

Quantifying systematic biases in black hole spectroscopy

Sebastian H. Völkel^{*} and Arnab Dhani[†]

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



(Received 6 August 2025; accepted 1 October 2025; published 29 October 2025)

How long after the merger of two black holes can one rely on linear perturbation theory, and how many quasinormal modes are in the ringdown? Such questions suggest that black hole spectroscopy suffers from systematic uncertainties that potentially spoil ringdown analyses, both from high-accuracy simulations and in data from gravitational wave detectors. In this work, we demonstrate that linear-signal analysis is a powerful tool for quantifying biases, allowing for detailed explorations that are computationally too expensive for traditional Bayesian injection and recovery approaches. We quantify the validity of the Fisher information matrix and bias formula by comparing it to robust but slow Bayesian sampling. Working with flat noise in the time domain, statistical errors and systematic biases can mostly be detected analytically. Due to its efficiency, we provide detailed parameter space analyses for potentially unmodeled small contributions from overtones, quadratic modes, and tails. We find linear signal analysis well suited for predicting biases in simple ringdown models at intermediate signal-to-noise ratios (SNRs) of order 50 when unmodeled effects are small. It is also valuable in explaining ongoing issues in extracting quasinormal modes from high-precision simulations, as one can understand them as high-SNR signals. Therefore, this approach offers promising prospects for improving ringdown models by efficiently identifying and incorporating systematic uncertainties, ultimately enhancing the accuracy and robustness of black hole spectroscopy.

DOI: [10.1103/g6sz-dw28](https://doi.org/10.1103/g6sz-dw28)

I. INTRODUCTION

The intriguing concept of studying the properties of a black hole through its gravitational wave spectrum is known as *black hole spectroscopy* [1–3]. It is a fundamental prediction of black hole perturbation theory [4–6]. A perturbed black hole, such as during a binary black hole merger, emits characteristic gravitational waves that can be well approximated by a sum of damped sinusoid modes, known as quasinormal modes (QNMs). For a comprehensive discussion of this topic, we refer to the literature, including Refs. [7–11]. According to general relativity (GR) and within the framework of no-hair theorems, the spectrum of QNMs is uniquely determined by the mass and spin of the resulting black hole, which is described by the Kerr metric [12–15]. As a result, measurements of multiple

QNMs can be used to test the assumptions of the Kerr hypothesis and even the validity of GR itself.

Despite its elegance and similarities with other spectroscopy methods in physics, black hole spectroscopy faces several complex challenges. One significant question is how long after the merger of two black holes one can rely on linear perturbation theory. Even after sufficient time has passed, quantifying the number of QNMs in a given signal remains challenging. This is in part because QNMs do not form a complete basis [16,17]. On the one hand, QNMs are only part of the solution of the perturbed Einstein's equations, which also include the transient early part of the signal, known as the prompt response, and power-law tails which dominate the signal at very late times [18–24]. On the other hand, the set of QNMs is overcomplete in the sense that any QNM can be expressed as a linear combination of the other QNMs, raising questions about their physical significance [25]. Standard black hole spectroscopy approaches, which analyze the intermediate part of the ringdown as a sum of damped sinusoids, encounter what is known as the short blanket problem. To minimize the effects of prompt responses and nonlinear contributions, the analysis must begin late enough but conclude before the signal's tails take over [19–21,23,26,27]. However, this reduced signal length requires high signal-to-noise ratio (SNR) signals. Unfortunately, increasing the

^{*}Contact author: sebastian.voelkel@aei.mpg.de

[†]Contact author: arnab.dhani@aei.mpg.de

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. Open access publication funded by the Max Planck Society.

SNR amplifies both the prompt response and nonlinear effects during the intermediate times. Pioneering work for the nonlinear case can be found in Refs. [28–30]. In recent years, several studies have assessed how effectively QNMs can be inferred from high-accuracy linear, nonlinear simulations, and actual data, with varying conclusions regarding the questions previously mentioned [31–46]. While earlier studies used a heuristic set of overtones and retrograde modes to fit the complex amplitudes to the data, more recent works use refined fitting approaches to identify multiple QNMs, including overtones, retrograde modes, and quadratic modes, demonstrating the importance of careful extraction techniques from numerical relativity (NR) simulations [40,42,43,46]. However, once the QNMs are identified, the complex amplitudes are extracted by fixing the identified QNM frequencies, similar to earlier studies. A ringdown surrogate model based on Gaussian processes has been introduced in Ref. [47].

One central aspect of ringdown analyses is controlling the number of free parameters. In a theory-agnostic approach, the goal is to use minimal information and treat all complex frequencies and amplitudes as independent parameters. This method enables a flexible extraction of signal features, thereby capturing possible modifications from GR. In contrast, a theory-specific approach requires choosing the underlying theory, such as GR, and modeling all possible QNMs as a function of a few parameters, like the mass and spin. Additionally, QNM excitation coefficients could be obtained from numerical relativity simulations or perturbative calculations to further reduce the number of free parameters [48–52]. While a smaller number of independent parameters reduces the complexity when considering a large number of QNMs, it is challenging to extend this to rotating black holes in beyond-GR theories.

Although each approach has its own advantages and disadvantages, both are prone to systematic errors in the extracted black hole parameters or agnostic QNMs. This is because, while almost all QNMs get excited in a binary merger, except for the ones excluded by the symmetries of the initial binary, it is challenging to calculate their amplitudes, except for simple cases [53,54]. Therefore, in both approaches, the number of QNMs needed to sufficiently describe a realistic ringdown signal is not known *a priori* [31,36,37,42,55–59]. Relying on metrics like mismatches does not address the risk of overfitting. Moreover, in both approaches, little attention has been given to systematic errors due to the presence of tails. Tails are usually not modeled in parameter estimation studies because they are considered a weak late-time effect that can be avoided by stopping the analysis early enough. However, recent studies suggest that tails might be more relevant than anticipated [60–66]. Other sources of systematics are the unmodeled prompt response, mode mixing, spin-precession from the progenitor system

[67,68], and detector calibration accuracy [69]. Nonlinear QNMs, so-called quadratic QNMs (which are sourced at quadratic order from linear perturbations), have also received significant attention recently and could be an important part of ringdown analyses, see Refs. [58,70,71] for more precise details, along with other nonlinear effects [72]. They might also be detectable by future detectors [73,74]. The importance of overtones in extracting the correct fundamental mode of collapsing neutron stars and binary-neutron star mergers forming a final black hole surrounded by leftover matter has recently been studied in Ref. [75].

Quantifying the impact of unmodeled contributions to realistic signals in terms of systematic errors on the underlying parameters is a very complex and vast problem. Using traditional Bayesian methods like Markov-chain-Monte-Carlo (MCMC) sampling for injection and recovery studies to tackle this question is computationally expensive and thus limited. Simulation-based inference methods are promising alternatives and have recently been applied to the ringdown [76,77].

In this work, we show that linear signal approximation (LSA) methods [78–82], based on the Fisher information matrix and the bias formula, are powerful tools for quantifying systematic errors from unmodeled effects to a simple ringdown model. For high-SNR signals from future detectors or high-accuracy simulations, LSA allows one to predict the biases in a model’s parameters as long as the signal being analyzed has small deviations from the model. Therefore, the method is well suited for detailed explorations of the vast zoo of possibly overlooked effects and for quantifying at which SNR systematic errors due to a specific unmodeled effect become important. This further allows one to improve ringdown models and shows great potential to incorporate systematic uncertainties on top of the statistical ones. As detectors become more sensitive, black hole spectroscopy needs to put more attention to dealing with such systematic uncertainties to prevent false violations from GR [83,84], or properly quantify new physics that could be waiting out there.

In Sec. II, we introduce the ringdown models and the statistical framework and provide a comprehensive application of our framework in Sec. III. We compare results obtained by LSA with those by Bayesian analysis for selected cases, and then explore a variety of effects, including unmodeled overtones, quadratic modes, and power-law tails. Our findings are discussed in Sec. IV, and our conclusions can be found in Sec. V. Complementary results can be found in the Appendix. In this work, we work in units with $G = c = 1$.

II. METHODS

In the following, we review ringdown models in Sec. II A, describe the Bayesian analysis framework and

statistical methods Sec. II B, and outline LSA methods in Sec. II C.

A. Modeling the ringdown

We model the ringdown signal as a sum of damped sinusoids, similar to the current ringdown analysis by the LVK Collaboration. The GW strain is then given by

$$h(t; \theta) = \sum_{\ell mn}^N A_{\ell mn} \exp[i(\omega_{\ell mn}^{\text{re}} t + \phi_{\ell mn}) - \omega_{\ell mn}^{\text{im}} t], \quad (1)$$

with time t (chosen such that the signal starts at $t = t_0 = 0$), QNMs $\omega_{\ell mn} = \omega_{\ell mn}^{\text{re}} + i\omega_{\ell mn}^{\text{im}}$, amplitudes $A_{\ell mn}$, and phases $\phi_{\ell mn}$. Note that t_0 should not be confused with the peak of a full inspiral-merger-ringdown waveform, as our analysis only employs a ringdown model. Instead, it sets a reference point with respect to amplitude ratios and phase shifts.

Throughout this work, we will distinguish between two different ways of employing Eq. (1): theory-agnostic and theory-specific. For the theory-agnostic case, all QNMs $\omega_{\ell mn}$, amplitudes $A_{\ell mn}$, and phases $\phi_{\ell mn}$ are treated as independent free parameters. Therefore, a ringdown waveform with N modes corresponds to $4N$ free parameters. On the other hand, for the theory-specific case, the QNMs $\omega_{\ell mn} = \omega_{\ell mn}(M, a)$ are determined in terms of the stationary black hole mass M and spin a , while the amplitudes $A_{\ell mn}$ and phases $\phi_{\ell mn}$ are still free parameters. In this case, a ringdown waveform with N modes has $2N + 2$ free parameters. The numerical data for the QNMs is taken from Refs. [3,9], which has been implemented in the parametrized QNM framework in Ref. [85] and is available on GITHUB [86].

To study the impact of unmodeled contributions to the ringdown signal, the signal, $h_{\text{signal}}(t; \theta_{\text{signal}})$, is given by the sum of the model and the unmodeled effect, $\Delta h(t; \theta_{\text{effect}})$, of interest. Therefore, it is described by

$$h_{\text{signal}}(t; \theta_{\text{signal}}) = h(t; \theta) + \Delta h(t; \theta_{\text{effect}}), \quad (2)$$

where $\theta_{\text{signal}} = [\theta, \theta_{\text{effect}}]$. In Sec. III, we study the effect of an overtone, a quadratic mode, and a power-law tail as the unmodeled feature. Whether its presence introduces biases at intermediate times, despite of not being a visible feature, is one of the question we will explore. The power-law tail is prescribed as

$$h_{\text{tail}}(t) = \frac{A}{(t - t_{\text{pole}})^7}, \quad (3)$$

where the exponent 7 is used to describe the $\ell = 2$ Price tail [19–21,23,26,27], and A and t_{pole} will be used as free parameters to control the contribution of the tail to the signal. These parameters, as well as the QNM amplitudes

and phases, depend on the initial data, i.e., the properties of the binary. However, in this work, we restrict ourselves to canonical values for these parameters and do not identify them with any binary configuration.

B. Bayesian framework

Below, we describe the standard time-domain statistical framework used in GW ringdown data analysis [87,88]. The log-likelihood describing stationary Gaussian noise is given by

$$\ln p(d|\theta) = -\frac{1}{2} \langle d - h(\theta) | d - h(\theta) \rangle, \quad (4)$$

where $h(\theta)$ is the ringdown model and d is the observed signal. We assume a flat noise power-spectral density (PSD), setting S_h to be a constant, which allows for simple analytic calculations. While this does not accurately model the frequency-dependent noise of a realistic GW detector, it allows us to isolate the effect of just the unmodeled physics. The inner product used in Eq. (4) then reduces to

$$\langle h_1 | h_2 \rangle = \frac{4}{S_h} \Re \left[\int_{t_{\text{start}}}^{t_{\text{end}}} h_1(t) \times h_2^*(t) dt \right], \quad (5)$$

which allows us to easily vary the start time t_{start} and the end time t_{end} . Using a realistic power-spectral density would not directly impact the validity of LSA, but rather assign different weights and might favor a frequency-domain analysis. Moreover, using Eq. (5) is common when extracting QNM information from numerical relativity simulations, e.g., to calibrate excitation coefficients by minimum chi-square estimation. In Sec. III, we will investigate different signal-to-noise ratios (SNRs) $\rho^2 \equiv \langle h|h \rangle$ by simply adjusting the value of S_h , instead of changing the signal itself. This is equivalent to adjusting the distance between the detector and the source.

In order to infer the parameters of a model, one can use Bayes' theorem

$$p(\theta|d) = \frac{p(d|\theta)p(\theta)}{p(d)}, \quad (6)$$

which relates the probability distribution of the parameters given the data $p(\theta|d)$, with the likelihood of the data given the parameters $p(d|\theta)$. The posterior further depends on our prior knowledge of the parameters $p(\theta)$, before analyzing the data, and the probability of the data itself $p(d)$, the evidence, which is just a normalization term. In this work, we assume uniform prior distributions for all parameters of our model.

A direct computation of the posterior distribution from Bayes' theorem is only possible in special cases. However, there is a wide range of statistical methods that allow one to draw direct samples from the posterior by providing

the likelihood and prior, independent from the evidence. The posterior distribution can then be approximated by obtaining many samples. Among these methods are MCMC techniques. In this work, we sample the likelihood Eq. (4) using the EMCEE sampler [89], which is a widely used MCMC code. Due to our simplifying assumptions of Gaussian noise in the time domain and using analytic functions (damped sinusoids), the computational cost of sampling is moderate. We use it to benchmark the results obtained from LSA, which is introduced next.

C. Linear signal analysis

Although MCMC methods are capable of providing posterior distributions, they are computationally expensive and thus quickly become a bottleneck for vast parameter space explorations. A well-established approach in gravitational wave data analysis, circumventing posterior sampling, is LSA. It is applied in the high-SNR limit, when an expansion of the likelihood around its maximum can be well approximated in terms of a linear expansion of the waveform model with respect to its parameters.

We start with the prescription of obtaining statistical errors and correlations in LSA. To estimate the statistical errors of the ringdown parameters θ , we utilize the Fisher information matrix. It represents a quadratic approximation of the log-likelihood around its maximum, depending on the parameters of interest, and is valid for large SNR. Its inverse is the covariance matrix $C = \Gamma^{-1}$ whose entries σ_{ij} contain the statistical errors $\Delta\theta_i \equiv \sqrt{\sigma_{ii}}$ and correlations. The elements of the Fisher information matrix are given by

$$\Gamma_{ij} = \langle \partial_i h | \partial_j h \rangle, \quad (7)$$

where the partial derivatives are with respect to the components of θ . Due to the simple structure of the ringdown model Eq. (1) and inner product, all of its elements can be computed analytically, both for the theory-specific and theory-agnostic case.

A second application of LSA is the prediction of systematic errors for θ , also known as biases $\delta\theta$. Biases are expected when the model used to analyze a signal is incomplete or inaccurate. Assuming that the difference between signal h_{signal} and model h is small but finite, one can use the so-called bias formula

$$\delta\theta_i = (\Gamma_{ij})^{-1} \langle \partial_j h | h_{\text{signal}} - h \rangle, \quad (8)$$

to predict the “shifts” in the parameters that one would obtain when performing inference (for large SNR) [81,82]. Note that when applied to bias computations across different waveform families, e.g., for precessing binary black hole systems, an important alignment procedure has been pointed out in Ref. [90] and adopted in Ref. [91]. Since our case allows for full analytic results for the elements of the Fisher information matrix Γ_{ij} , we only need to invert it

numerically to obtain its inverse. Note that the bias formula does not require one to have an analytic representation of the signal. Because we also want to consider signals for which there is no analytic expression and be flexible for future applications, we compute the inner product for the bias formula numerically. Note that typical computation of the Fisher information matrix and bias formula is very fast for the ringdown model, i.e., many orders of magnitude faster than MCMC sampling.

Finally, to quantify the relevance of biases in a compact form, we introduce the bias ratio $\delta\theta_i/\Delta\theta_i$. This quantity is smaller than one when parameters are dominated by statistical errors, and larger than one if dominated by systematic errors.

III. APPLICATIONS

In this section, we apply our framework to different types of scenarios that are interesting for black hole spectroscopy. To set the stage for comprehensive parameter space explorations for which MCMC analyses become too expensive, we first quantify the agreement between LSA and MCMC analysis for selected cases in Sec. III A. We then proceed with different types of detailed parameter space explorations that are only feasible in a reasonable time with LSA in Sec. III B. The latter includes how biases from unmodeled QNMs depend on their phase difference, the impact of varying the start time of the ringdown analyses, and how the large zoo of possibly unmodeled QNMs introduces biases when being modeled one at a time, see Refs. [92,93] for recent examples using LSA for LISA-related applications.

A. LSA versus Bayesian analysis

We consider specific examples to demonstrate that the predictions from LSA can be accurate when compared to the full Bayesian MCMC analysis. We consider a signal composed of a fundamental mode QNM along with either an overtone, a quadratic mode, or a power-law tail. The complex frequencies correspond to the prograde quadrupolar mode of a BH with unit mass and dimensionless spin of 0.7. The amplitude of the additional mode is fixed at 10% of the fundamental mode at the beginning of the signal. Note that the beginning of the signal is not intended to represent the actual peak amplitude of a binary merger ringdown. In this case, the amplitudes would typically be much larger, and it would be important to incorporate additional linear and quadratic QNMs simultaneously. Instead, our example rather refers to a situation where the analysis is started when one has either subtracted other modes (but not completely) or assumes they are negligible. Our values are chosen as qualitative examples and do not correspond to any particular binary configuration. We summarize the different contributions in Fig. 1, which shows the fundamental mode (ω_{220}), overtone (ω_{221}),

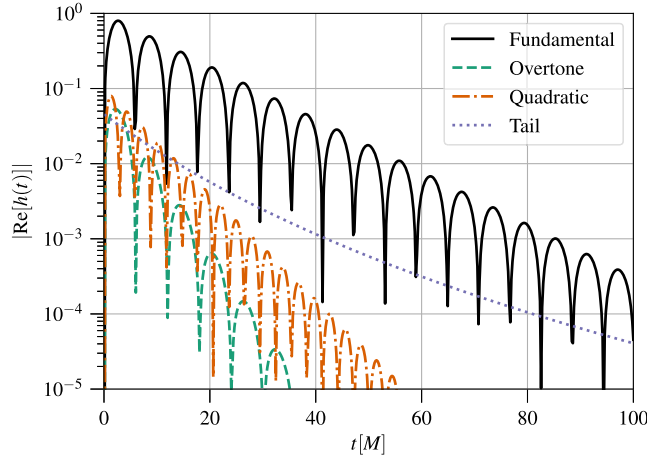


FIG. 1. *Ringdown signal and model*: We show the fundamental mode ω_{220} along with the contribution of an unmodeled overtone ω_{221} , unmodeled quadratic QNM $\omega_{220 \times 220}$, and unmodeled power-law tail. Numerical values of the parameters are reported in Table I.

quadratic QNM ($\omega_{220 \times 220}$), and the power-law tail. Note that the quadratic mode oscillates and decays twice as fast as the fundamental mode but is still longer lived than the overtone. All numerical values are reported in Table I.

We inject a signal with $\text{SNR} = 50$ containing the fundamental model and an additional effect, which could be either an overtone, a quadratic QNM, or a power-law tail. We analyze this signal using a single QNM model corresponding to the fundamental mode. We report the comparison between the Bayesian analysis and the Fisher estimates in Fig. 2, with the top panel displaying the theory-specific case and the theory-agnostic case in the bottom panel. The crosshairs depict the true values of

TABLE I. *Parameters of the ringdown signals*: Values of each case are in terms of the theory-specific parametrization. Theory-agnostic real and imaginary parts of the QNMs are completely determined by mass and spin, as shown in the table. Their numerical values are $\omega_{220} = 0.53260024 + i0.08079287$, $\omega_{221} = 0.52116077 + i0.24423832$, $\omega_{220 \times 220} = 1.06520049 + i0.16158575$.

Parameters	Case 1	Case 2	Case 3
M	1	1	1
a	0.7	0.7	0.7
A_{220}	1.0	1.0	1.0
ϕ_{220}	$\pi/2$	$\pi/2$	$\pi/2$
A_{221}	0.1	...	...
ϕ_{221}	$\pi/2$	...	...
$A_{220 \times 220}$	...	0.1	...
$\phi_{220 \times 220}$	...	$\pi/2$	...
A_{tail}	...	...	1×10^{11}
$t_{\text{pole}}[M]$	...	...	-58

the parameters, while the different cases are shown in different colors.

As evident from Fig. 2, the overall agreement between the full Bayesian posteriors and LSA predictions is excellent. In the theory-specific case, we observe that mass M and spin a posteriors follow a similar correlation in all the cases. In the theory-agnostic case, the posteriors of the real and imaginary part of ω do not follow such a correlation. Instead, the real part is less sensitive to the unmodeled effects, while the imaginary part is similarly biased as mass and spin in the theory-specific case. Although the biases in

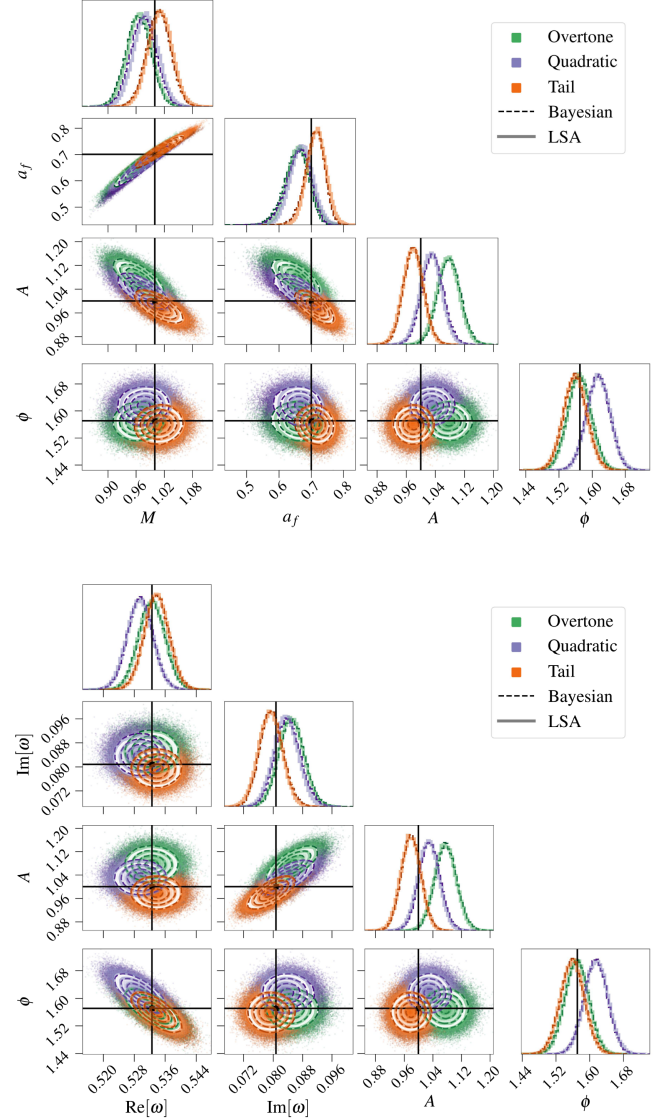


FIG. 2. *Comparison of LSA versus Bayesian analysis*: Here we compare the theory-specific (top panel) and theory-agnostic (bottom panel) biases for three unmodeled effects (different colors) described in the main text. In each panel, we compare posteriors obtained using the Bayesian (thin dashed lines) and LSA (thick solid lines) methods, respectively. The values of the injected parameters are shown for reference (black crosshair). In this example, the signal has $\text{SNR} = 50$.

the examples presented here are not dramatic, they would already be problematic for tests of GR. Finally, we note that the biases in amplitude and phase are very similar in the theory-specific and theory-agnostic cases. In fact, the largest bias appears for the amplitude in the case of an unmodeled overtone. This could potentially be interesting in the context of amplitude-phase consistency tests of GR as proposed in Ref. [94]. We provide additional results for the same unmodeled effects, but $\text{SNR} = 20$ and $\text{SNR} = 100$ in the Appendix in Fig. 6. We find that the accuracy of LSA degrades at $\text{SNR} = 20$, with a small but notable better performance in the theory agnostic case. At $\text{SNR} = 100$, the systematic errors typically exceed the statistical ones, which would be highly problematic for tests of GR.

In Fig. 7 of the Appendix, we provide a comprehensive comparison of bias ratios as a function of selected SNRs starting from 20 to 300. They correspond to the same effects as studied previously, but condense the median and 68% of the Bayesian results for the 1D marginalized posteriors of the parameters of the theory-specific and theory-agnostic model, respectively. In summary, as expected, the agreement between LSA and Bayesian results improves with increasing SNR. Remarkably, the quantitative details, i.e., which of the parameters is better approximated by LSA or experiences the largest biases, is very case dependent.

B. Parameter space explorations

1. Exploring QNM phase differences

Having established the validity of the LSA, we explore the impact of the phase difference between the fundamental mode and an unmodeled QNM on systematic biases. In Fig. 3, we show the bias ratios of the theory-specific and theory-agnostic case when using only the fundamental mode ω_{220} as model, but allowing either an unmodeled overtone ω_{221} or an unmodeled quadratic mode $\omega_{220 \times 220}$ with an initial amplitude of 10% of the fundamental, respectively.

The results in Fig. 3 show nontrivial features that might easily remain unnoticed in a less detailed analysis. We find that the phase difference has a huge impact on bias ratios, which vary by many orders of magnitude. Moreover, it also significantly changes which of the parameters is most or least biased. For instance, let us first compare results for the phase with those in Fig. 2. We observe that the phase bias is almost vanishing for the unmodeled overtone ω_{221} (shown in the top row) only due to the specific choice of injecting it with zero phase difference. In contrast, the phase bias from the unmodeled overtone ω_{221} would be negligible around other phase differences. Moreover, we find that the bias ratios for amplitude and phase are very similar across the theory-specific and theory-agnostic models, as long as the unmodeled effect is the same. This generalizes the specific finding from Fig. 2 that amplitude and phase posteriors are very similar for the same unmodeled effect. Finally, the

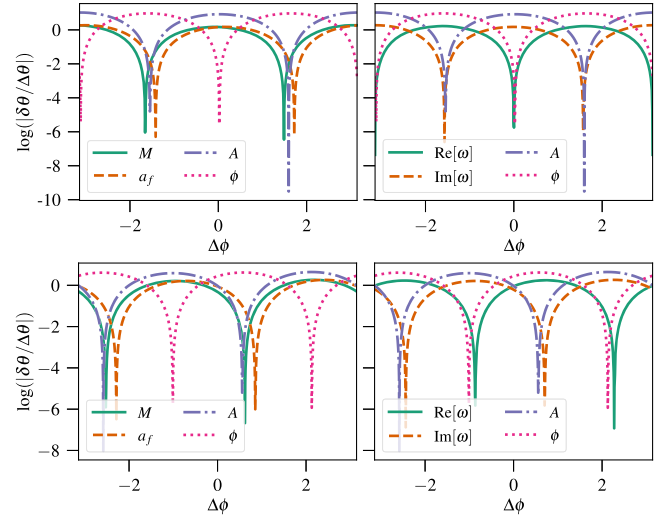


FIG. 3. *Bias ratios as a function of QNM phase differences:* In all panels, we show the bias ratio originating from an unmodeled QNM with a relative amplitude of 0.1 to the ω_{220} fundamental mode as a function of their phase difference $\Delta\phi$ for $\text{SNR} = 50$. In the top row, we consider ω_{221} , and in the bottom row, we consider the quadratic $\omega_{220 \times 220}$ as an unmodeled QNM. Left and right columns correspond to theory-specific and theory-agnostic models, respectively.

dependency of the bias ratios for the mass M and spin a or real and imaginary part of ω is quite different from each other. The value of phase differences at which one of its bias ratios vanishes seems quite unrelated.

2. Exploring start times of ringdown analysis

In our next example, we quantify how starting the ringdown analysis at later times changes the bias ratios. Similar to varying the phase shifts in Sec. III B, each choice for the starting time corresponds to an independent analysis. One would expect that any effect decaying faster than the fundamental mode should result in smaller biases, while the statistical errors become larger. In the presence of tails, the same scaling reasoning does not apply to the biases. The power-law tail will take over the signal at very late times, but depending on when it starts in the ringdown, and what value one chooses for t_{pole} , it may also be relevant at early times.

In Fig. 4, we show the bias ratio as a function of the starting time for the same cases of unmodeled overtone, quadratic mode, and power-law tail as in Sec. III A. Starting with the overtone and quadratic mode, we find striking differences in the qualitative behavior of the bias ratios. Although all bias ratios tend to decrease toward later times, which is expected due to the faster decay of the effects compared to the fundamental mode, the oscillatory patterns are very different. In the case of the unmodeled overtone, the bias ratios do not seem to oscillate at all, while the ones from the unmodeled quadratic oscillate significantly.

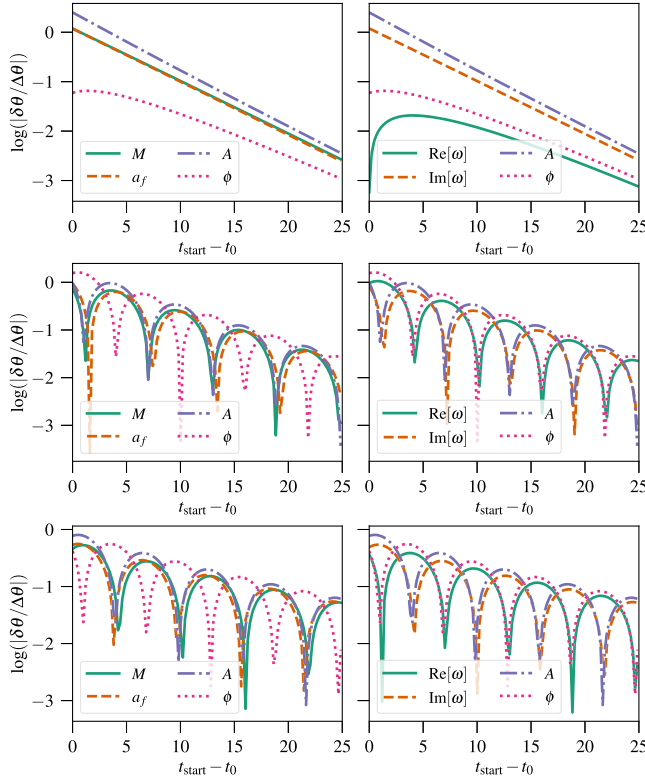


FIG. 4. *Bias ratios as a function of starting time:* In all panels, we show the bias ratio originating from an unmodeled QNM with a relative amplitude of 0.1 to the ω_{220} fundamental mode as a function of the starting time $t_{\text{start}} - t_{\text{peak}}$ for $\text{SNR} = 50$. The top, middle, and bottom rows show biases from unmodeled ω_{221} , quadratic $\omega_{220 \times 220}$, and power-law tail contributions, respectively. Left and right columns correspond to theory-specific and theory-agnostic models, respectively.

This is qualitatively in agreement with the observation that the difference in the real part of the overtone with respect to the fundamental mode is much smaller than the one between the quadratic mode and the fundamental mode. A quick look at the bias formula (8) in connection with the simple model (1) (using only two modes) reveals that the modulation is either very small (overtone case), or of the same frequency as the fundamental mode (quadratic case). The bias ratios of the power-law tail shown in the bottom row admit qualitatively similar oscillations as the quadratic case, but decays slower. Also, this can be understood if one realizes that the tail is a “low-frequency” effect and the only oscillation comes from the fundamental mode. A careful look reveals that the bias ratios do not exactly decay exponentially but show indications of a slower, power-law behavior.

We provide complementary results for $\text{SNR} = 100$ in the Appendix in Fig. 9 which, due to the validity of LSA, is just a simple rescaling.

3. Exploring the zoo of possibly unmodeled QNMs

In Fig. 5, we show the predicted biases for a variety of possibly unmodeled linear QNMs. Here we assume that the

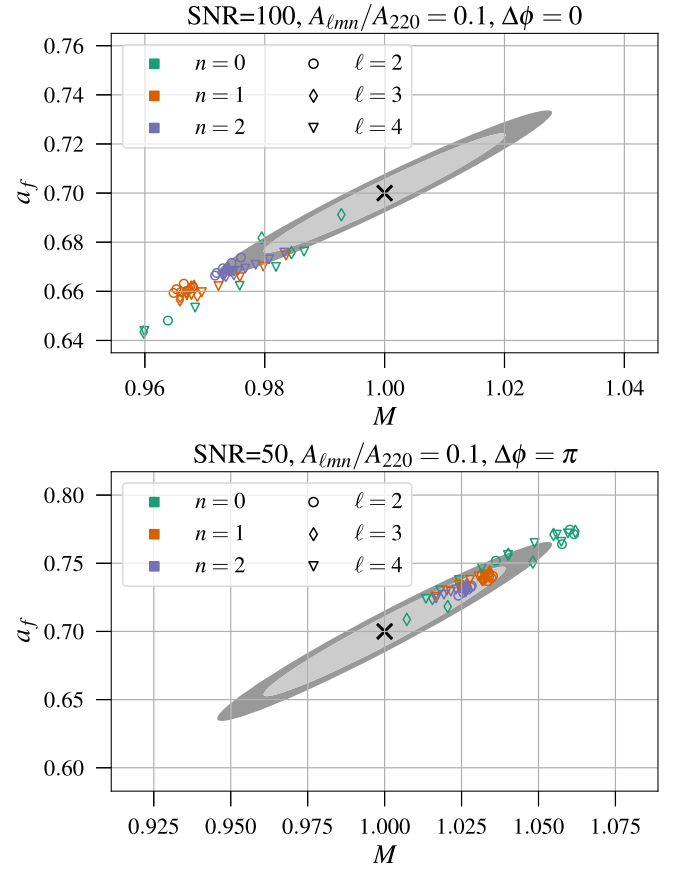


FIG. 5. *Exploring biases from unmodeled QNMs:* Here we show biases of various unmodeled linear QNMs described in the main text using LSA. The contours indicate 70% (light gray) and 90% (dark gray) probability levels using the LSA prediction (Fisher) of the theory-specific fundamental mode, i.e., the marginalized errors on the mass and spin. The true value is shown in the center (black cross). In all cases, biases are computed one at a time with a relative mode amplitude of $A_{\ell mn}/A_{220} = 0.1$ at $\text{SNR} = 50$. The top panel shows modes with phase difference $\Delta\phi = 0$, and the bottom panel with $\Delta\phi = \pi$.

signal contains only one unmodeled QNM at a time, and it has an amplitude ratio of $A_{\ell mn}/A_{220} = 0.1$ and the SNR is 50. The uncertainty region corresponds to 70% and 90% confidence intervals. We show cases with phase difference between the 220 fundamental mode and any other QNM of $\Delta\phi = 0$ and $\Delta\phi = \pi$. We include $\ell = [2, 3, 4]$, $m \in [-\ell, \dots, \ell]$ and $n \in [0, 1, 2]$. Complementary results for $\text{SNR} = 100$ can be found in the Appendix in Fig. 10. To ease the presentation and avoid cluttering, we now focus on the theory-specific case and only show the bias of the final mass and spin.

Overall, we find that biases are correlated and underestimate the mass and spin when being in phase ($\Delta\phi = 0$), and overestimate both when out of phase ($\Delta\phi = \pi$), which is similar to the findings reported in Ref. [95]. Moreover, likely due to the strong increase in damping time, biases from unmodeled overtones $n > 0$ are less important overall

compared to most, but not all, fundamental modes. Within the assumptions of LSA, one could, in principle, add biases from each QNM independently. Although this strategy might, in practice, break the assumption that the total effect is small, it is quite evident that biases from unmodeled QNMs of this amplitude are a significant challenge for tests of GR.

IV. DISCUSSION

We discuss the validity of LSA in Sec. IV A, the extension to realistic detectors in Sec. IV B, and future application of the LSA framework in Sec. IV C.

A. Validity of LSA

Within the range of applications considered in this work, we find LSA to be in very good agreement with accurate Bayesian MCMC results when considering SNRs of around 50. However, we have only considered cases in which the unmodeled effects are small. The bias formula is derived under the assumption that differences between signal and model are small and that the likelihood can be locally approximated by a multivariate Gaussian. In the following, we want to describe two cases for which LSA should be used with caution.

First, if one has a model with several QNMs, and some of them have low SNR, then the posterior distribution of a subset of parameters may not be well approximated. This may cause inaccurate predictions when evaluating the bias formula, due to the contraction with the covariance matrix. In such a case, the bias calculation may be limited to those parameters well captured by the Fisher information matrix.

Second, we comment on a subtle aspect regarding the high-SNR limit of the bias formula. In contrast to the Fisher information matrix, which should become more accurate with increasing SNR, the accuracy of the bias formula also depends on how similar the signal and model are. For very high SNRs [$\mathcal{O}(10^4) - \mathcal{O}(10^5)$], even a small but finite difference becomes relevant, i.e., the prediction of the bias itself has a finite error. Simply speaking, the predicted biases are a direct measure for the displacement of the means of the multivariate Gaussian describing the posterior, and the statistical errors simply encode its widths. Thus, for sufficiently high SNR, the true posterior and the one predicted from LSA would not overlap anymore, due to the finite error in predicting the systematic errors while statistical errors decrease. One may call this a systematic error of the systematic error itself, see also Ref. [84]. In our cases, it is much smaller than the statistical one and ellipses for the posteriors would only stop overlapping at much higher SNR (decreasing statistical errors) or if the strength of the unmodeled effect would be increased (increasing errors of the LSA prediction).

B. Extension to realistic detectors

As outlined in Sec. II B, we assume flat noise in the time domain to allow for the analytic computation of the Fisher information matrix. This is well suited for the purpose of analyzing high-accuracy simulations to obtain best-fit parameters and their systematic errors. However, this simplification is not adequate to make quantitative predictions for realistic gravitational wave detectors, as those are characterized by a frequency-dependent power-spectral density. Since our work does not focus on specific sources and detectors but scales the detector directly by the signal SNRs, it rather demonstrates the methodology. Note that this aspect is not important when comparing Bayesian analysis and LSA, as we use the same approximation in both analyses.

Future work focusing on detector-specific questions could either work with a more complicated detector response in the time domain, e.g., as described here [87,88], or transform the model and signal into the frequency domain. Although both approaches are, in principle, straightforward, one should not underestimate subtle complications due to the finite length of the signal, especially if the start time or end time are being varied [33,96–98]. A main complication would be that the Fisher information matrix cannot, in general, be computed analytically. This would reduce the overall efficiency of computations for large parameter space explorations and calls for a careful numerical treatment when exploring different regions of the parameter space. Nevertheless, LSA still provides an efficient way to complement traditional Bayesian MCMC injection and recovery campaigns.

C. Future applications of LSA framework

Given the computational advantages and robustness of LSA in our applications, there is a range of possible extensions that could be considered in future work. Since our study demonstrates the performance of LSA for several examples, one obvious extension could be to consider amplitudes and phases as extracted from NR simulations. This would allow one to connect the relevance of the various harmonics ℓmn with the properties of BBH systems before merger. Especially in light of large mass ratios, eccentricity, and precession, it is far from trivial to rank subdominant excitations. The systematic errors originating from not including or mismodeling the amplitudes and phases themselves, e.g., how accurate are their predictions, have not been explored in depth. We expect such challenging BBH systems with high SNR to be particularly important for next-generation detectors like LISA. In this context, LSA was recently applied to study ringdown systematics for the LISA detector in Refs. [92,93], which demonstrates that future black hole spectroscopy relies crucially on identifying how many QNMs need to be included, see also Ref. [99]. Finally, it would be intriguing to extend our analysis to the more general scenario in which

multiple QNMs are being modeled and multiple QNMs are treated as unmodeled effects. Such an analysis using LSA requires added care because the SNRs in the subleading order modes can be of $\mathcal{O}(\text{few})$ even when the leading fundamental mode has large SNR (~ 50), as can be seen in the recent event GW250114 [100].

V. CONCLUSIONS

Black hole spectroscopy is a cornerstone of gravitational wave astronomy and opens up precision tests of strong gravity and dynamic spacetimes of black holes. Extracting multiple QNMs from the ringdown of two merging compact objects allows one to extract the source's properties and test GR's predictions. However, conceptually difficult questions like how many QNMs need to be included in a ringdown model make it clear that theoretical uncertainties need to be taken seriously. A large volume of studies, partially with conflicting results and differences in working principles, support this further [31,36,37,42,55–59,96,97,101,102]. Although improvements in fitting routines allow one to extract more QNMs, the problem remains challenging, even in ideal conditions [40,42,43,46]. Thus, black hole spectroscopy must improve in quantifying theoretical uncertainties to provide robust results. Such steps are essential for unbiased parameter estimation and crucial for accurate tests of GR with the ringdown.

In this work, we demonstrated that the Fisher information matrix and bias formula, both primary tools from LSA, can yield fast and accurate estimates of theoretical uncertainties. In particular, we have focused on theory-specific and theory-agnostic ringdown tests in which a small unmodeled contribution contaminates the analyzed signal. To quantify the validity of LSA, we first studied representative examples with traditional Bayesian MCMC sampling in Sec. III A. LSA can, in most cases, accurately describe biases from an unmodeled overtone, quadratic QNM, or power-law tail at an SNR of 50. While such ringdown SNRs are not common with current GW detectors, the LVK Collaboration has already observed one such event—GW250114 [100]—which has a ringdown SNR of ~ 40 . With increasing detector sensitivity in upcoming observing runs, such events—expected to be 1% of the total detections [103]—will have ringdown SNRs > 50 . Furthermore, future observatories, such as the Einstein Telescope and Cosmic Explorer, will routinely observe BBH mergers with large ringdown SNRs [104,105].

In Sec. III B, we move to vast parameter space explorations for which traditional Bayesian MCMC sampling is computationally prohibitive. In Sec. III B 1 and Sec. III B 2, we provide a detailed overview of how phase differences of an unmodeled QNM and different choices of starting time impact biases, respectively. LSA explains typical fitting instabilities/inaccuracies in the presence of ringdown data with a tail contribution. Thus, even if tails are not visible in the analyzed signal, they can cause significant biases. This observation may

explain why QNM fitting in actual black hole signals, including a tail (even if not visible in the fitting window), is more difficult than for signals without a tail (either by construction or due to the underlying wave equation).

In Sec. III B 3, we explore the zoo of possibly unmodeled QNMs of various ℓmn and how they bias our estimates for mass and spin. We find that unmodeled QNMs with a relative amplitude ratio of only 0.1 are typically sufficient to obtain theoretical errors that exceed the statistical ones. Such contributions are clearly relevant if one starts the ringdown analysis later, when such QNMs may falsely be considered too small. Focusing on the fundamental mode, we find that biases from unmodeled overtones are less important overall compared to most, but not all, fundamental modes, if they are injected with the same relative amplitude.

In Sec. IV, we discussed that quantifying ringdown biases with LSA opens up many possible applications, ranging from analyzing high-precision simulations to practically mitigating theoretical errors in actual ringdown models. While some aspects of our work could already be studied using high-precision simulations, extending it to detector-specific studies with realistic noise properties will require some basic extensions.

ACKNOWLEDGMENTS

The authors thank Keefe Mitman for useful discussions, and Emanuele Berti, Nicola Franchini, and Saul Teukolsky for valuable comments on the manuscript. The authors also thank the anonymous referee for providing feedback that improved the presentation of this work. S. H. V. acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG): Project No. 386119226.

DATA AVAILABILITY

The data are not publicly available. The data are available from the authors upon reasonable request.

APPENDIX: ADDITIONAL RESULTS

In the following, we provide additional results to those presented in the main text. In Fig. 6, we compare LSA with Bayesian analysis for the same cases as in the main text, but for SNR = 20 and SNR = 100. In Fig. 7, we provide a detailed comparison of the bias ratio predicted from LSA and Bayesian analysis for selected SNRs. We define the Bayesian bias as the difference between the true parameter and the median of the MCMC samples. The statistical error is defined as half of the 68% highest density interval, which for a Gaussian distribution (valid at high SNR) corresponds to one sigma. In Fig. 8, we show the bias ratio as a function of phase difference for SNR = 100. In Fig. 9, we show the bias ratio as a function of starting time for SNR = 100. In Fig. 10, we provide the biases from various unmodeled QNMs at SNR = 100.

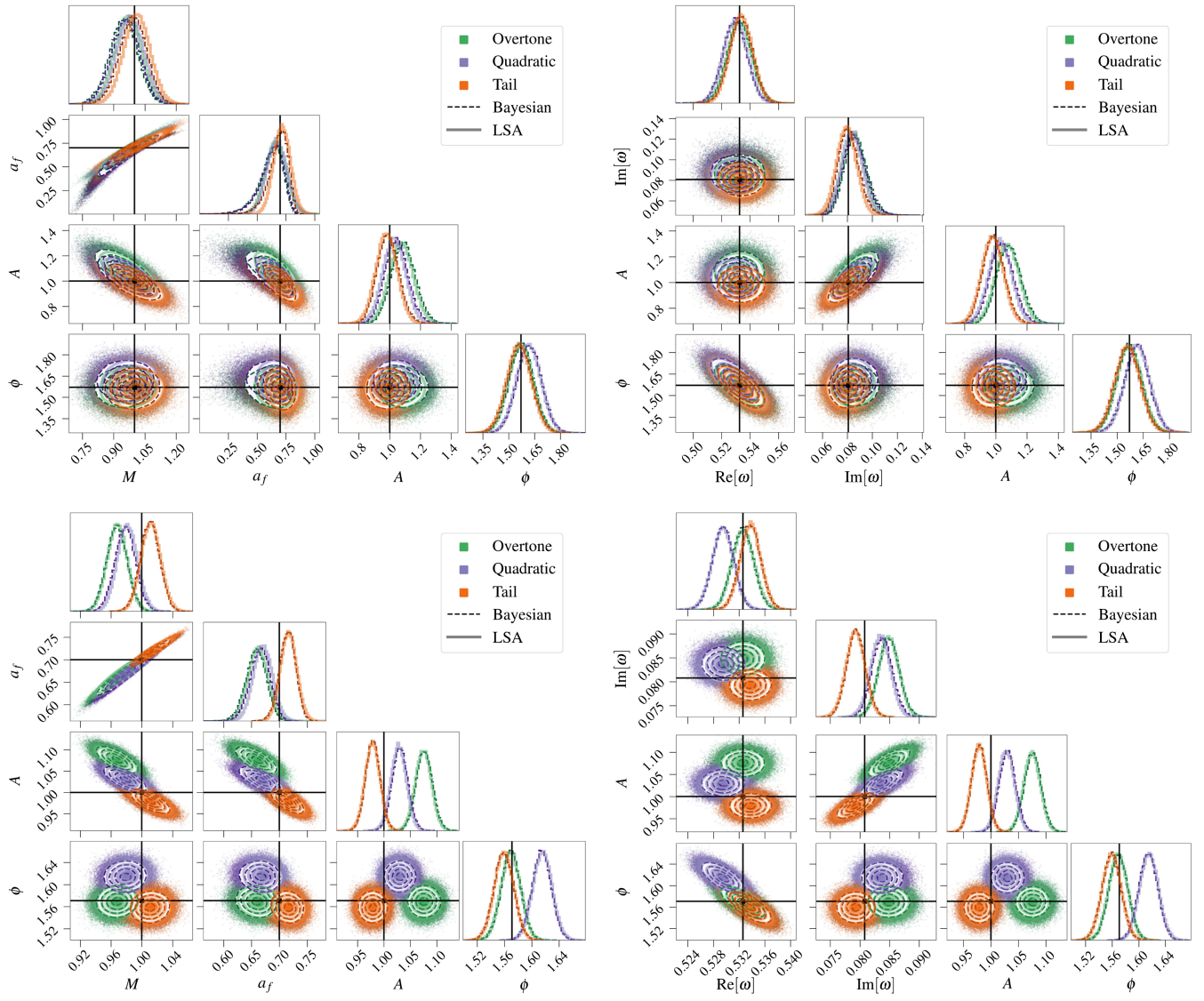


FIG. 6. *Comparison of LSA versus Bayesian analysis:* Same cases of unmodeled effects as studied in Fig. 2, but for lower SNR = 20 and higher SNR = 100.

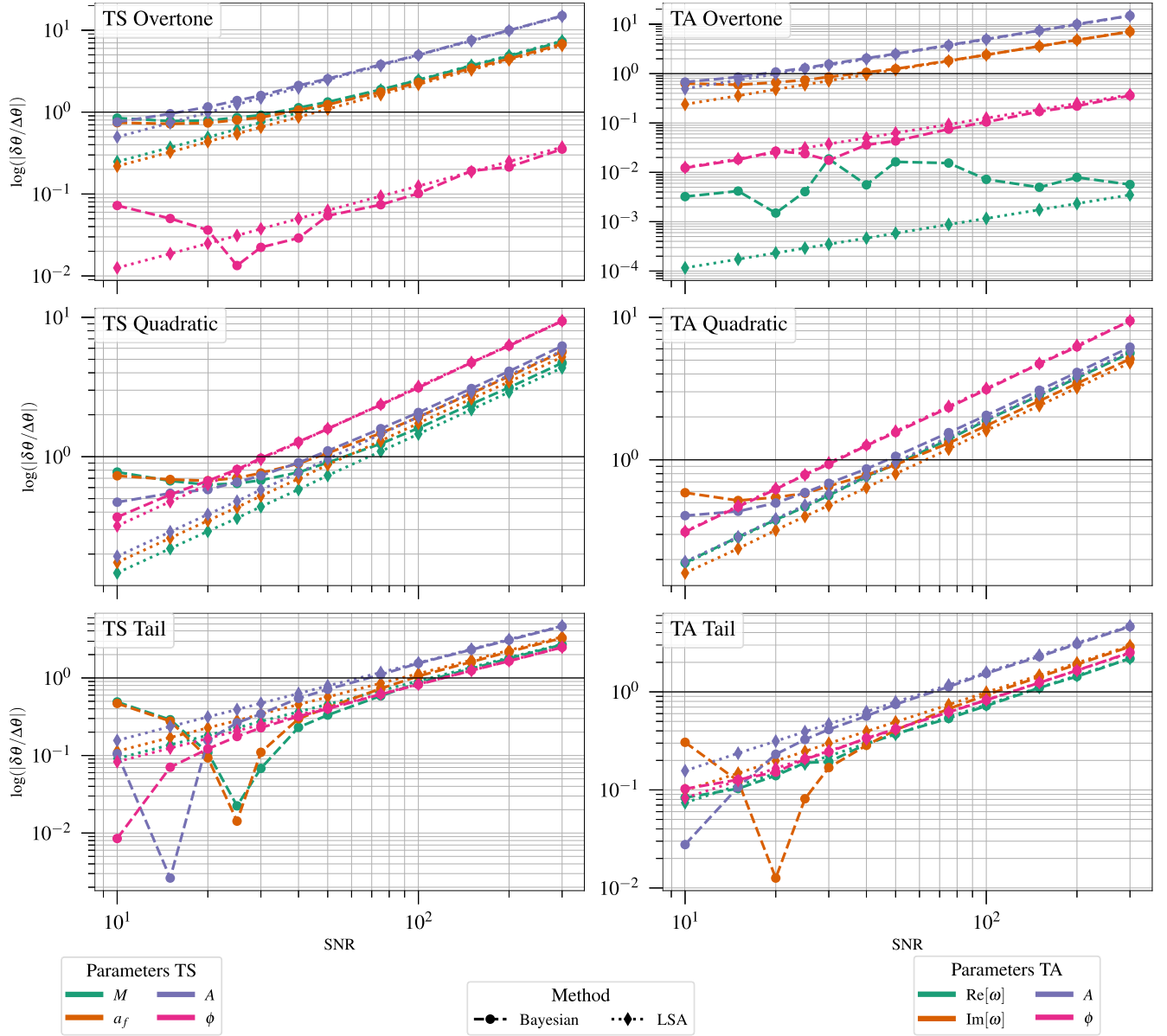


FIG. 7. Bias ratios obtained with Bayesian analysis and LSA as a function of SNR: Here we show the bias ratio for the theory-specific (left column) and theory-agnostic (right column) approach as a function of SNR.

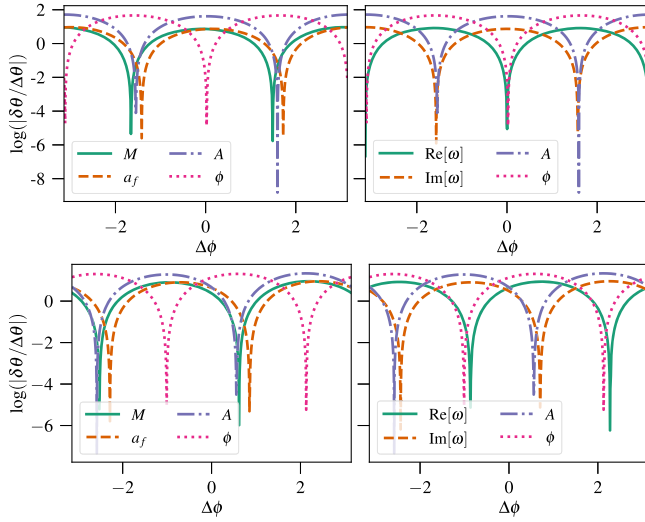


FIG. 8. *Bias ratios as a function of QNM phase differences:* Complementary results to those presented in Fig. 3, but for SNR = 100.

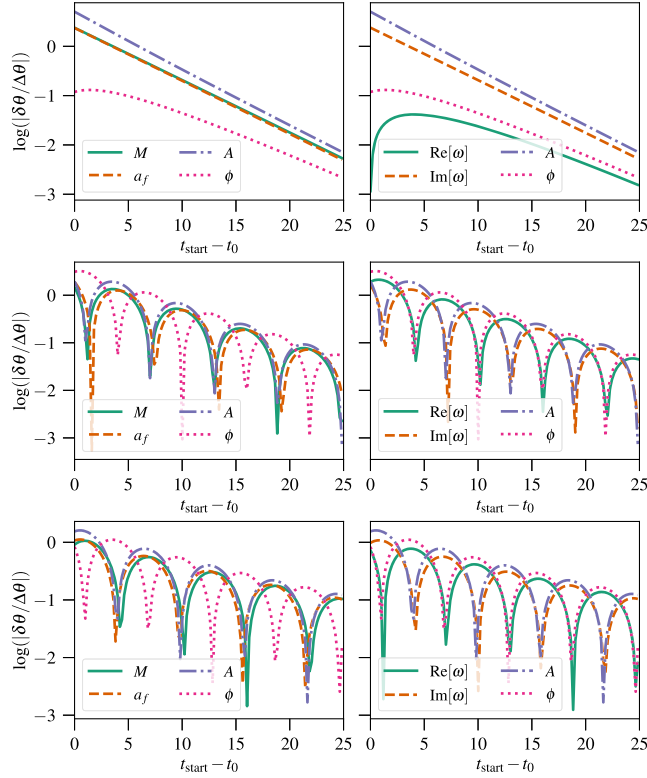


FIG. 9. *Bias ratios as a function of stating time:* Complementary results to those presented in Fig. 4, but for SNR = 100.

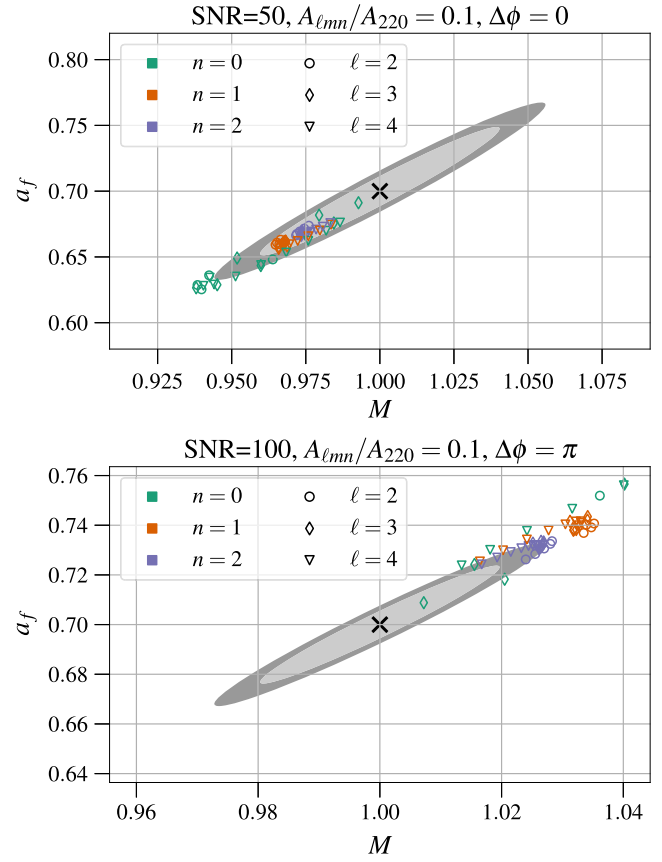


FIG. 10. *Exploring biases from unmodeled QNMs:* Here we show biases of various unmodeled linear QNMs with the same properties as those presented in Fig. 5 but for SNR = 100.

- [1] S. L. Detweiler, Black holes and gravitational waves. III—The resonant frequencies of rotating holes, *Astrophys. J.* **239**, 292 (1980).
- [2] O. Dreyer, B. J. Kelly, B. Krishnan, L. S. Finn, D. Garrison, and R. Lopez-Aleman, Black hole spectroscopy: Testing general relativity through gravitational wave observations, *Classical Quantum Gravity* **21**, 787 (2004).
- [3] E. Berti, V. Cardoso, and C. M. Will, On gravitational-wave spectroscopy of massive black holes with the space interferometer LISA, *Phys. Rev. D* **73**, 064030 (2006).
- [4] T. Regge and J. A. Wheeler, Stability of a Schwarzschild singularity, *Phys. Rev.* **108**, 1063 (1957).
- [5] F. J. Zerilli, Effective potential for even parity Regge-Wheeler gravitational perturbation equations, *Phys. Rev. Lett.* **24**, 737 (1970).
- [6] S. A. Teukolsky, Perturbations of a rotating black hole. 1. Fundamental equations for gravitational electromagnetic and neutrino field perturbations, *Astrophys. J.* **185**, 635 (1973).
- [7] K. D. Kokkotas and B. G. Schmidt, Quasinormal modes of stars and black holes, *Living Rev. Relativity* **2**, 2 (1999).
- [8] H.-P. Nollert, Topical review: Quasinormal modes: The characteristic ‘sound’ of black holes and neutron stars, *Classical Quantum Gravity* **16**, R159 (1999).
- [9] E. Berti, V. Cardoso, and A. O. Starinets, Quasinormal modes of black holes and black branes, *Classical Quantum Gravity* **26**, 163001 (2009).
- [10] R. A. Konoplya and A. Zhidenko, Quasinormal modes of black holes: From astrophysics to string theory, *Rev. Mod. Phys.* **83**, 793 (2011).
- [11] E. Berti *et al.*, Black hole spectroscopy: From theory to experiment, [arXiv:2505.23895](https://arxiv.org/abs/2505.23895).
- [12] W. Israel, Event horizons in static vacuum space-times, *Phys. Rev.* **164**, 1776 (1967).
- [13] S. W. Hawking, Black holes in general relativity, *Commun. Math. Phys.* **25**, 152 (1972).
- [14] B. Carter, Axisymmetric black hole has only two degrees of freedom, *Phys. Rev. Lett.* **26**, 331 (1971).
- [15] D. C. Robinson, Uniqueness of the Kerr black hole, *Phys. Rev. Lett.* **34**, 905 (1975).
- [16] A. Bachelot and A. Motet-Bachelot, Les résonances d’un trou noir de Schwarzschild, *Ann. L’Inst. Henri Poincaré Sect. (A) Phys. Theor.* **59**, 3 (1993).
- [17] H. R. Beyer, On the completeness of the quasinormal modes of the Poschl-Teller potential, *Commun. Math. Phys.* **204**, 397 (1999).
- [18] C. V. Vishveshwara, Scattering of gravitational radiation by a Schwarzschild black-hole, *Nature (London)* **227**, 936 (1970).
- [19] R. H. Price, Nonspherical perturbations of relativistic gravitational collapse. 1. Scalar and gravitational perturbations, *Phys. Rev. D* **5**, 2419 (1972).
- [20] E. W. Leaver, Spectral decomposition of the perturbation response of the Schwarzschild geometry, *Phys. Rev. D* **34**, 384 (1986).
- [21] C. Gundlach, R. H. Price, and J. Pullin, Late time behavior of stellar collapse and explosions: 1. Linearized perturbations, *Phys. Rev. D* **49**, 883 (1994).
- [22] C. Gundlach, R. H. Price, and J. Pullin, Late time behavior of stellar collapse and explosions: 2. Nonlinear evolution, *Phys. Rev. D* **49**, 890 (1994).
- [23] L. Barack, Late time dynamics of scalar perturbations outside black holes. 2. Schwarzschild geometry, *Phys. Rev. D* **59**, 044017 (1999).
- [24] R. F. Rosato and P. Pani, On the universality of late-time ringdown tail, *Phys. Rev. D* **112**, 024080 (2025).
- [25] H.-P. Nollert and R. H. Price, Quantifying excitations of quasinormal mode systems, *J. Math. Phys. (N.Y.)* **40**, 980 (1999).
- [26] S. Hod, The radiative tail of realistic gravitational collapse, *Phys. Rev. Lett.* **84**, 10 (2000).
- [27] S. Hod, Mode coupling in rotating gravitational collapse: Gravitational and electromagnetic perturbations, *Phys. Rev. D* **61**, 064018 (2000).
- [28] A. Buonanno, G. B. Cook, and F. Pretorius, Inspiral, merger and ring-down of equal-mass black-hole binaries, *Phys. Rev. D* **75**, 124018 (2007).
- [29] E. Berti, V. Cardoso, J. A. Gonzalez, U. Sperhake, M. Hannam, S. Husa, and B. Bruegmann, Inspiral, merger and ringdown of unequal mass black hole binaries: A multipolar analysis, *Phys. Rev. D* **76**, 064034 (2007).
- [30] L. London, D. Shoemaker, and J. Healy, Modeling ringdown: Beyond the fundamental quasinormal modes, *Phys. Rev. D* **90**, 124032 (2014); **94**, 069902(E) (2016).
- [31] M. Giesler, M. Isi, M. A. Scheel, and S. Teukolsky, Black hole ringdown: The importance of overtones, *Phys. Rev. X* **9**, 041060 (2019).
- [32] E. Finch and C. J. Moore, Searching for a ringdown overtone in GW150914, *Phys. Rev. D* **106**, 043005 (2022).
- [33] M. Isi and W. M. Farr, Revisiting the ringdown of GW150914, [arXiv:2202.02941](https://arxiv.org/abs/2202.02941).
- [34] S. Ma, K. Mitman, L. Sun, N. Deppe, F. Hébert, L. E. Kidder, J. Moxon, W. Throwe, N. L. Vu, and Y. Chen, Quasinormal-mode filters: A new approach to analyze the gravitational-wave ringdown of binary black-hole mergers, *Phys. Rev. D* **106**, 084036 (2022).
- [35] S. Ma, L. Sun, and Y. Chen, Black hole spectroscopy by mode cleaning, *Phys. Rev. Lett.* **130**, 141401 (2023).
- [36] V. Baibhav, M. H.-Y. Cheung, E. Berti, V. Cardoso, G. Carullo, R. Cotesta, W. Del Pozzo, and F. Duque, Agnostic black hole spectroscopy: Quasinormal mode content of numerical relativity waveforms and limits of validity of linear perturbation theory, *Phys. Rev. D* **108**, 104020 (2023).
- [37] P. J. Nee, S. H. Völkel, and H. P. Pfeiffer, Role of black hole quasinormal mode overtones for ringdown analysis, *Phys. Rev. D* **108**, 044032 (2023).
- [38] M. Isi and W. M. Farr, Comment on Analysis of ringdown overtones in GW150914, *Phys. Rev. Lett.* **131**, 169001 (2023).
- [39] G. Carullo, R. Cotesta, E. Berti, and V. Cardoso, Reply to Comment on “Analysis of ringdown overtones in GW150914”, *Phys. Rev. Lett.* **131**, 169002 (2023).
- [40] M. H.-Y. Cheung, E. Berti, V. Baibhav, and R. Cotesta, Extracting linear and nonlinear quasinormal modes from black hole merger simulations, *Phys. Rev. D* **109**, 044069 (2024); **110**, 049902(E) (2024).

- [41] A. Maselli, S. Yi, L. Pierini, V. Vellucci, L. Reali, L. Gualtieri, and E. Berti, Black hole spectroscopy beyond Kerr: Agnostic and theory-based tests with next-generation interferometers, *Phys. Rev. D* **109**, 064060 (2024).
- [42] M. Giesler *et al.*, Overtones and nonlinearities in binary black hole ringdowns, *Phys. Rev. D* **111**, 084041 (2025).
- [43] K. Mitman *et al.*, Probing the ringdown perturbation in binary black hole coalescences with an improved quasinormal mode extraction algorithm, *Phys. Rev. D* **112**, 064016 (2025).
- [44] F. Nobili, S. Bhagwat, C. Pacilio, and D. Gerosa, Ringdown mode amplitudes of precessing binary black holes, *Phys. Rev. D* **112**, 044058 (2025).
- [45] S. Thomopoulos, S. H. Völkel, and H. P. Pfeiffer, Ringdown spectroscopy of phenomenologically modified black holes, *Phys. Rev. D* **112**, 064054 (2025).
- [46] L. Gao *et al.*, Robustness of extracting quasinormal mode information from black hole merger simulations, *Phys. Rev. D* **112**, 024025 (2025).
- [47] L. Magaña Zertuche *et al.*, High-precision ringdown surrogate model for non-precessing binary black holes, *Phys. Rev. D* **112**, 024077 (2025).
- [48] N. Oshita, Ease of excitation of black hole ringing: Quantifying the importance of overtones by the excitation factors, *Phys. Rev. D* **104**, 124032 (2021).
- [49] H. Motohashi, Resonant excitation of quasinormal modes of black holes, *Phys. Rev. Lett.* **134**, 141401 (2025).
- [50] N. Oshita and V. Cardoso, Reconstruction of ringdown with excitation factors, *Phys. Rev. D* **111**, 104043 (2025).
- [51] N. Oshita, E. Berti, and V. Cardoso, Unstable chords and destructive resonant excitation of black hole quasinormal modes, *Phys. Rev. Lett.* **135**, 031401 (2025).
- [52] R. K. L. Lo, L. Sabani, and V. Cardoso, Quasinormal modes and excitation factors of Kerr black holes, *Phys. Rev. D* **111**, 124002 (2025).
- [53] E. Berti and V. Cardoso, Quasinormal ringing of Kerr black holes. I. The excitation factors, *Phys. Rev. D* **74**, 104020 (2006).
- [54] Z. Zhang, E. Berti, and V. Cardoso, Quasinormal ringing of Kerr black holes. II. Excitation by particles falling radially with arbitrary energy, *Phys. Rev. D* **88**, 044018 (2013).
- [55] A. Dhani, Importance of mirror modes in binary black hole ringdown waveform, *Phys. Rev. D* **103**, 104048 (2021).
- [56] A. Dhani and B. S. Sathyaprakash, Overtones, mirror modes, and mode-mixing in binary black hole mergers, *arXiv:2107.14195*.
- [57] X. J. Forteza and P. Mourier, High-overtone fits to numerical relativity ringdowns: Beyond the dismissed $n = 8$ special tone, *Phys. Rev. D* **104**, 124072 (2021).
- [58] K. Mitman *et al.*, Nonlinearities in black hole ringdowns, *Phys. Rev. Lett.* **130**, 081402 (2023).
- [59] G. B. Cook, Aspects of multimode Kerr ringdown fitting, *Phys. Rev. D* **102**, 024027 (2020).
- [60] G. Carullo and M. De Amicis, Late-time tails in nonlinear evolutions of merging black hole binaries, *arXiv:2310.12968*.
- [61] V. Cardoso, G. Carullo, M. De Amicis, F. Duque, T. Katagiri, D. Pereniguez, J. Redondo-Yuste, T. F. M. Spieksma, and Z. Zhong, Hushing black holes: Tails in dynamical spacetimes, *Phys. Rev. D* **109**, L121502 (2024).
- [62] M. De Amicis, S. Albanesi, and G. Carullo, Inspiral-inherited ringdown tails, *Phys. Rev. D* **110**, 104005 (2024).
- [63] T. Islam, G. Faggioli, G. Khanna, S. E. Field, M. van de Meent, and A. Buonanno, Phenomenology and origin of late-time tails in eccentric binary black hole mergers, *Phys. Rev. D* **112**, 024061 (2025).
- [64] S. Ma, K. C. Nelli, J. Moxon, M. A. Scheel, N. Deppe, L. E. Kidder, W. Throwe, and N. L. Vu, Einstein–Klein–Gordon system via Cauchy-characteristic evolution: Computation of memory and ringdown tail, *Classical Quantum Gravity* **42**, 055006 (2025).
- [65] M. De Amicis *et al.*, Late-time tails in nonlinear evolutions of merging black holes, *arXiv:2412.06887*.
- [66] S. Ma, M. A. Scheel, J. Moxon, K. C. Nelli, N. Deppe, L. E. Kidder, W. Throwe, and N. L. Vu, Merging black holes with Cauchy-characteristic matching: Computation of late-time tails, *Phys. Rev. D* **112**, 024003 (2025).
- [67] R. O’Shaughnessy, L. London, J. Healy, and D. Shoemaker, Precession during merger: Strong polarization changes are observationally accessible features of strong-field gravity during binary black hole merger, *Phys. Rev. D* **87**, 044038 (2013).
- [68] H. Zhu *et al.*, Black hole spectroscopy for precessing binary black hole coalescences, *Phys. Rev. D* **111**, 064052 (2025).
- [69] M. R. Sinha, L. Sun, and S. Ma, Impact of detector calibration accuracy on black hole spectroscopy, *arXiv:2506.15979*.
- [70] M. H.-Y. Cheung *et al.*, Nonlinear effects in black hole ringdown, *Phys. Rev. Lett.* **130**, 081401 (2023).
- [71] J. Redondo-Yuste, G. Carullo, J. L. Ripley, E. Berti, and V. Cardoso, Spin dependence of black hole ringdown nonlinearities, *Phys. Rev. D* **109**, L101503 (2024).
- [72] L. Sberna, P. Bosch, W. E. East, S. R. Green, and L. Lehner, Nonlinear effects in the black hole ringdown: Absorption-induced mode excitation, *Phys. Rev. D* **105**, 064046 (2022).
- [73] S. Yi, A. Kuntz, E. Barausse, E. Berti, M. H.-Y. Cheung, K. Kritos, and A. Maselli, Nonlinear quasinormal mode detectability with next-generation gravitational wave detectors, *Phys. Rev. D* **109**, 124029 (2024).
- [74] N. Khera, S. Ma, and H. Yang, Quadratic mode couplings in rotating black holes and their detectability, *Phys. Rev. Lett.* **134**, 211404 (2025).
- [75] O. Steppohn, S. H. Völkel, and T. Dietrich, Black hole spectroscopy of collapsing and merging neutron stars, *arXiv:2508.15534*.
- [76] M. Crisostomi, K. Dey, E. Barausse, and R. Trotta, Neural posterior estimation with guaranteed exact coverage: The ringdown of GW150914, *Phys. Rev. D* **108**, 044029 (2023).
- [77] C. Pacilio, S. Bhagwat, and R. Cotesta, Simulation-based inference of black hole ringdowns in the time domain, *Phys. Rev. D* **110**, 083010 (2024).
- [78] L. S. Finn and D. F. Chernoff, Observing binary inspiral in gravitational radiation: One interferometer, *Phys. Rev. D* **47**, 2198 (1993).
- [79] A. Krolak, J. A. Lobo, and B. J. Meers, Estimation of the parameters of the gravitational wave signal of a coalescing binary system, *Phys. Rev. D* **48**, 3451 (1993).

- [80] K. D. Kokkotas, G. Tsegas, and A. Krolak, Statistical analysis of the estimators of the parameters of the gravitational wave signal from a coalescing binary, *Classical Quantum Gravity* **11**, 1901 (1994).
- [81] E. E. Flanagan and S. A. Hughes, Measuring gravitational waves from binary black hole coalescences: 2. The waves' information and its extraction, with and without templates, *Phys. Rev. D* **57**, 4566 (1998).
- [82] C. Cutler and M. Vallisneri, LISA detections of massive black hole inspirals: Parameter extraction errors due to inaccurate template waveforms, *Phys. Rev. D* **76**, 104018 (2007).
- [83] A. Gupta *et al.*, Possible causes of false general relativity violations in gravitational wave observations, *SciPost Phys. Comm. Rep.* **5** (2024).
- [84] R. S. Chandramouli, K. Prokup, E. Berti, and N. Yunes, Systematic biases due to waveform mismodeling in parametrized post-Einsteinian tests of general relativity: The impact of neglecting spin precession and higher modes, *Phys. Rev. D* **111**, 044026 (2025).
- [85] P. A. Cano, L. Capuano, N. Franchini, S. Maenaut, and S. H. Völkel, Parametrized quasinormal mode framework for modified Teukolsky equations, *Phys. Rev. D* **110**, 104007 (2024).
- [86] S. Völkel and N. Franchini, sebastianvoelkel/parametrized_qnm_framework: Parametrized qnm framework (2024).
- [87] G. Carullo, W. Del Pozzo, and J. Veitch, Observational black hole spectroscopy: A time-domain multimode analysis of GW150914, *Phys. Rev. D* **99**, 123029 (2019); **100**, 089903(E) (2019).
- [88] M. Isi and W. M. Farr, Analyzing black-hole ringdowns, [arXiv:2107.05609](https://arxiv.org/abs/2107.05609).
- [89] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, EMCEE: The MCMC Hammer, *Publ. Astron. Soc. Pac.* **125**, 306 (2013).
- [90] A. Dhani, S. H. Völkel, A. Buonanno, H. Estelles, J. Gair, H. P. Pfeiffer, L. Pompili, and A. Toubiana, Systematic biases in estimating the properties of black holes due to inaccurate gravitational-wave models, *Phys. Rev. X* **15**, 031036 (2025).
- [91] V. Kapil, L. Reali, R. Cotesta, and E. Berti, Systematic bias from waveform modeling for binary black hole populations in next-generation gravitational wave detectors, *Phys. Rev. D* **109**, 104043 (2024).
- [92] S. Yi, F. Iacovelli, S. Marsat, D. Wadekar, and E. Berti, Systematic biases from the exclusion of higher harmonics in parameter estimation on LISA binaries, [arXiv:2502.12237](https://arxiv.org/abs/2502.12237).
- [93] L. Capuano, M. Vaglio, R. S. Chandramouli, C. L. Pitte, A. Kuntz, and E. Barausse, Systematic bias in LISA ringdown analysis due to waveform inaccuracy, [arXiv:2506.21181](https://arxiv.org/abs/2506.21181).
- [94] X. J. Forteza, S. Bhagwat, S. Kumar, and P. Pani, Novel ringdown amplitude-phase consistency test, *Phys. Rev. Lett.* **130**, 021001 (2023).
- [95] E. Berti, J. Cardoso, V. Cardoso, and M. Cavaglia, Matched-filtering and parameter estimation of ringdown waveforms, *Phys. Rev. D* **76**, 104044 (2007).
- [96] M. Isi, M. Giesler, W. M. Farr, M. A. Scheel, and S. A. Teukolsky, Testing the no-hair theorem with GW150914, *Phys. Rev. Lett.* **123**, 111102 (2019).
- [97] R. Cotesta, G. Carullo, E. Berti, and V. Cardoso, Analysis of ringdown overtones in GW150914, *Phys. Rev. Lett.* **129**, 111102 (2022).
- [98] Y.-F. Wang, C. D. Capano, J. Abedi, S. Kastha, B. Krishnan, A. B. Nielsen, A. H. Nitz, and J. Westerweck, A gating-and-inpainting perspective on GW150914 ringdown overtone: Understanding the data analysis systematics, [arXiv:2310.19645](https://arxiv.org/abs/2310.19645).
- [99] A. Toubiana, L. Pompili, A. Buonanno, J. R. Gair, and M. L. Katz, Measuring source properties and quasinormal mode frequencies of heavy massive black-hole binaries with LISA, *Phys. Rev. D* **109**, 104019 (2024).
- [100] The LIGO Scientific Collaboration, the Virgo Collaboration, and the KAGRA Collaboration, Black hole spectroscopy and tests of general relativity with GW250114, [arXiv:2509.08099](https://arxiv.org/abs/2509.08099).
- [101] A. Correia, Y.-F. Wang, J. Westerweck, and C. D. Capano, Low evidence for ringdown overtone in GW150914 when marginalizing over time and sky location uncertainty, *Phys. Rev. D* **110**, L041501 (2024).
- [102] Y.-F. Wang, C. D. Capano, J. Abedi, S. Kastha, B. Krishnan, A. B. Nielsen, A. H. Nitz, and J. Westerweck, A gating-and-inpainting perspective on GW150914 ringdown overtone: Understanding the data analysis systematics, [arXiv:2310.19645](https://arxiv.org/abs/2310.19645).
- [103] LIGO Scientific, VIRGO, and Kagra Collaborations *et al.*, GW230814: Investigation of a loud gravitational-wave signal observed with a single detector, [arXiv:2509.07348](https://arxiv.org/abs/2509.07348).
- [104] M. Maggiore *et al.* (ET Collaboration), Science case for the Einstein Telescope, *J. Cosmol. Astropart. Phys.* **03** (2020) 050.
- [105] S. Bhagwat, X. J. Forteza, P. Pani, and V. Ferrari, Ringdown overtones, black hole spectroscopy, and no-hair theorem tests, *Phys. Rev. D* **101**, 044033 (2020).