{(f-(f_{\min}+w_l))k} + T_l, & f \in [f_{\min}, f_{\min} + w_l] \\ S(f), & f \in [f_{\min} + w_l, f_{\max} - w_h] \\ S(f) \cdot e^{-(f-(f_{\max}-w_h))k} + T_h, & f \in [f_{\max} - w_h, f_{\max}] \end{cases} \quad (\text{A6})$$

where

$$\begin{cases} T_l = S(f_{\min}) \cdot \mathcal{I}_l \cdot \frac{1-e^{-(f_{\min}+w_l)k}}{1-e^{-w_l k}}, \\ T_h = S(f_{\max}) \cdot \mathcal{I}_h \cdot \frac{1-e^{-(f_{\max}-w_h)k}}{1-e^{-w_h k}}, \end{cases} \quad (\text{A7})$$

and $\mathcal{I}_l, \mathcal{I}_h \geq 1$ are constants. The parameters (w_l, w_h) specify the window size, and k determines its steepness.

In Fig. 10 we plot an example of a windowed PSD, along with the frequency-domain signals introduced before. We can now try to determine a set of parameters $\underline{W} = (w_l, w_h, k, \mathcal{I}_l, \mathcal{I}_h)$ that maximizes the agreement between the frequency-domain and time-domain results by computing

$$p_{\text{TD-FD}}(\underline{W}) = \frac{|\mathcal{M}_{\text{TD}}(\underline{W}) - \mathcal{M}_{\text{FD}}|}{\mathcal{M}_{\text{FD}}}, \quad (\text{A8})$$

where $\mathcal{M}_{\text{FD}}$ was kept fixed, since the application of the window function only affects the time-domain calculation.

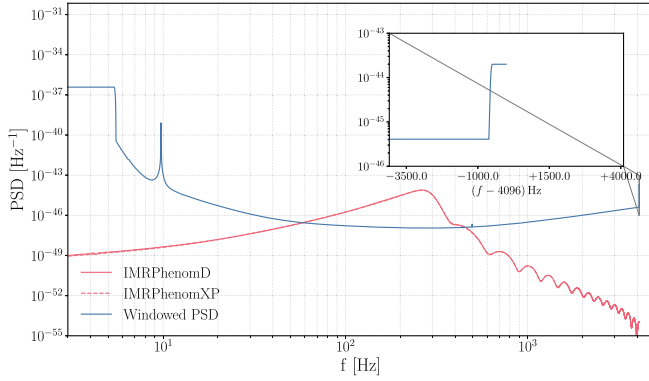


FIG. 10. Illustrative windowed PSD as a function of the frequency (blue curve), generated starting from the PSD in Ref. [101]. The window parameters are $w_l = w_h = 0.5$ Hz, $k = 234$, $\mathcal{I}_l = 100$, and $\mathcal{I}_h = 49$; the inset is a zoom-in of the windowed PSD at high frequencies. The squared absolute value of two different waveform models in the frequency domain (IMRPhenomD and IMRPhenomXP) is shown in red.

To minimize changes in the PSD morphology we set $w_l = w_h = 0.5$ Hz (i.e., we make the window as small as possible, but sufficiently large to guarantee a smooth transition with a reasonable number of points). This can be problematic if, for instance, the ringdown signal in the frequency domain overlaps with one “edge” of the PSD (for instance, massive events have most power at low frequencies).

We then scan the $(k, \mathcal{I}_l, \mathcal{I}_h)$ parameter space. We vary k and $\mathcal{I}_{l,h}$ in the ranges $k \in [0.1, 10^3]$, $\mathcal{I}_{l,h} \in [1, 100]$ and we find that the quantity defined in Eq. (A8) has a minimum

$$p_{\text{TD-FD}}(\underline{W}_{\text{opt}}) = 1.8 \times 10^{-8} \quad (\text{A9})$$

for the window parameters $\underline{W}_{\text{opt}}$ reported in Table IV and used in the main text. We find that varying k has only a mild effect on $p_{\text{TD-FD}}(\underline{W})$. However, the application of the window on a Welch PSD after bandpassing, where its edges are automatically set to lower values, can have a larger effect [90]. We considered different binary parameters and found similar (small) improvements from the application of a window function, so our conclusions are expected to be general.

TABLE IV. Optimal window parameters for the mismatch comparison between the ringdown-only signals of IMRPhenomD and IMRPhenomXP.

w_l [Hz]	w_h [Hz]	k	$\mathcal{I}_l$	$\mathcal{I}_h$
0.5	0.5	234.4	1.0	49.0

APPENDIX B: ADDITIONAL RESULTS

In this appendix, to improve readability, we collect additional results complementing the analysis presented in Sec. III.

In Fig. 11 we show the mismatch distributions for different multipoles (ℓ, m) and different starting times for the London, Cheung, and TEOBPM models. This plot is similar to Fig. 3, but now the origin of the starting time $t_{\text{start}} = 0$ for each (ℓ, m) refers to the peak of the $(2, 2)$ multipole. The plot shows that the TEOBPM mismatches are worse at early times for higher multipoles, because the specific mode is not excited yet and the prepeak emission is not included.

In Figs. 12–16, we plot the Cheung mismatch distributions for $t_{\text{start}} = 15$ M in various two-dimensional projections of the (η, χ_+, χ_-) parameter space. Different rows show the impact of removing specific QNMs. Note that: (1) removing the overtones has a mild impact on the $(2, 2)$ and $(3, 3)$ multipoles; (2) the overtone removal for the $(2, 1)$ multipole actually improves the accuracy, as its fit is inaccurate; (3) removing the quadratic mode has a mild but noticeable effect. This is consistent with Fig. 17, where we plot the ratio between the $(4, 4, 0)^+$ and the $(4, 4, 0)^+ + (2, 2, 0)^+ \times (2, 2, 0)^+$ mismatches at $t_{\text{start}} = 15$ M.

Then, in Fig. 18 we illustrate the accuracy of the interpolation across the parameter space by plotting the distributions of the residuals on t_{start} , denoted as Δt_{start} , for the validation set ($\sim 20\%$ of the SXS simulations). This quantity is defined as

$$\Delta t_{\text{start}} = t_{\text{start}}^{\text{int}} - t_{\text{start}}^{\text{true}}, \quad (\text{B1})$$

where $t_{\text{start}}^{\text{int}}$ and $t_{\text{start}}^{\text{true}}$ are the starting time provided by the interpolant and the true starting time, respectively. We plot the residual distributions for both the London and Cheung models. The histograms show that the residuals peak at 0. For the higher modes, however, some simulations present larger residuals.

In Fig. 19 we show the mismatch distributions of the Cheung model for $t_{\text{start}} = 10, 20$ M considering two different sets of simulation: (1) $\mathcal{I}_C$, i.e., the nonprecessing, quasicircular SXS simulations used for calibrating the Cheung model; (2) $\mathcal{I}_{\text{tot-C}}$, i.e., all the nonprecessing, quasicircular SXS simulations that were not used to calibrate the model. The two distributions have a similar behavior.

Finally, in Fig. 20 we compare the Cheung model with the new, recently released version of the SXS catalog [106]. The distributions are roughly the same, because improvements in the new catalog are mostly related to the number of simulations used to span the parameter space. The main improvement in the new catalog concerns the $(3, 2)$ mode, for which a subset of simulations in the old catalog have $\mathcal{M} = \mathcal{O}(1)$.

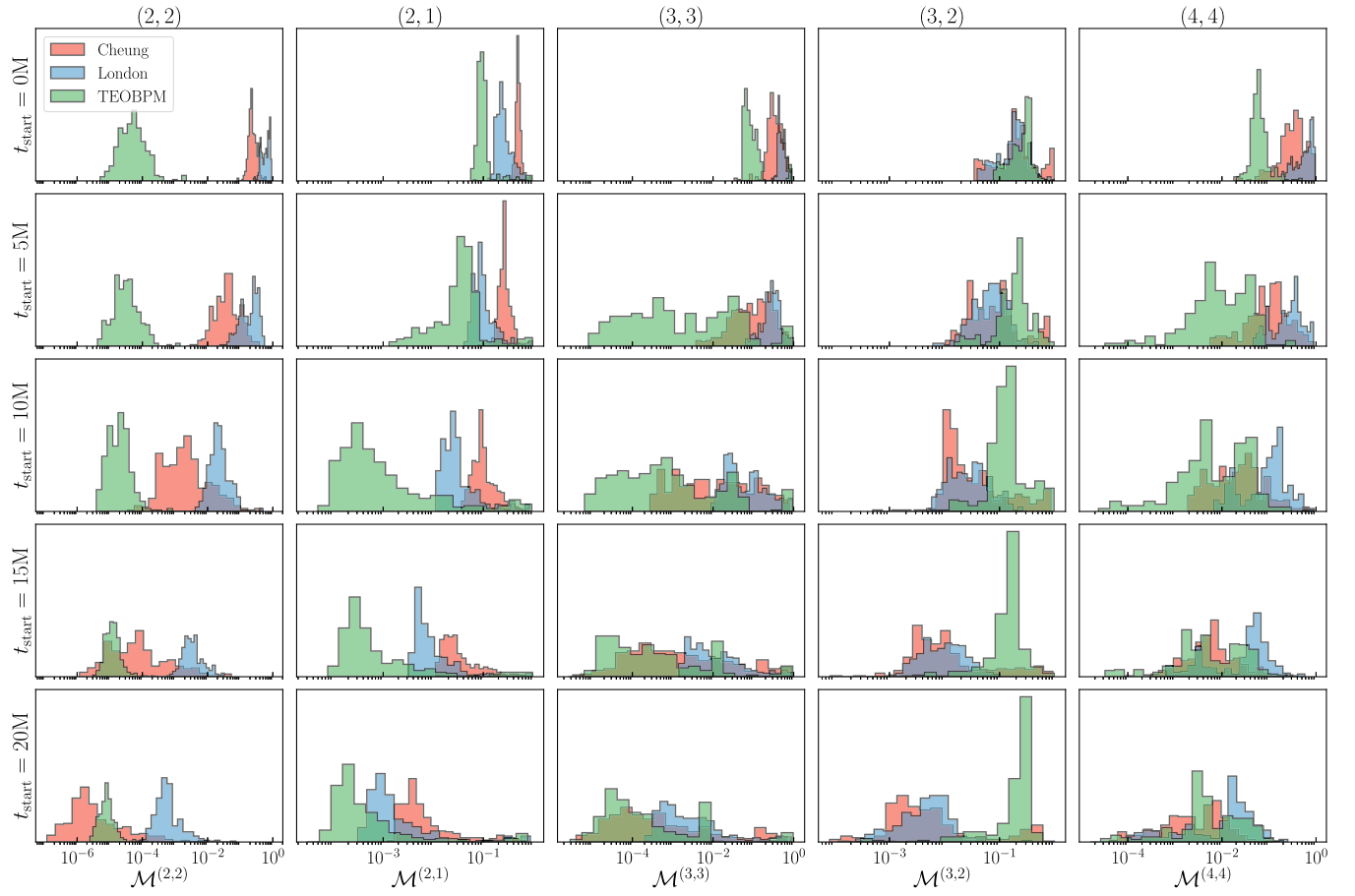


FIG. 11. Same as Fig. 3, but with t_{start} indicating the starting time defined with respect to the $(2, 2)$ peak.

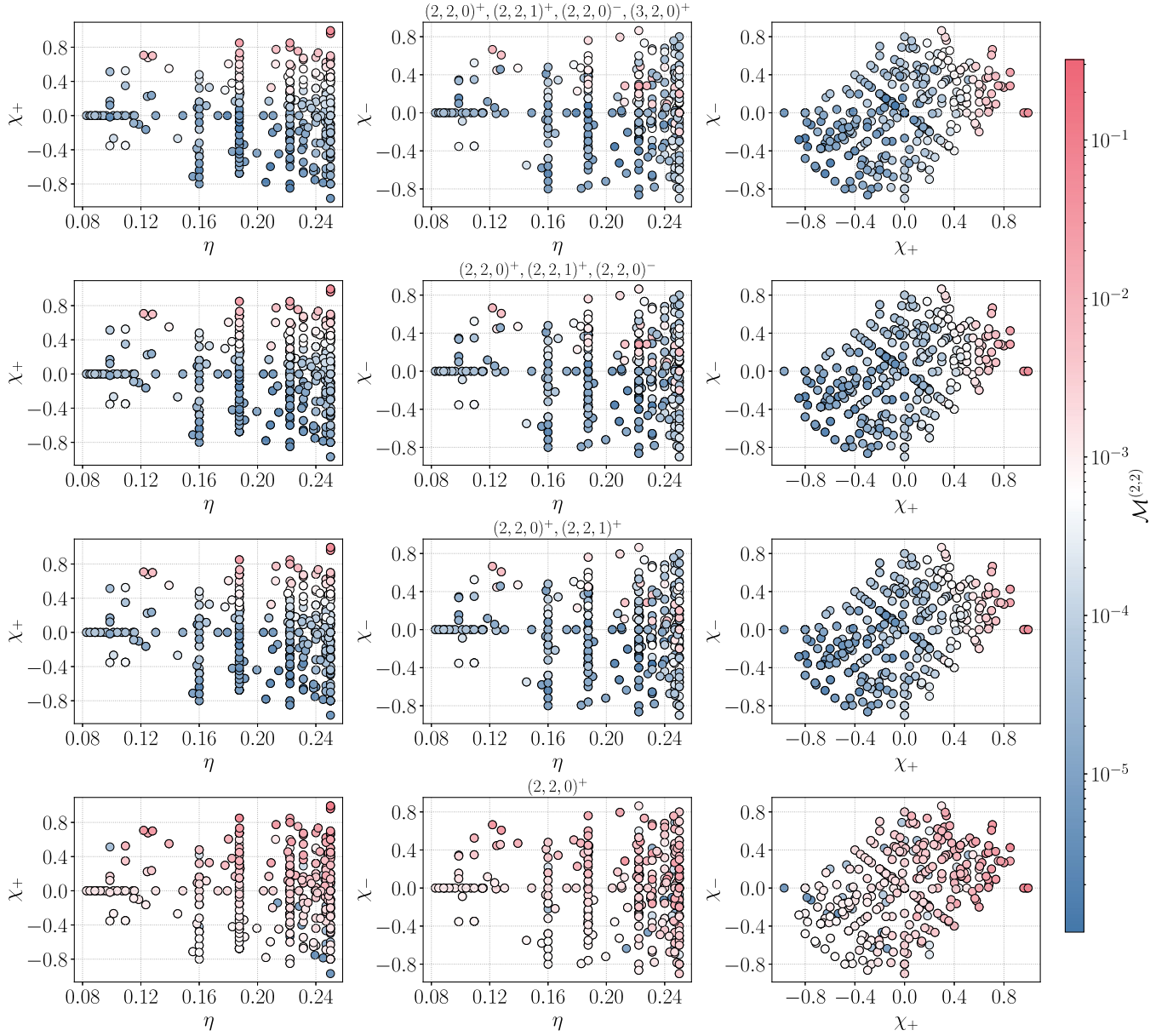


FIG. 12. Scatter plots of the mismatch $\mathcal{M}^{(2,2)}$ of the Cheung model at $t_{\text{start}} = 15$ M as a function of (η, χ_+) , (η, χ_-) , and (χ_+, χ_-) . Each row shows the effect of removing specific QNMs combinations from the model.

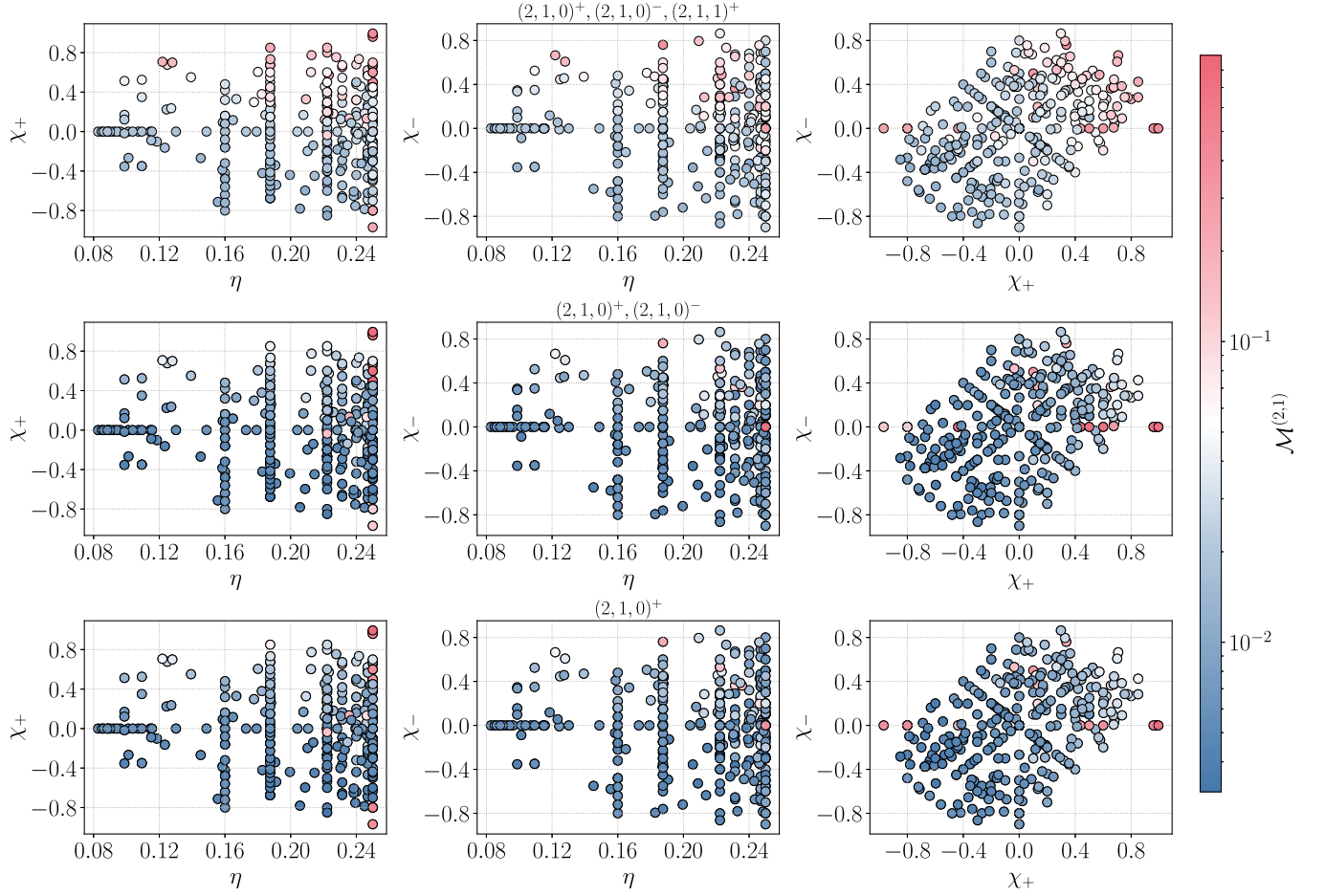


FIG. 13. Scatter plots of the mismatch $\mathcal{M}^{(2,1)}$ of the Cheung model at $t_{\text{start}} = 15$ M as a function of (η, χ_+) , (η, χ_-) , and (χ_+, χ_-) . Each row shows the effect of removing specific QNMs combinations from the model.

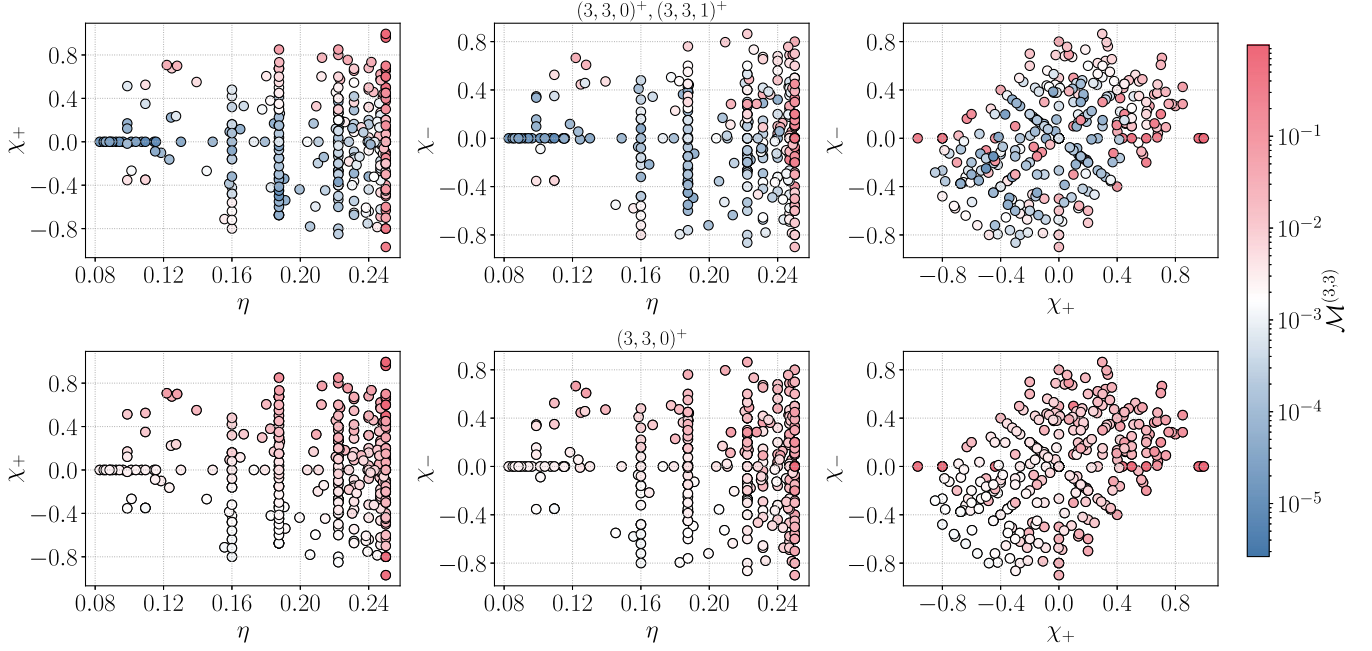


FIG. 14. Scatter plots of the mismatch $\mathcal{M}^{(3,3)}$ of the Cheung model at $t_{\text{start}} = 15$ M as a function of (η, χ_+) , (η, χ_-) , and (χ_+, χ_-) . Each row shows the effect of removing specific QNMs combinations from the model.

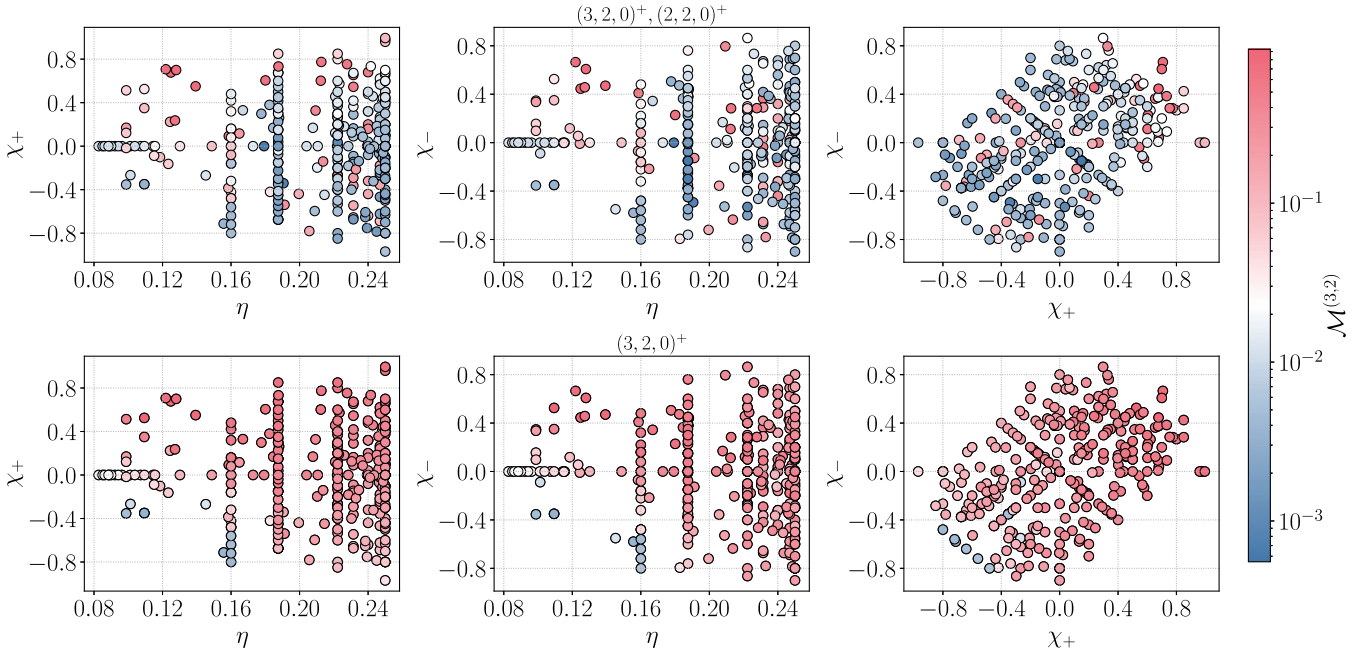


FIG. 15. Scatter plots of the mismatch $\mathcal{M}^{(3,2)}$ of the Cheung model at $t_{\text{start}} = 15$ M as a function of (η, χ_+) , (η, χ_-) , and (χ_+, χ_-) . Each row shows the effect of removing specific QNMs combinations from the model.

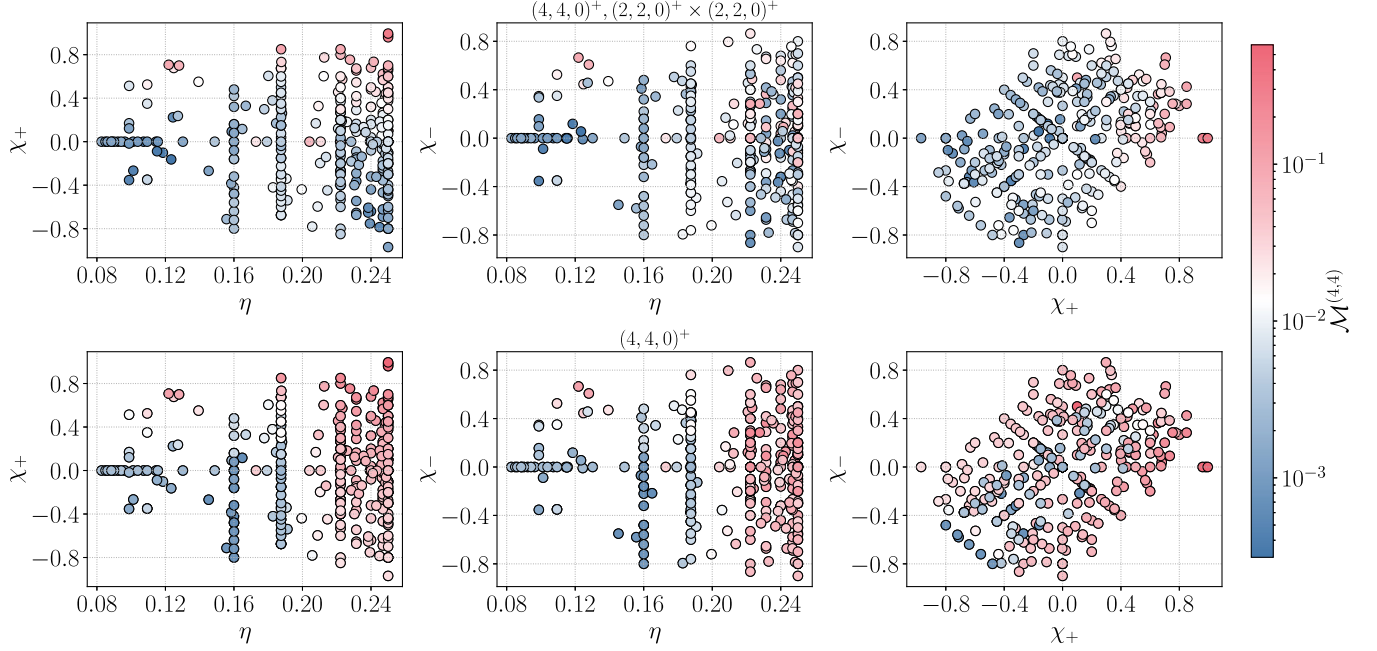


FIG. 16. Scatter plots of the mismatch $\mathcal{M}^{(4,4)}$ of the Cheung model at $t_{\text{start}} = 15$ M as a function of (η, χ_+) , (η, χ_-) , and (χ_+, χ_-) . Each row shows the effect of removing specific QNMs combinations from the model.

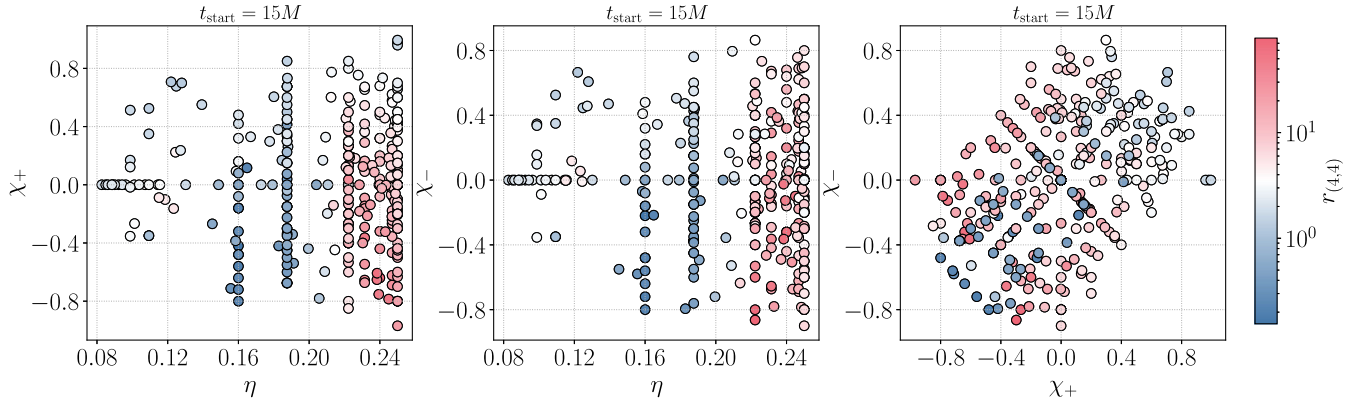


FIG. 17. Ratio $r_{(4,4)}$, as defined in Eq. (22), for $t_{\text{start}} = 15$ M, and for the different parameter space projections.

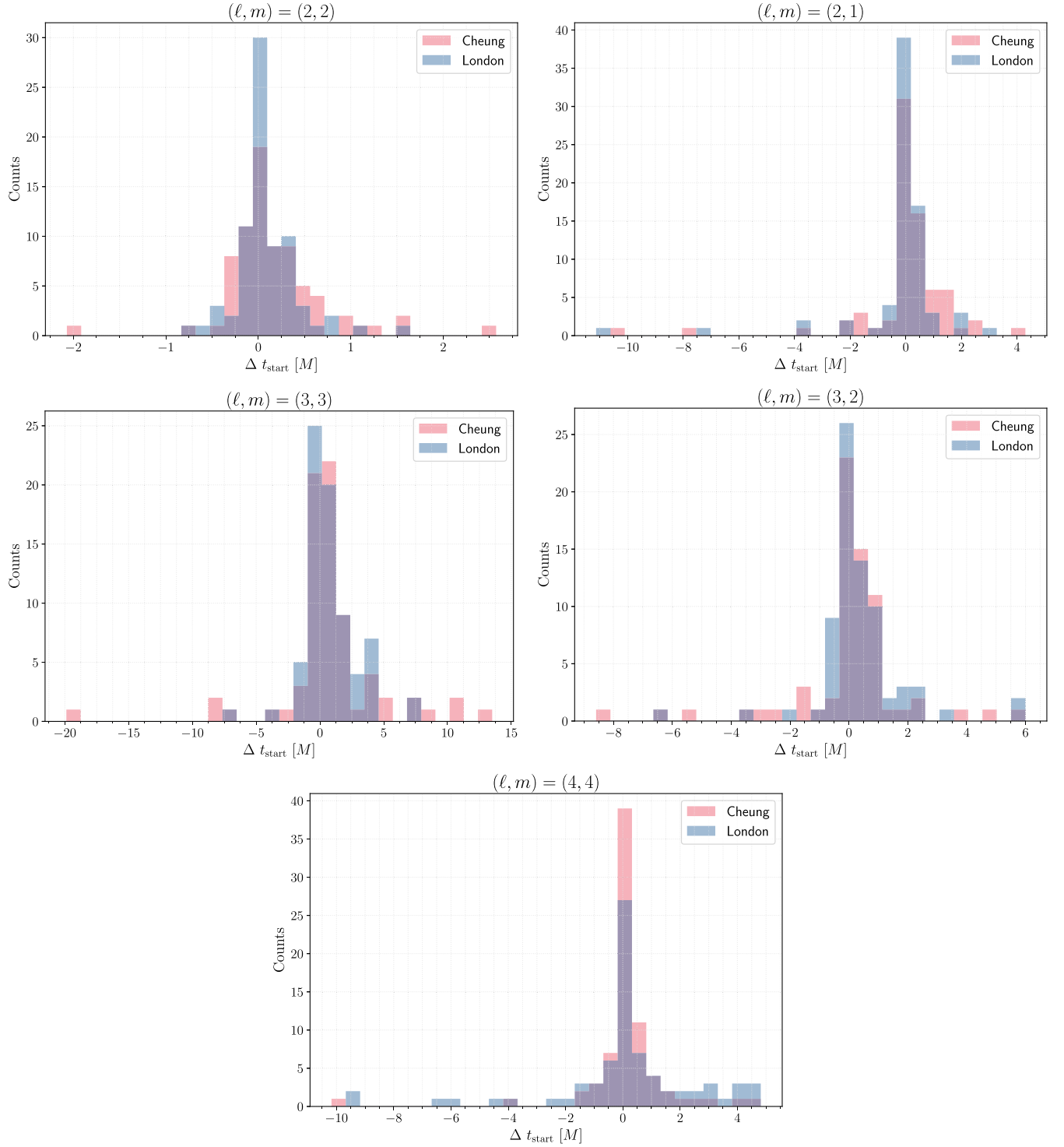


FIG. 18. Distribution of the residual starting time Δt_{start} , as defined in Eq. (B1), for 20% of the simulations of the $\mathcal{I}_{\text{tot}}$ dataset, for both the London (blue) and Cheung (red) models and for different multipoles.

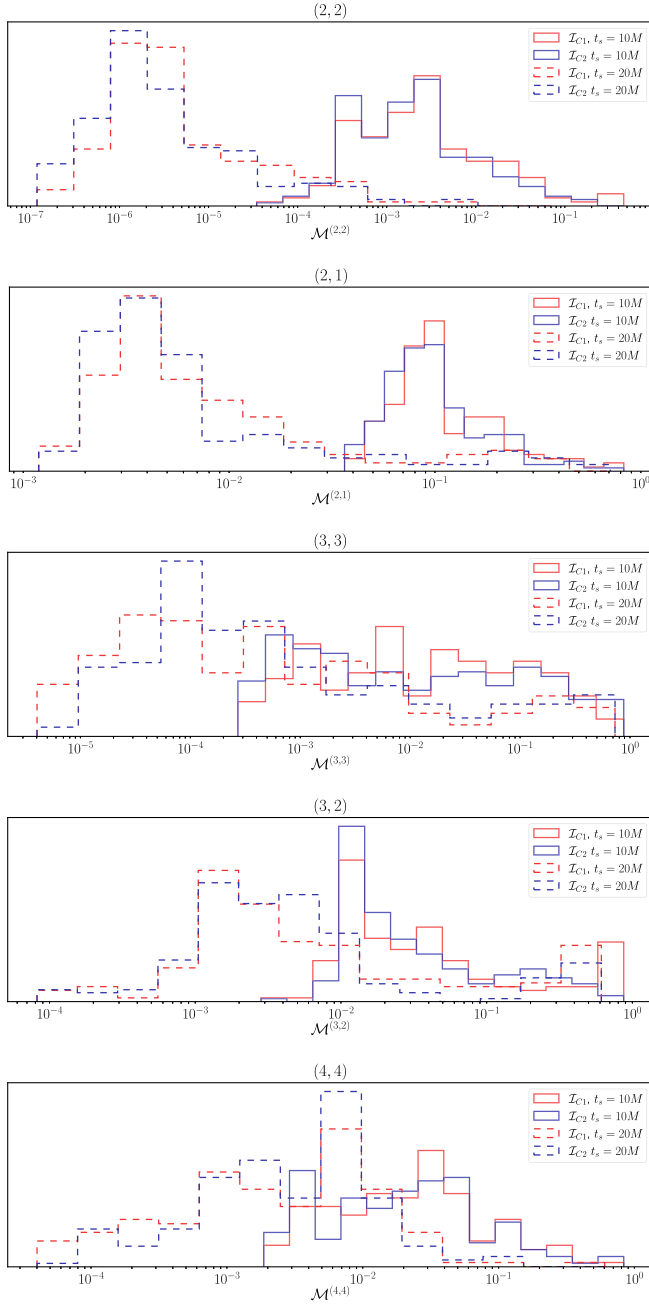


FIG. 19. Effect of using either the ensemble of simulations $\mathcal{I}_C$ (red) or $\mathcal{I}_{\text{tot-C}}$ (blue). Different rows correspond to different $(\ell, |m|)$ multipoles. For each panel, we report the results at $t_{\text{start}} = 10$ M (continuous lines) and 20M (dashed lines).

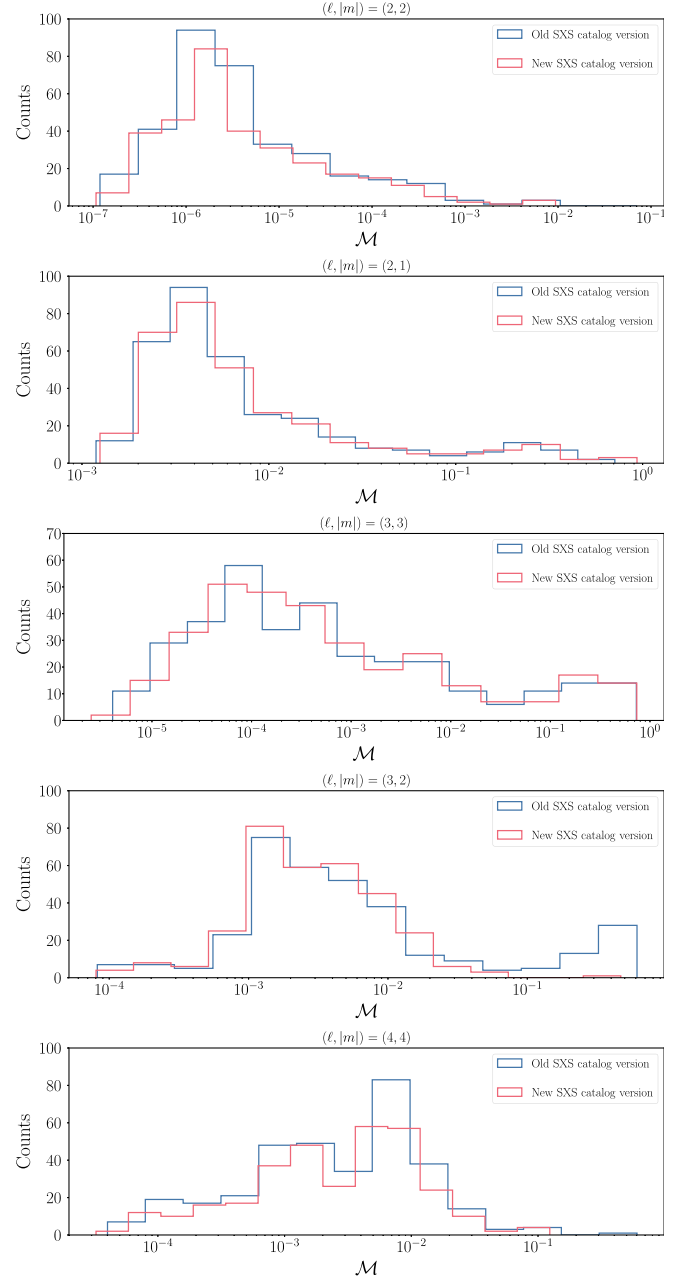


FIG. 20. Mismatch distribution for the Cheung model by comparing the old (blue) and new (red) SXS catalog versions for the different $(\ell, |m|)$ modes. The simulations are plotted for $t_{\text{start}} = 20$ M.

- [1] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **116**, 221101 (2016).
- [2] R. Abbott *et al.* (KAGRA, Virgo, and LIGO Scientific Collaborations), *Phys. Rev. X* **13**, 041039 (2023).
- [3] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), [arXiv:2112.06861](#).
- [4] R. Abbott *et al.* (KAGRA, Virgo, and LIGO Scientific Collaborations), *Phys. Rev. X* **13**, 011048 (2023).
- [5] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), *Astrophys. J.* **949**, 76 (2023).
- [6] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. D* **100**, 104036 (2019).
- [7] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. D* **103**, 122002 (2021).
- [8] E. Berti *et al.*, *Classical Quantum Gravity* **32**, 243001 (2015).
- [9] E. Berti, K. Yagi, H. Yang, and N. Yunes, *Gen. Relativ. Gravit.* **50**, 49 (2018).
- [10] V. Cardoso and P. Pani, *Living Rev. Relativity* **22**, 4 (2019).
- [11] T. Regge and J. A. Wheeler, *Phys. Rev.* **108**, 1063 (1957).
- [12] F. J. Zerilli, *Phys. Rev. D* **2**, 2141 (1970).
- [13] S. A. Teukolsky, *Phys. Rev. Lett.* **29**, 1114 (1972).
- [14] S. A. Teukolsky, *Astrophys. J.* **185**, 635 (1973).
- [15] C. V. Vishveshwara, *Nature (London)* **227**, 936 (1970).
- [16] W. H. Press, *Astrophys. J. Lett.* **170**, L105 (1971).
- [17] E. W. Leaver, *Phys. Rev. D* **34**, 384 (1986).
- [18] W. Israel, *Phys. Rev.* **164**, 1776 (1967).
- [19] S. W. Hawking, *Commun. Math. Phys.* **25**, 152 (1972).
- [20] B. Carter, *Phys. Rev. Lett.* **26**, 331 (1971).
- [21] D. C. Robinson, *Phys. Rev. Lett.* **34**, 905 (1975).
- [22] S. L. Detweiler, *Astrophys. J.* **239**, 292 (1980).
- [23] O. Dreyer, B. J. Kelly, B. Krishnan, L. S. Finn, D. Garrison, and R. Lopez-Aleman, *Classical Quantum Gravity* **21**, 787 (2004).
- [24] E. Berti, V. Cardoso, and C. M. Will, *Phys. Rev. D* **73**, 064030 (2006).
- [25] E. Berti, J. Cardoso, V. Cardoso, and M. Cavaglia, *Phys. Rev. D* **76**, 104044 (2007).
- [26] J. Meidam, M. Agathos, C. Van Den Broeck, J. Veitch, and B. S. Sathyaprakash, *Phys. Rev. D* **90**, 064009 (2014).
- [27] E. Berti, V. Cardoso, G. Carullo *et al.*, [arXiv:2505.23895](#).
- [28] A. G. Abac *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), [arXiv:2509.08099](#).
- [29] A. G. Abac *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), *Phys. Rev. Lett.* **135**, 111403 (2025).
- [30] M. De Amicis, E. Cannizzaro, G. Carullo, and L. Sberna, [arXiv:2506.21668](#).
- [31] I. Kamaretsos, M. Hannam, S. Husa, and B. S. Sathyaprakash, *Phys. Rev. D* **85**, 024018 (2012).
- [32] I. Kamaretsos, M. Hannam, and B. Sathyaprakash, *Phys. Rev. Lett.* **109**, 141102 (2012).
- [33] E. Barausse, J. Bellovary, E. Berti, K. Holley-Bockelmann, B. Farris, B. Sathyaprakash, and A. Sesana, *J. Phys. Conf. Ser.* **610**, 012001 (2015).
- [34] S. J. Miller, M. Isi, K. Chatziioannou, V. Varma, and S. Hourihane, [arXiv:2505.14573](#).
- [35] S. Hadar and B. Kol, *Phys. Rev. D* **84**, 044019 (2011).
- [36] Z. Zhang, E. Berti, and V. Cardoso, *Phys. Rev. D* **88**, 044018 (2013).
- [37] L. Küchler, G. Compère, and A. Pound, [arXiv:2506.02189](#).
- [38] E. Maggio, H. O. Silva, A. Buonanno, and A. Ghosh, *Phys. Rev. D* **108**, 024043 (2023).
- [39] X. J. Forteza, S. Bhagwat, S. Kumar, and P. Pani, *Phys. Rev. Lett.* **130**, 021001 (2023).
- [40] A. Toubiana, L. Pompili, A. Buonanno, J. R. Gair, and M. L. Katz, *Phys. Rev. D* **109**, 104019 (2024).
- [41] A. Toubiana and J. R. Gair, [arXiv:2401.06845](#).
- [42] A. Gupta *et al.*, [arXiv:2405.02197](#).
- [43] S. Yi, F. Iacovelli, S. Marsat, D. Wadekar, and E. Berti, [arXiv:2502.12237](#).
- [44] L. Pompili, E. Maggio, H. O. Silva, and A. Buonanno (2025), [arXiv:2504.10130](#).
- [45] Z. Hu, D. D. Doneva, Z. Wang, V. Paschalidis, G. Bozzola, S. S. Yazadjiev, and L. Shao, [arXiv:2509.07111](#).
- [46] M. Colpi *et al.* (LISA Collaboration), [arXiv:2402.07571](#).
- [47] A. Abac *et al.*, [arXiv:2503.12263](#).
- [48] M. Evans *et al.*, [arXiv:2306.13745](#).
- [49] L. London and E. Fauchon-Jones, *Classical Quantum Gravity* **36**, 235015 (2019).
- [50] X. Jiménez Forteza, S. Bhagwat, P. Pani, and V. Ferrari, *Phys. Rev. D* **102**, 044053 (2020).
- [51] M. H.-Y. Cheung, E. Berti, V. Baibhav, and R. Cotesta, *Phys. Rev. D* **109**, 044069 (2024); **110**, 049902(E) (2024).
- [52] C. Pacilio, S. Bhagwat, F. Nobili, and D. Gerosa, *Phys. Rev. D* **110**, 103037 (2024).
- [53] F. Nobili, S. Bhagwat, C. Pacilio, and D. Gerosa, [arXiv:2504.17021](#).
- [54] L. Magaña Zertuche, L. Gao, E. Finch, and G. B. Cook, *Classical Quantum Gravity* **42**, 197001 (2025).
- [55] K. Mitman *et al.*, *Phys. Rev. D* **112**, 064016 (2025).
- [56] R. J. Gleiser, C. O. Nicasio, R. H. Price, and J. Pullin, *Phys. Rev. Lett.* **77**, 4483 (1996).
- [57] L. London, D. Shoemaker, and J. Healy, *Phys. Rev. D* **90**, 124032 (2014); **94**, 069902(E) (2016).
- [58] M. H.-Y. Cheung *et al.*, *Phys. Rev. Lett.* **130**, 081401 (2023).
- [59] K. Mitman *et al.*, *Phys. Rev. Lett.* **130**, 081402 (2023).
- [60] 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).
- [61] B. Bucciotti, A. Kuntz, F. Serra, and E. Trincherini, *J. High Energy Phys.* **12** (2023) 048.
- [62] D. Perrone, T. Barreira, A. Kehagias, and A. Riotto, *Nucl. Phys. B* **999**, 116432 (2024).
- [63] J. Redondo-Yuste, G. Carullo, J. L. Ripley, E. Berti, and V. Cardoso, *Phys. Rev. D* **109**, L101503 (2024).
- [64] Y. Qiu, X. J. Forteza, and P. Mourier, *Phys. Rev. D* **109**, 064075 (2024).
- [65] S. Ma and H. Yang, *Phys. Rev. D* **109**, 104070 (2024).
- [66] B. Bucciotti, L. Juliano, A. Kuntz, and E. Trincherini, *Phys. Rev. D* **110**, 104048 (2024).
- [67] P. Bourg, R. Panosso Macedo, A. Spiers, B. Leather, B. Bonga, and A. Pound, *Phys. Rev. Lett.* **134**, 061401 (2025).
- [68] H. Zhu *et al.*, *Phys. Rev. D* **109**, 104050 (2024).
- [69] G. Carullo, *J. Cosmol. Astropart. Phys.* **10** (2024) 061.
- [70] R. H. Price, S. Nampalliwar, and G. Khanna, *Phys. Rev. D* **93**, 044060 (2016).

- [71] M. Lagos and L. Hui, *Phys. Rev. D* **107**, 044040 (2023).
- [72] N. Oshita, S. Ma, Y. Chen, and H. Yang, *arXiv:2509.09165*.
- [73] N. Lu, S. Ma, O. J. Piccinni, Y. Chen, and L. Sun, *arXiv:2510.01001*.
- [74] L. Sberna, P. Bosch, W. E. East, S. R. Green, and L. Lehner, *Phys. Rev. D* **105**, 064046 (2022).
- [75] J. Redondo-Yuste, D. Pereñíguez, and V. Cardoso, *Phys. Rev. D* **109**, 044048 (2024).
- [76] T. May, S. Ma, J. L. Ripley, and W. E. East, *Phys. Rev. D* **110**, 084034 (2024).
- [77] L. Capuano, L. Santoni, and E. Barausse, *Phys. Rev. D* **110**, 084081 (2024).
- [78] A. Buonanno and T. Damour, *Phys. Rev. D* **59**, 084006 (1999).
- [79] A. Buonanno and T. Damour, *Phys. Rev. D* **62**, 064015 (2000).
- [80] T. Damour, *Phys. Rev. D* **64**, 124013 (2001).
- [81] T. Damour and A. Nagar, *Phys. Rev. D* **90**, 024054 (2014).
- [82] W. Del Pozzo and A. Nagar, *Phys. Rev. D* **95**, 124034 (2017).
- [83] A. Bohé *et al.*, *Phys. Rev. D* **95**, 044028 (2017).
- [84] R. Cotesta, A. Buonanno, A. Bohé, A. Taracchini, I. Hinder, and S. Ossokine, *Phys. Rev. D* **98**, 084028 (2018).
- [85] A. Nagar, G. Pratten, G. Riemenschneider, and R. Gamba, *Phys. Rev. D* **101**, 024041 (2020).
- [86] A. Nagar, G. Riemenschneider, G. Pratten, P. Retegno, and F. Messina, *Phys. Rev. D* **102**, 024077 (2020).
- [87] 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).
- [88] L. Pompili *et al.*, *Phys. Rev. D* **108**, 124035 (2023).
- [89] M. Isi and W. M. Farr, *arXiv:2107.05609*.
- [90] H. Siegel, M. Isi, and W. M. Farr, *arXiv:2410.02704*.
- [91] L. T. London, *Phys. Rev. D* **102**, 084052 (2020).
- [92] V. Baibhav and E. Berti, *Phys. Rev. D* **99**, 024005 (2019).
- [93] V. Baibhav, E. Berti, and V. Cardoso, *Phys. Rev. D* **101**, 084053 (2020).
- [94] L. Capuano, M. Vaglio, R. S. Chandramouli, C. L. Pitte, A. Kuntz, and E. Barausse, *arXiv:2506.21181*.
- [95] F. Crescimbeni, Interpolating function of ringdown starting time <https://github.com/francesco-crescimbeni/Interpolating-function-of-ringdown-starting-time> (2025), *gitHub repository*.
- [96] W. H. Press and S. A. Teukolsky, *Astrophys. J.* **185**, 649 (1973).
- [97] E. Berti, V. Cardoso, and M. Casals, *Phys. Rev. D* **73**, 024013 (2006); **73**, 109902(E) (2006).
- [98] E. Berti and A. Klein, *Phys. Rev. D* **90**, 064012 (2014).
- [99] E. Berti, V. Cardoso, J. A. Gonzalez, U. Sperhake, M. Hannam, S. Husa, and B. Bruegmann, *Phys. Rev. D* **76**, 064034 (2007).
- [100] L. Blanchet, *Living Rev. Relativity* **17**, 2 (2014).
- [101] LIGO Scientific Collaboration, Instrument science white paper 2018, LIGO Document T1800044, 2018, <https://dcc.ligo.org/LIGO-T1800044/public>.
- [102] <https://numpy.org/doc/stable/reference/routines.fft.html>.
- [103] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Classical Quantum Gravity* **37**, 055002 (2020).
- [104] G. Carullo *et al.*, *Phys. Rev. D* **98**, 104020 (2018).
- [105] M. Boyle *et al.*, *Classical Quantum Gravity* **36**, 195006 (2019).
- [106] M. A. Scheel *et al.*, *arXiv:2505.13378*.
- [107] M. Boyle, K. Mitman, M. Scheel, and L. Stein, The sxs package (2025) [10.5281/ZENODO.15412737](https://doi.org/10.5281/ZENODO.15412737).
- [108] G. Carullo, M. De Amicis, and J. Redondo-Yuste, *bayRing* github.com/GCarullo/bayRing (2023).
- [109] G. Carullo, W. Del Pozzo, and J. Veitch, A time-domain ringdown analysis PYTHON package *pyRing*: git.ligo.org/lscsoft/pyring (2023).
- [110] L. C. Stein, *J. Open Source Software* **4**, 1683 (2019).
- [111] K. Jani, J. Healy, J. A. Clark, L. London, P. Laguna, and D. Shoemaker, *Classical Quantum Gravity* **33**, 204001 (2016).
- [112] V. Gennari, G. Carullo, and W. Del Pozzo, *Eur. Phys. J. C* **84**, 233 (2024).
- [113] E. Finch and C. J. Moore, *Phys. Rev. D* **103**, 084048 (2021).
- [114] H. Zhu *et al.*, *Phys. Rev. D* **111**, 064052 (2025).
- [115] E. E. Flanagan and S. A. Hughes, *Phys. Rev. D* **57**, 4566 (1998).
- [116] L. Lindblom, B. J. Owen, and D. A. Brown, *Phys. Rev. D* **78**, 124020 (2008).
- [117] R. S. Chandramouli, K. Prokup, E. Berti, and N. Yunes, *Phys. Rev. D* **111**, 044026 (2025).
- [118] J. E. Thompson, C. Hoy, E. Fauchon-Jones, and M. Hannam, *Phys. Rev. D* **112**, 064011 (2025).
- [119] S. Bhagwat, C. Pacilio, P. Pani, and M. Mapelli, *Phys. Rev. D* **108**, 043019 (2023).
- [120] E. Berti, A. Sesana, E. Barausse, V. Cardoso, and K. Belczynski, *Phys. Rev. Lett.* **117**, 101102 (2016).
- [121] S. Bhagwat, C. Pacilio, E. Barausse, and P. Pani, *Phys. Rev. D* **105**, 124063 (2022).
- [122] K. Mitman *et al.*, *Phys. Rev. D* **103**, 024031 (2021).
- [123] J. Blackman, S. E. Field, C. R. Galley, B. Szilagy, M. A. Scheel, and M. Tiglio, *Phys. Rev. Lett.* **115**, 121102 (2015).
- [124] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jiménez Forteza, and A. Bohé, *Phys. Rev. D* **93**, 044007 (2016).
- [125] G. Pratten *et al.*, *Phys. Rev. D* **103**, 104056 (2021).
- [126] 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).
- [127] PyCBC, Calculating the match between waveforms (2025), <https://pycbc.org/pycbc/latest/html/waveform.html#calculating-the-match-between-waveforms> (accessed: 2025-02-09).
- [128] M. Pürrer and C.-J. Haster, *Phys. Rev. Res.* **2**, 023151 (2020).
- [129] <https://www.arch.jhu.edu/>.