Black hole spectroscopy beyond Kerr: Agnostic and theory-based tests with next-generation interferometers

Andrea Maselli[1,2] Sophia Yi,[3] Lorenzo Pierini,[4,5] Vania Vellucci,[6,7,8] Luca Reali,[3]
Leonardo Gualtieri,[9,10] and Emanuele Berti[3]

[1]*Gran Sasso Science Institute (GSSI), I-67100 L'Aquila, Italy*
[2]*INFN, Laboratori Nazionali del Gran Sasso, I-67100 Assergi, Italy*
[3]*William H. Miller III Department of Physics and Astronomy, Johns Hopkins University,*
3400 North Charles Street, Baltimore, Maryland 21218, USA
[4]*Dipartimento di Fisica, "Sapienza" Universitá di Roma*
[5]*Sezione INFN Roma 1, P. Aldo Moro 5, 00185 Roma, Italy*
[6]*SISSA, International School for Advanced Studies, via Bonomea 265, 34136 Trieste, Italy*
[7]*INFN, Sezione di Trieste, via Valerio 2, 34127 Trieste, Italy*
[8]*IFPU, Institute for Fundamental Physics of the Universe, via Beirut 2, 34014 Trieste, Italy*
[9]*Dipartimento di Fisica, Universitá di Pisa,*
Edificio C–Polo Fibonacci Largo B. Pontecorvo 3, 56127 Pisa, Italy
[10]*Sezione INFN Pisa, L. Bruno Pontecorvo 3, 56127 Pisa, Italy*

Black hole spectroscopy is a clean and powerful tool to test gravity in the strong-field regime and to probe the nature of compact objects. Next-generation ground-based detectors, such as the Einstein Telescope and Cosmic Explorer, will observe thousands of binary black hole mergers with large signal-to-noise ratios, allowing for accurate measurements of the remnant black hole quasinormal mode frequencies and damping times. In previous work we developed an observable-based parametrization of the quasinormal mode spectrum of spinning black holes beyond general relativity (ParSpec). In this paper we use this parametrization to ask the following question: can next-generation detectors detect or constrain deviations from the Kerr spectrum by stacking multiple observations of binary mergers from astrophysically motivated populations? We focus on two families of tests: (i) agnostic (null) tests, and (ii) theory-based tests, which make use of quasinormal frequency calculations in specific modified theories of gravity. We consider in particular two quadratic gravity theories (Einstein-scalar-Gauss-Bonnet and dynamical Chern-Simons gravity) and various effective field theory-based extensions of general relativity. We find that robust inference of hypothetical corrections to general relativity requires pushing the slow-rotation expansion to high orders. Even when high-order expansions are available, ringdown observations alone may not be sufficient to measure deviations from the Kerr spectrum for theories with dimensionful coupling constants. This is because the constraints are dominated by "light" black hole remnants, and only few of them have sufficiently high signal-to-noise ratio in the ringdown. Black hole spectroscopy with next-generation detectors may be able to set tight constraints on theories with dimensionless coupling, as long as we assume prior knowledge of the mass and spin of the remnant black hole.

I. INTRODUCTION

The LIGO-Virgo-KAGRA gravitational-wave (GW) detector network has already observed about 100 events produced by merging compact binaries [1]. This number is expected to grow to several hundreds during the ongoing fourth observing run (O4), and by orders of magnitude when next-generation (XG) detectors, such as the Einstein Telescope (ET) in Europe [2] and Cosmic Explorer (CE) in the U.S. [3], will start taking data. The improved noise power spectral density (PSD) and observing volume of these detectors (see Fig. 1) will open the door to new and

exciting science, ranging from studies of the astrophysical formation scenarios of compact object binaries to precise tests of general relativity (GR) and cosmology [4–9].

A promising strategy to test gravity in the strong-field regime and to check if the remnants of these mergers really are the Kerr black holes (BHs) of general relativity is based on "black hole spectroscopy." The idea is that the entire quasinormal mode (QNM) frequency spectrum of Kerr BHs is uniquely determined by their mass M and dimensionless spin χ [10–12]. The measurement of the frequency and damping time of a single QNM yields the BH mass and spin. At least in principle, the measurement of three or more
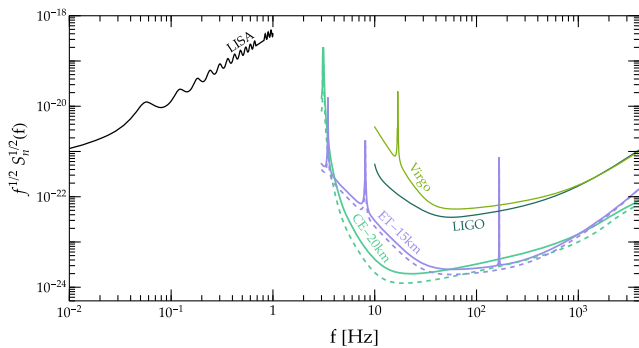
FIG. 1. Characteristic strain for ET and CE, compared to LIGO and Virgo at design sensitivity, as well as to the space-based interferometer Laser Interferometer Space Antenna (LISA). For ET (CE), solid and dashed curves correspond to detector configurations with 15 km (20 km) and 20 km (40 km) armlength, respectively. In this work we assume a 40 km CE [8], and a network of two L-shaped aligned 15 km ET [6].

quantities (either frequencies or damping times) in the so-called ringdown signal following a binary BH merger can be a test for consistency with the predictions of GR [13–15]. In practice, precision BH spectroscopy will probably require the detection of several ringdown events with the large signal-to-noise ratios (SNRs) expected from XG detectors, or from the space-based Laser Interferometer Space Antenna [16–26].

Tests of modified gravity theories can be either "theory agnostic" or "theory specific."

Agnostic tests are similar in spirit to the Will-Nordtvedt "parametrized-post-Newtonian" formalism [27]: the deviations of observable quantities from the GR predictions are described through a set of free parameters that can either be constrained from the data, or measured if a deviation is found. Theory-agnostic parametrizations are useful as long as the parameters can be mapped to the predictions of specific beyond-GR theories. In the parametrized-post-Newtonian expansion, the parametrization involves the spacetime metric [28]. In the so-called parametrized post-Einsteinian framework, the parametrization involves the amplitude and phase of the (inspiral) gravitational waveform [5,29–34]: see, e.g., [35,36] for reviews.

In the theory-specific approach, the goal is to compute the observable quantities (in our context, for example, the QNM frequencies) within each modified gravity theory. This is not always possible or practical. When it is, one can place bounds (or measure) the parameters of the theory by direct comparison with the data.

There have been several recent attempts at parametrizing the ringdown waveform (see, e.g., [37–43]). The special nature of the Kerr metric in GR implies that the perturbation equations are "miraculously" separable [44], leading to decoupled ordinary differential equations for the radial and angular dependence of the perturbations (the "Teukolsky equations" [45]). Separability is usually lost in beyond-GR theories, and the calculation of QNM frequencies usually

proceeds in one of two ways. The most common procedure is to first look for spherically symmetric BH solution, and then consider rotational corrections to the background metric and to the perturbation equations by working in a small-spin expansion. This approach is fairly common in GR [46,47], and it has recently been extended to theories such as Einstein-scalar-Gauss-Bonnet (EsGB) gravity [48,49], dynamical Chern-Simons (dCS) gravity [50–52], and certain classes of effective field theory (EFT) modifications of GR [53,54]. Other approaches try to bypass the small-spin expansion using either numerical or analytical methods. For example, one can use spectral methods to find the QNM spectrum by solving the (nonseparable) partial differential equations that describe perturbations of the Kerr-Newman metric in GR [55–58] (see also [59]). More recently, various groups have developed perturbation schemes leading to "modified Teukolsky equations" in modified theories of gravity [60–65]. These schemes have been applied to EFT modifications of GR in Ref. [66]; in this case the small-spin expansion is performed on the perturbations but not on the background.

The ParSpec parametrization of Ref. [67] is motivated, at least in part, by the small-spin approach to the problem of computing the QNM spectrum of BHs in modified gravity theories, but both approaches have found application in GW data analysis (see, e.g., [68] for an application of the ParSpec formalism to GW data and [69] for constraints on the Kerr-Newman QNMs computed via spectral methods).

The ParSpec formalism also relies on the assumption that modified gravity corrections can be treated perturbatively: while one can contemplate the possibility of nonperturbative corrections that can dramatically alter the spectrum [70], there are two strong arguments for working within a perturbative framework. The first argument is theoretical: most modified theories of gravity only make sense as EFTs, and a perturbative expansion around GR usually determines their domain of validity (see, e.g., Refs. [71–73] for discussions in the context of EsGB and dCS gravity, respectively). The second argument is experimental: GR is so well constrained that the perturbative approach used in this paper is general enough for most applications (see, e.g., [74]).

In this paper we focus on the observable-based ParSpec parametrization of the QNM spectrum developed in Ref. [67]. While the parametrization is theory agnostic, specific theories can be mapped to a given subset of the ParSpec parameters. In fact, in this paper we will discuss the implications of a ParSpec-based analysis of XG data on the three specific classes of theories listed above: EsGB, dCS, and various EFT theories. Our goal is to assess whether theory-agnostic and theory-specific BH spectroscopy tests are feasible with XG detectors.

The plan of the paper is as follows. In Sec. II we briefly review the ParSpec parametrization, our data analysis framework, and the population models used in our forecast.

In Sec. III we describe our implementation of theory-agnostic and theory-specific tests. In Sec. IV we present the results of our analysis, and in Sec. V we summarize our main conclusions. To improve readability, we relegate to the Appendices some of the most technical aspects, including the mapping between the ParSpec expansion and the QNM frequencies in specific theories (Appendix A); the remnant mass and spin accuracy achievable with ringdown observations (Appendix B); differences in ringdown SNR between ET and CE (Appendix C); and a list of the Fisher matrix components used in our parameter estimation calculations (Appendix D). Throughout the paper we use geometrical units ($G = c = 1$).

II. THE PARAMETRIZED SPECTROSCOPY PIPELINE

A. ParSpec

The key idea behind ParSpec is to express the QNM frequencies and damping times of a set of N observed ringdown events in terms of a two-parameter Taylor expansion, such that the (ℓ, m) QNM frequency $\omega_i^{\ell m}$ and damping time $\tau_i^{\ell m}$ of the ith source ($i = 1, \ldots, N$) read

$$M_i \omega_i^{\ell m} = \sum_{k_1=0}^{n_1} \chi_i^{k_1} \bar{\omega}_{\ell m}^{(k_1)} + \sum_{k_2=0}^{n_2} \chi_i^{k_2} \bar{\omega}_{\ell m}^{(k_2)} \gamma_i \delta\omega_{\ell m}^{(k_2)}, \quad (1)$$

$$\tau_i^{\ell m}/M_i = \sum_{k_1=0}^{n_1} \chi_i^{k_1} \bar{\tau}_{\ell m}^{(k_1)} + \sum_{k_2=0}^{n_2} \chi_i^{k_2} \bar{\tau}_{\ell m}^{(k_2)} \gamma_i \delta\tau_{\ell m}^{(k_2)}, \quad (2)$$

where M_i and χ_i are the remnant BH's detector-frame mass and dimensionless spin, and $\bar{\omega}_{\ell m}^{(k_1)}, \bar{\tau}_{\ell m}^{(k_1)}$ are the k_1-order dimensionless coefficients in a small-spin ($\chi_i \ll 1$) expansion of the mode's frequency and damping time for a Kerr BH. Beyond-Kerr corrections at order k_2 in the spin expansion are specified by the ParSpec dimensionless parameters $\delta\omega_{\ell m}^{(k_2)}$ and $\delta\tau_{\ell m}^{(k_2)}$. Given some "fundamental" modification of GR, these deviations are universal, i.e., they do not depend on the source.

We assume that the deviation from GR is characterized by a fundamental coupling constant α, with dimensions $[\alpha] = (\text{mass})^p$ ($p \geq 0$), which is considered as a small expansion parameter. The dimensionless coupling constants γ_i are linear in the coupling. They can depend on the source i (e.g., they can scale with some power of the remnant BH's mass in theories of gravity where the action is quadratic or of higher order in the curvature), but not on the specific QNM. At leading order in our perturbative scheme, they are given by

$$\gamma_i = \frac{\alpha}{M_i^p}(1 + z_i)^p, \quad (3)$$

where M_i is the ADM mass of the BH in the detector frame, and z_i is the redshift of the source. We will focus on three representative cases: $p = 0$, $p = 4$, and $p = 6$.

Within this expansion, GR is recovered (by definition) in the limit $\gamma_i \to 0$. Assuming GR corrections to be perturbative also requires that $|\gamma_i \delta\omega_{\ell m}^{(k_2)}| \lesssim 1$ and $|\gamma_i \delta\tau_{\ell m}^{(k_2)}| \lesssim 1$. Note that the order of the spin expansion of the QNM frequencies in GR, denoted by n_1 in Eqs. (1) and (2), can be different from the order of the spin expansion that allows for deviations from GR, n_2.

Consider N ringdown events in which we extract from the signal the frequencies and damping times of q QNMs for each source. This yields $\mathcal{O} = 2N \times q$ observables that can be used, at least in principle, to determine the beyond-GR parameters of the expansion in Eqs. (1) and (2), i.e., the quantities $\vec{\zeta} = (\delta\omega_{\ell m}^{(k_2)}, \delta\tau_{\ell m}^{(k_2)})$. As discussed in Ref. [67], there are two possible strategies to determine these quantities: (i) we assume that the (unperturbed) masses and spins of the sources, $(M_i, \chi_i)_{i=1\ldots N}$, are known a priori, or (ii) we extract these quantities from the ringdown signal by considering them as part of the parameter set to be determined, i.e., $\vec{\zeta} = (\delta\omega_{\ell m}^{(k_2)}, \delta\tau_{\ell m}^{(k_2)}, M_i, \chi_i)$.

Of course, the first strategy is preferable whenever possible, since it generally leads to smaller errors. This approach requires a priori knowledge of the masses and spins of the remnant. In GR, these quantities be obtained (for example) by using the parameters of the binary progenitors and semianalytical fits of numerical relativity simulations [75]. Going beyond GR, if we are interested in tests of a specific gravity theory, numerical relativity simulations in that theory would be necessary in order to derive similar fits. For tests not based on specific theories—hereafter, "agnostic" tests—and as long as the deviations from GR are small, we can assume $(M_i, \chi_i)_{i=1\ldots N}$ to be the GR masses and spins, obtained for example from the inspiral signal through the analytical fits derived within GR [75]. In this way, the shifts between the "true" masses and spins of the remnant and those obtained from the GR fits are absorbed in the unknown parameters $\delta\omega_{\ell m}^{(k_2)}$ (see the discussion in Ref. [67]). In summary: when performing theory-specific tests in a theory for which numerical relativity simulations are not available, we are forced to follow the second strategy, but the larger number of parameters is expected to lead to larger errors.

If we know in advance the masses and spins of the sources [strategy (i) above], then we can use the $\mathcal{O} = 2N \times q$ observables to constrain $2(n_2 + 1)q$ parameters, therefore we must have $N \geq n_2 + 1$. If, instead, we use only the ringdown signal [strategy (ii) above], there are now $2(n_2 + 1)q + 2N$ free parameters, and thus the minimum number of sources is $N > q(1 + n_2)/(q - 1)$. Constraining the GR modifications requires in this case at least two modes, i.e., $q > 1$, since the first mode is used to determine

the mass and spin. In this paper we will investigate both of these methodologies, extending the analysis of Ref. [67].

To infer the beyond-Kerr corrections we adopt a hierarchical inference scheme. For a given set of N astrophysical sources we compute the corresponding QNM frequencies and damping times, and assume they are observed by a XG detector, which can be either ET or CE. We compute the joint posterior probability distribution associated with the mode measurements using a Fisher Information Matrix (FIM) approach.

The observed values of $(\omega^{\ell m}, \tau^{\ell m})_{i=1...N}$ are interpreted in terms of the ParSpec template by sampling the $\vec{\zeta}$ parameters of Eqs. (1) and (2) through Markov chain Monte Carlo (MCMC) simulations. Below we give more details about each step of the analysis.

B. Quasinormal mode templates

For a given multipolar component of the radiation with angular indices (ℓm), the "plus" and "cross" polarizations of the ringdown waveform are given by

$$h_+(t) = \frac{M\mathcal{A}_{\ell m}Y_+^{\ell m}}{r}\mathrm{Re}[e^{-t/\tau_{\ell m}+i(\omega_{\ell m}t+\phi_{\ell m}-m\varphi)}], \quad (4)$$

$$h_\times(t) = \frac{M\mathcal{A}_{\ell m}Y_\times^{\ell m}}{r}\mathrm{Im}[e^{-t/\tau_{\ell m}+i(\omega_{\ell m}t+\phi_{\ell m}-m\varphi)}], \quad (5)$$

where the amplitude $\mathcal{A}_{\ell m}$ and the phase $\phi_{\ell m}$ are real quantities, we have defined

$$Y_+^{\ell m}(\iota) = {}_{-2}Y^{\ell m}(\iota,0) + (-1)^\ell {}_{-2}Y^{\ell-m}(\iota,0), \quad (6)$$

$$Y_\times^{\ell m}(\iota) = {}_{-2}Y^{\ell m}(\iota,0) - (-1)^\ell {}_{-2}Y^{\ell-m}(\iota,0), \quad (7)$$

ι is the inclination angle of the source with respect to the detector, and ${}_{-2}Y^{\ell m}(\iota,\varphi)$ are the spin-weighted spherical harmonics of spin-weight -2 [22]. To compute the signal in the frequency domain we follow the Flanagan-Hughes convention, i.e., we assume that the waveform for $t < 0$ is found by reflection through the origin of the waveform for $t > 0$, and then we divide the amplitude by a factor of $\sqrt{2}$ to compensate for the doubling [15,76]. Under this approximation the Fourier transforms of Eqs. (4) and (5) can be computed analytically, and they read

$$\tilde{h}_+(f) = \frac{M\mathcal{A}_{\ell m}}{\sqrt{2}r}[\hat{Y}_+^{\ell m}e^{i\phi_{\ell m}}b_+ + \hat{Y}_+^{\ell m\star}e^{-i\phi_{lm}}b_-], \quad (8)$$

$$\tilde{h}_\times(f) = \frac{M\mathcal{A}_{\ell m}}{i\sqrt{2}r}[\hat{Y}_\times^{\ell m}e^{i\phi_{\ell m}}b_+ - \hat{Y}_\times^{\ell m\star}e^{-i\phi_{\ell m}}b_-], \quad (9)$$

where $\hat{Y}_{+,\times}^{\ell m} = Y_{+,\times}^{\ell m}e^{-im\varphi}$ and $b_\pm$ are the Breit-Wigner functions,

$$b_\pm = \frac{1/\tau_{lm}}{\tau_{\ell m}^{-2} + (\omega \pm \omega_{\ell m})^2}. \quad (10)$$

The full waveform is then given by

$$\tilde{h}(f) = F_+\tilde{h}_+(f) + F_\times\tilde{h}_\times(f), \quad (11)$$

where $F_{+,\times}$ are the pattern functions of the detector, which depend on the source position in the sky and on the polarization angle [77].

Given the Fourier-domain amplitude of Eq. (11), we can compute the SNR ρ for a GW interferometer with noise power spectral density $S_n(f)$ from the relation

$$\rho^2 = 4\int_{f_{\min}}^{f_{\max}} \frac{\tilde{h}(f)\tilde{h}^\star(f)}{S_n(f)}df. \quad (12)$$

Hereafter we average over the detector and BH angular variables. Making use of the identities

$$\langle F_{+,\times}^2\rangle = \frac{1}{5}, \quad \langle F_+F_\times\rangle = 0, \quad \langle (Y_{\ell m}^+)^2 + (Y_{\ell m}^\times)^2\rangle = \frac{1}{\pi},$$

from Eq. (12) we find

$$\rho^2 = \frac{2}{5\pi}\frac{M^2\mathcal{A}_{\ell m}^2}{r^2}\int_{f_{\min}}^{f_{\max}} \frac{b_+^2 + b_-^2}{S_n(f)}df. \quad (13)$$

The SNR can be expressed in terms of the GW energy spectrum dE/df,

$$\rho^2 = \frac{2}{5\pi^2 r^2}\int_{f_{\min}}^{f_{\max}} \frac{1}{f^2 S_n(f)}\frac{dE}{df}df, \quad (14)$$

which in turn is related to the radiation efficiency

$$\epsilon_{\mathrm{RD}} = \frac{1}{M}\int_{f_{\min}}^{f_{\max}} \frac{dE}{df}df, \quad (15)$$

a measure of the amount of energy radiated by the (ℓ, m) mode. From Eqs. (13) and (15) we find

$$\epsilon_{\mathrm{RD}} = \pi M\mathcal{A}_{\ell m}^2\int_{f_{\min}}^{f_{\max}} dff^2(b_+^2 + b_-^2). \quad (16)$$

For a given choice of the source parameters $(M, r, \omega_{\ell m}, \tau_{\ell m}, \phi_{\ell m})$, and of the efficiency ϵ_{RD}, we can numerically solve Eq. (16) to find the signal amplitude $\mathcal{A}_{\ell m}$. We compute ϵ_{RD} for a given (ℓ, m) multipolar component from the semianalytic formula of Ref. [21]:

$$\epsilon_{\mathrm{RD}} = [a_{\ell m} + b_{\ell m}\chi_+ + c_{\ell m}\chi_-]^2, \quad (17)$$

where $a_{\ell m}$, $b_{\ell m}$, and $c_{\ell m}$ are functions of the binary mass ratio $q = m_1/m_2 \geq 1$ and of two spin parameters,

$$\chi_{\pm} = \frac{m_1\chi_1 \pm m_2\chi_2}{m_1 + m_2}, \tag{18}$$

which in turn depend on the masses $m_{1,2}$ and on the dimensionless spins $\chi_{1,2}$ of the progenitors.

C. Parameter estimation of quasinormal mode frequencies

To estimate the measurement accuracy of QNM frequencies and damping times by XG detectors, we use the waveform template of Eq. (11) and a FIM analysis [78].

The interferometer output $s(t) = h(t, \vec{\theta}) + n(t)$ is the sum of the detector's (stationary) noise $n(t)$ and of the actual signal. The probability distribution of the source parameters $\vec{\theta}$, up to a normalization constant, is

$$p(\vec{\theta}|s) \propto p^{(0)}(\vec{\theta})e^{-\frac{1}{2}(h(\vec{\theta})-s|h(\vec{\theta})-s)}, \tag{19}$$

where $p^{(0)}(\vec{\theta})$ is the prior on the source parameters [79], and we have introduced the inner product between two waveforms

$$(h_1|h_2) = 2\int_{f_{\min}}^{f_{\max}} \frac{\tilde{h}_1(f)\tilde{h}_2^{\star}(f) + \tilde{h}_1^{\star}(f)\tilde{h}_2(f)}{S_n(f)}df. \tag{20}$$

For GW signals with large SNRs, we expect $p(\vec{\theta}|s)$ to be peaked around the true values $\vec{\xi}$. Therefore, we can expand Eq. (19) as a Taylor series up to second order in $\Delta\vec{\theta} = \vec{\theta} - \vec{\xi}$, such that

$$p(\vec{\theta}|s) \propto p^{(0)}(\vec{\theta})e^{-\frac{1}{2}\Gamma_{ij}\Delta\theta^i\Delta\theta^j}, \tag{21}$$

where

$$\Gamma_{ij} = \left(\frac{\partial h}{\partial\theta^i}\bigg|\frac{\partial h}{\partial\theta^j}\right) \tag{22}$$

is the FIM, computed at $\vec{\theta} = \vec{\xi}$. The covariance matrix Σ^{ij} of the waveform parameters is then given by

$$\Sigma^{ij} = (\Gamma^{-1})^{ij}, \tag{23}$$

where Γ^{-1} is the inverse of the FIM. In the single-mode analysis, the waveform $\tilde{h}(f)$ is completely specified by four parameters: $\vec{\theta} = (\bar{\mathcal{A}}_{\ell m}, \phi_{\ell m}, \omega_{\ell m}, \tau_{\ell m})$, where we have defined the effective amplitude $\bar{\mathcal{A}}_{\ell m} = M\mathcal{A}_{\ell m}/r$. Hereafter, following the conventions of Ref. [15], we fix $\phi_{\ell m} = 0$. The components of Γ_{ij} are listed in Appendix D.

For the two-mode study, we replace the polarizations (4) and (5) with a sum over two multipolar components of the form [15]

$$h_+(t) = \sum_{\ell m} \frac{M\mathcal{A}_{\ell m}Y_+^{\ell m}}{r}\text{Re}\left[e^{-\frac{t}{\tau_{\ell m}}+i(\omega_{\ell m}t+\phi_{\ell m})}\right], \tag{24}$$

$$h_\times(t) = \sum_{\ell m} \frac{M\mathcal{A}_{\ell m}Y_\times^{\ell m}}{r}\text{Im}\left[e^{-\frac{t}{\tau_{\ell m}}+i(\omega_{\ell m}t+\phi_{\ell m})}\right]. \tag{25}$$

Here we consider two scenarios in which, along with the fundamental $\ell = m = 2$ mode, we include either the $(\ell, m) = (21)$ or the $(\ell, m) = (33)$ component (note that we do not include overtones in our analysis [26,80]). In this two-mode model, the Fisher matrix (22) becomes an 8×8 matrix in the variables $\vec{\theta} = (\bar{\mathcal{A}}_{\ell_1 m_1}, \phi_{\ell_1 m_1}, \omega_{\ell_1 m_1}, \tau_{\ell_1 m_1}, \bar{\mathcal{A}}_{\ell_2 m_2}, \phi_{\ell_2 m_2}, \omega_{\ell_2 m_2}, \tau_{\ell_2 m_2})$. Since the two modes have different harmonic indices (ℓ, m), by angular averaging and using the orthogonality of the spherical harmonics we can decouple the two modes. This property simplifies our calculations, since the FIM is block diagonal and can be written as a combination of the matrices $\Gamma_{ij}^{(1,2)}$ corresponding to each mode:

$$\Gamma_{ij} = \begin{pmatrix} \Gamma_{ij}^{(1)} & 0 \\ 0 & \Gamma_{ij}^{(2)} \end{pmatrix}. \tag{26}$$

Similarly, the total SNR is given by the sum in quadrature of the two modes $\rho = \sqrt{\rho_{(1)}^2 + \rho_{(2)}^2}$. These considerations are easily generalized to q modes with different values of (ℓ, m), as long as we do not consider overtones: in this case, the FIM is simply the block-diagonal combination of q matrices corresponding to the different QNMs.

In the evaluation of SNRs and Fisher matrix components, we fix the integration limits to $f_{\min} = 3$ Hz and $f_{\max} = 5$ kHz. For ET we consider a network of two L-shaped aligned detectors with 15 km armlength [6], while for CE we focus on single 40 km interferometer [8].

D. Sampling the beyond-Kerr parameters

From the covariance matrices for $\omega_{\ell m}$ and $\tau_{\ell m}$ we can infer the probability distribution of the beyond-Kerr parameters using a Bayesian approach based on MCMC simulations. In the most general case, given the set of parameters

$$\vec{\zeta} = \left(\delta\omega_{\ell_j m_j}^{(k)}, \delta\tau_{\ell_j m_j}^{(k)}, M_i, \chi_i\right) \quad \begin{cases} i = 1, \dots, N \\ j = 1, \dots, q \\ k = 1, \dots, n_2 \end{cases}, \tag{27}$$

for a set $\vec{d}$ of N ringdown observations consisting of $2N \times q$ frequencies and damping times, the posterior probability distribution of $\vec{\zeta}$ is given by

$$p(\vec{\zeta}|\vec{d}) \propto \mathcal{L}(\vec{d}|\vec{\zeta}) p_0(\vec{\zeta}), \tag{28}$$

where $\mathcal{L}(\vec{d}|\vec{\zeta})$ is the likelihood, which we choose to be a Gaussian for each event:

$$\mathcal{L}(\vec{d}|\vec{\zeta}) = \mathcal{N}(\vec{\mu}_i, \Sigma_i). \tag{29}$$

The vector $\vec{\mu}_i$ is defined as

$$\vec{\mu}_i = (\vec{\mu}_i^{(1)}, \dots, \vec{\mu}_i^{(q)})^T. \tag{30}$$

Each $\vec{\mu}_i^{(j)}$ is a two-component vector that depends on the difference between the observed mode frequencies ($j = 1, \dots, q$) and the parametrized templates in Eqs. (1) and (2):

$$\vec{\mu}_i^{(j)} = \begin{bmatrix} \omega_i^{\ell_j m_j} - \omega_{i,\text{obs}}^{\ell_j m_j} \\ \tau_i^{\ell_j m_j} - \tau_{i,\text{obs}}^{\ell_j m_j} \end{bmatrix}. \tag{31}$$

In Eq. (29) Σ_i denotes the covariance matrix (23), which includes errors on (and correlations between) the frequencies and damping times of the ith source. For each value of i, Σ_i is a block-diagonal ($4q \times 4q$) square matrix. Because the covariance matrix is block diagonal, the likelihood function for a q-mode analysis (neglecting the overtones) can be recast as a product of Gaussian distributions:

$$\mathcal{N}(\vec{\mu}_i, \Sigma_i) = \prod_{j=1}^{q} \mathcal{N}(\mu_i^{(j)}, \Sigma_i^{(j)}). \tag{32}$$

Finally, since the N ringdown observations are all independent, the combined likelihood function of the parameters can be factorized as

$$\mathcal{L}(\vec{d}|\vec{\zeta}) = \prod_{i=1}^{N} \mathcal{L}_i(\vec{d}|\vec{\zeta}) = \prod_{i=1}^{N} \prod_{j=1}^{q} \mathcal{N}(\mu_i^{(j)}, \Sigma_i^{(j)}). \tag{33}$$

We sample the posterior distribution using the EMCEE algorithm with stretch move [81]. For each dataset, we run $n \gg N$ walkers of 2×10^6 samples, discarding the first half as burn-in and applying a thinning factor of 0.2 for the remaining samples.

E. Astrophysical population models

We construct our binary black hole (BBH) population by sampling the masses from the POWER LAW+PEAK phenomenological model favored by the latest LIGO/Virgo/KAGRA catalog, GWTC-3 [82].

The primary mass m_1 follows a truncated power law with the addition of a Gaussian peak and an exponential tapering at low masses:

$$P(m_1) \propto [(1 - \lambda)P_{\text{law}}(m_1|\gamma_1, m_{\text{max}})$$
$$+ \lambda G(m_1|\mu_m, \sigma_m)]S(m_1|m_{\text{min}}, \delta_m). \tag{34}$$

Here, $P_{\text{law}}(m_1|\gamma_1, m_{\text{max}})$ is a power-law distribution with slope $\gamma_1 = -3.40$ and cutoff at $m_{\text{max}} = 86.85 M_\odot$, $G(m_1|\mu_m, \sigma_m)$ is a Gaussian distribution with mean $\mu_m = 33.73 M_\odot$ and standard deviation $\sigma_m = 3.36 M_\odot$, and $S(m_1|m_{\text{min}}, \delta_m)$ is a smoothing function that rises monotonically from 0 to 1 within $[m_{\text{min}}, m_{\text{min}} + \delta_m]$, with $m_{\text{min}} = 5.08 M_\odot$ and $\delta_m = 4.83 M_\odot$.

The secondary mass m_2 is obtained from the mass ratio $q = m_2/m_1$, sampled from a smoothed power law

$$P(q) \propto q^{\gamma_q} S(m_1 q|m_{\text{min}}, \delta_m), \tag{35}$$

with $\gamma_q = 1.08$.

We sample the BBH population up to redshift $z = 10$. The BBH redshift distribution is assumed to follow the Madau-Dickinson cosmic star-formation rate [83], with parameters taken from the phenomenological fit of Ref. [84]. The normalization is set by the local BBH merger rate $R_m = 28.3 \text{ Gpc}^{-3} \text{ yr}^{-1}$, which is favored by the GWTC-3 catalog for the POWER LAW+PEAK model [82].

We assume BH spins aligned or antialigned with the binary orbital angular momentum, and we sample their magnitudes from two different distributions. In model I, the dimensionless spin components $\chi_{1,2}$ follow a beta distribution with parameters $\alpha = 2$ and $\beta = 5$: this model is meant to qualitatively reproduce the Default spin model of the GWTC-3 catalog [82]. In model II, the individual dimensionless spin magnitudes $\chi_{1,2}$ are uniformly sampled within the range $[-1, 1]$.

III. CASE STUDIES

We consider here different and complementary case studies. The main purpose of these studies is to compare "agnostic" null tests of the Kerr QNM spectrum against theory-based approaches, which exploit direct calculations of the frequencies and damping times of spinning BH solutions in beyond-GR theories.

A. Agnostic tests

We focus first on agnostic null-hypothesis tests of GR, with the goal of understanding the largest deviation in the QNM spectrum compatible with Kerr predictions. In this case we assume the injected values of $(\omega_{i,\text{obs}}^{\ell_j m_j}, \tau_{i,\text{obs}}^{\ell_j m_j})$ in Eq. (31) to be computed in GR, and interpret them in terms of the ParSpec template. We consider two families of

beyond-Kerr corrections, specified by different mass dimensions of the fundamental coupling of the theory, α.

The simplest case we consider is that of a dimensionless coupling constant, i.e., $p = 0$ in Eq. (3). In this case, γ_i is the same for all sources ($\gamma = \alpha$), and it can be reabsorbed in the definitions of $\delta\omega^{(k_2)}_{\ell_j m_j}$ and $\delta\tau^{(k_2)}_{\ell_j m_j}$. This case includes certain families of scalar-tensor theories of gravity [85], and it was investigated in Ref. [67] under the assumption of a random distribution of the mass and SNR of the ringdown events (rather than the astrophysically motivated population models used in this work).

As a second scenario we consider the case of a coupling constant with dimensions $(\text{mass})^4$, i.e., $p = 4$ (3):

$$\gamma_i = \frac{\alpha}{M_i^4}(1 + z_i)^4. \tag{36}$$

Again, α can be absorbed in the definitions of the frequency and damping time deviation coefficients, so the number of parameters that must be constrained is unchanged. This case is particularly relevant because it includes notable examples of beyond-GR theories, such as EsGB gravity [48,49,86–88] and dCS gravity [50–52,73,89,90], as discussed in Sec. III B below.

For both parametrizations ($p = 0$ and $p = 4$), we perform the Bayesian analysis discussed in Sec. II D. We compare the inference of the ParSpec parameters in the "optimistic" case in which the masses and spins of the remnant BHs are inferred from the inspiral-merger phase against the "pessimistic" case in which we only use the ringdown phase, and therefore we need (at a minimum) two modes to perform a test. For the latter case, we will investigate in detail how the analysis depends on the specific choice of the "secondary" QNM used in the test.

In the $p = 0$ case, we assume that the ParSpec parameters $(\delta\omega^{(k_2)}_{\ell m}, \delta\tau^{(k_2)}_{\ell m})$ are drawn from uniform priors in the range $[-1, 1]$. In the $p = 4$ case, we use flat priors for $(\delta\omega^{(k_2)}_{\ell m}, \delta\tau^{(k_2)}_{\ell m})$ in the range $[-3 \times 10^4, 3 \times 10^4]$ km^4. We also impose the conditions $|\gamma_i(\delta\omega^{(k_2)}_{\ell m}, \delta\tau^{(k_2)}_{\ell m})| \lesssim 1$, as required by the perturbative character of the ParSpec expansion. Finally, when masses and spins are allowed to vary in the MCMC simulations, they are sampled from uniform distributions in the ranges $[20, 500]M_\odot$ and $[0.5, 1]$, respectively.

B. Theory-specific tests

As a top-down approach, we study deviations from Kerr ringdown waveforms focusing on three families of modified gravity theories: EsGB gravity, dCS gravity, and the same class of EFT models considered in Ref. [66]. These theories are particularly appealing in the context of our study for two reasons: (i) the BH solutions in all of these theories are different from those of GR (i.e., the no-hair

theorems do not apply), and (ii) the QNM frequencies of these BH solutions have been explicitly computed in the slow-rotation approximation at different orders in the slow-rotation expansion, as we discuss below.

In EsGB and dCS gravity, strong-field, large-curvature modifications are induced by a scalar field φ that is nonminimally coupled to the gravitational sector. In both classes of theories, the deviations from GR depend on a fundamental coupling constant (α_{GB} for EsGB gravity, α_{CS} for dCS gravity) with dimensions of $(\text{mass})^2$.

The action of EsGB gravity is

$$\mathcal{S}_{\text{GB}} = \int d^4x\sqrt{-g}\left(R - \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi + \frac{\alpha_{\text{GB}}}{4}e^\varphi\mathcal{R}^2_{\text{GB}}\right), \tag{37}$$

where $\mathcal{R}^2_{\text{GB}} = R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} - 4R_{\mu\nu}R^{\mu\nu} + R^2$ is the Gauss-Bonnet invariant, and $R_{\mu\nu\rho\sigma}, R_{\mu\nu}, R$ are the Riemann tensor, the Ricci tensor and the Ricci scalar, respectively [86].

The action of dCS gravity is

$$\mathcal{S}_{\text{CS}} = \int d^4x\sqrt{-g}\left(R - \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi + \frac{\alpha_{\text{CS}}}{4}{}^\star RR\varphi\right), \tag{38}$$

where ${}^\star RR = {}^\star R^\mu{}_\nu{}^{\kappa\delta}R^\nu{}_{\mu\kappa\delta}$ is the Pontryagin density, and ${}^\star R^{\mu\nu\kappa\delta}$ is the dual of the Riemann tensor [73]. In Eqs. (37) and (38) we do not consider the matter sector of the theory, because we are interested in astrophysical processes involving BHs. For simplicity, we also neglect the possibility of a nonzero scalar field potential.

In both theories, corrections with respect to the BH solutions in GR scale as $\beta_{\text{GB,CS}} = \alpha_{\text{GB,CS}}/M^2$. In the limit $\beta_{\text{GB,CS}} \to 0$ we recover the Kerr metric. The shifts in the QNM frequencies are proportional to the square of $\beta_{\text{GB,CS}}$, i.e., to the square of the fundamental coupling constant. Therefore, in both cases the parameter α appearing in Eq. (3)—say, $\alpha = \alpha^2_{\text{GB}}$ for EsGB gravity, and $\alpha = \alpha^2_{\text{CS}}$ for dCS gravity—has dimensions of $(\text{mass})^4$.

The gravitational QNM spectrum in EsGB gravity was computed in Ref. [91] for static BH solutions, and in Ref. [49] for slowly rotating BHs at second order in a small-spin expansion. For dCS gravity, the QNMs of static BHs coincide with those of GR. The QNM spectrum of slowly rotating BHs was computed in Refs. [50–52] at first order in a small-spin expansion. These calculations also assume a perturbative expansion in the coupling constant $\alpha_{\text{GB,CS}}$. In the spirit of the ParSpec framework, we truncate this weak-coupling expansion at leading order.

Following Ref. [66], for EFT gravity we consider the most general extension of GR specified by the following action:

$$\mathcal{S}_{\text{EFT}} = \frac{1}{16\pi}\int d^4x\sqrt{-g}(R + \ell^4[\lambda_{\text{evn}}\mathcal{R}^3 + \lambda_{\text{odd}}\tilde{\mathcal{R}}^3]$$
$$+ \ell^6[\epsilon_1\mathcal{C}^2 + \epsilon_2\tilde{\mathcal{C}}^2 + \epsilon_3\tilde{\mathcal{C}}\mathcal{C}]). \tag{39}$$

The action contains up to eight derivatives, encoded within the curvature scalars

$$\mathcal{R}^3 = R^{\rho\sigma}{}_{\mu\nu} R_{\rho\sigma}{}^{\delta\gamma} R_{\delta\gamma}{}^{\mu\nu}, \qquad \tilde{\mathcal{R}}^3 = R^{\rho\sigma}{}_{\mu\nu} R_{\rho\sigma}{}^{\delta\gamma} \tilde{R}_{\delta\gamma}{}^{\mu\nu},$$

$$\mathcal{C} = R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma}, \qquad \tilde{\mathcal{C}} = R_{\mu\nu\rho\sigma} \tilde{R}^{\mu\nu\rho\sigma}, \qquad (40)$$

where $\tilde{R}_{\alpha\beta\mu\nu} = \epsilon_{\alpha\beta\delta\sigma} R^{\delta\sigma}{}_{\mu\nu}$. Corrections to GR are controlled by the length scale $\ell \sim \Lambda_{\rm cut}^{-1}$, which is related to the EFT cutoff, and by the dimensionless coefficients $\{\lambda_{\rm evn,odd}, \epsilon_{1,2,3}\}$.

The QNM frequencies for slowly rotating BH solutions in these theories have recently been computed at high order in a small-spin expansion, including terms up to $\mathcal{O}(\chi^{12})$ [66]. For cubic and quartic theories, the deviations from the Kerr spectrum depend on the coupling constants

$$\beta_q^{\rm cubic} = \ell^4 \lambda_q / M^4 = \alpha_q^4 / M^4 \qquad q = \{{\rm evn, odd}\},$$

$$\beta_q^{\rm quartic} = \ell^6 \lambda_q / M^6 = \alpha_q^6 / M^6 \qquad q = \{1, 2, 3\}. \qquad (41)$$

The validity of the EFT approach requires that these coupling constants satisfy the conditions $\beta_q^{\rm cubic, quartic} \ll 1$. Within the ParSpec framework, the parameter α appearing in the coupling (3) is mapped to $\alpha = \alpha_q^4$ and $\alpha = \alpha_q^6$ for the cubic and quartic models, respectively. In this work we consider eight different EFT gravity theories, for which QNMs have been explicitly computed in Ref. [66]. In particular, we focus on the following: (i) two parity-violating models, identified by $q = \{{\rm odd}^+, 3^+\}$, and (ii) six parity-preserving theories, labeled by $q = \{{\rm evn}^+, {\rm evn}^-, 1^+, 1^-, 2^+, 2^-\}$.

At variance with the agnostic, null test approach of Sec. III A, for this analysis we assume that the injected values of $(\omega_{i,{\rm obs}}^{\ell_j m_j}, \tau_{i,{\rm obs}}^{\ell_j m_j})$ are given by the true values obtained in EsGB, dCS, or EFT gravity for a given combination of the BH mass, the BH spin, and the coupling constant.

We map the numerical frequencies and damping times computed in Refs. [49–52,66] to the template (1) and (2), i.e., we determine the values of $\delta\omega_{\ell_j m_j}^{(k_2)}$ and $\delta\tau_{\ell_j m_j}^{(k_2)}$ for each theory (and for each GW source) by setting

$$\gamma_i^{\rm cubic} = \beta_q^{\rm cubic} = \frac{\alpha_q^4}{M_i^4}, \qquad \gamma_i^{\rm quartic} = \beta_q^{\rm quartic} = \frac{\alpha_q^6}{M_i^6},$$

$$\gamma_{{\rm GB},i} = \beta_{\rm GB}^2 = \frac{\alpha_{\rm GB}^2}{M_i^4}, \qquad \gamma_{{\rm CS},i} = \beta_{\rm CS}^2 = \frac{\alpha_{\rm CS}^2}{M_i^4}. \qquad (42)$$

The technical details of the mapping can be found in Appendix A. With the shifts $(\delta\omega_{\ell_j m_j}^{(k_2)}, \delta\tau_{\ell_j m_j}^{(k_2)})$ fixed by the theory, we recover the injected QNMs using the ParSpec template by MCMC sampling on the coupling $\alpha_{\rm GB,CS}$, and also on (M_i, χ_i). This is because, at variance with the

"agnostic" case—where M_i, χ_i in Eqs. (1) and (2) are the GR masses and spin, and the shifts between those and the "true" masses and spins of the remnant are absorbed in the parameters $(\delta\omega_{\ell m}^{(k_2)}, \delta\tau_{\ell m}^{(k_2)})$—in theory-specific tests the BH mass and spin enter in Eqs. (1) and (2) with their physical values.

For this reason we must infer M_i and χ_i directly from each ringdown observation, through a two-mode analysis. In other words, we are forced to follow the second strategy discussed in Sec. II A. Fortunately, in this case the inclusion of the secondary QNM introduces a smaller set of parameters to be constrained compared to the agnostic approach. This is because the shifts are fixed, and the only free parameter beyond (M_i, χ_i) is the coupling constant $\alpha_{\rm GB,CS,q}$. We impose flat priors on the coupling constants of the three theories in the range $\alpha_{\rm GB,CS} \sim \mathcal{U}_{[0,100]}$ and $\alpha_q \sim \mathcal{U}_{[0,50]}$. Masses and spins are sampled from uniform distributions within $[3, 150]M_\odot$ and [0.5, 1], respectively.

C. Numerical setup

With all the ingredients described in the previous sections, we can now forecast the constraints on beyond-Kerr QNMs according to the following procedure:

(1) We compute the mass and the spin of the remnant for the binary systems of our catalog exploiting semi-analytic relations derived from NR [75]. For each BH we derive the fundamental mode ($\ell = m = 2$) and the secondary modes (either $\ell = m = 3$ or $\ell = 2$, $m = 1$) through the templates of Eqs. (1) and (2), both in GR and in the gravity theories described in Sec. III B. We choose independently the order of the spin expansion for the GR (n_1) and beyond-GR (n_2) coefficients. In particular we set $n_1 = 12$, as this is sufficient to reproduce the QNMs of the full Kerr solution for a typical merger remnant ($\chi \sim 0.7$) with an accuracy better than 1% [67]: see also the discussion in Appendix A. In the agnostic case we vary n_2 to study the constraints that can be inferred on different spin-dependent ParSpec coefficients. In the theory-specific study, n_2 is necessarily limited by the accuracy of the QNM calculations available so far, i.e., $n_2 = 2$, $n_2 = 1$, and $n_2 = 12$ for EsGB, dCS, and EFT gravity, respectively.

(2) We compute the radiation efficiency of all mode/binary configurations (and hence the corresponding amplitudes $\mathcal{A}_{\ell m}$) using Eq. (17). We then determine the SNR of each multipolar component (ℓm), and select only events for which the SNR of the (22) QNM is larger than the detectability threshold $\rho = 12$, as measured by ET or CE. From now on, we will refer to this subset as the "detectable" events. We repeat this procedure for both spin distributions ("model I" and "model II" of Sec. II E).

(3) We compute the FIM for each of the selected ringdown events, evaluating the joint likelihood distribution of frequencies and damping times, in both the single-mode and two-mode configurations. For theory-based tests we apply the FIM approach to QNMs computed in EsGB, dCS, and EFT gravity, which are uniquely determined by the coupling constant of the theory, together with the remnant's masses and spins. For simplicity, we select the same ensemble of events—i.e., those for which the SNR of the dominant mode in GR is larger than 12. This is justified, as the SNR computed in these theories of gravity typically differs by less than a percent from the values computed in GR.

(4) We use the likelihood functions as seeds for our Bayesian analysis to sample the ParSpec parameters for the agnostic tests (with $p = 0$ and $p = 4$), and the fundamental couplings $\alpha_{\mathrm{GB/CS}}$ for the theory-based tests. The remnant masses and spins are kept fixed in the single-mode analysis, while they can be allowed to vary together with the beyond-GR parameters in the two-mode approach.

When quoting constraints on all parameters of interest, we refer to the 90% highest posterior density intervals, except for distributions bounded by the prior, for which we report the one-side 90% probability interval. We show posterior distributions as histograms of the samples inferred from the MCMC simulations (or kernel density estimations of the latter, computed in *Mathematica* assuming a Gaussian kernel).

IV. RESULTS

The two panels of Fig. 2 show the SNR distribution as a function of the remnant mass for the events with a detectable ringdown, assuming the spin distribution of model I. Within the 10^5 BBH binaries of the catalog we find $N = 13547$ and $N = 11126$ detectable events for ET and CE, respectively. The mean SNR of observable events is $\rho \simeq 21$ ($\rho \simeq 19$) for ET (CE). The qualitative behavior for the model II population is similar, although the number of selected events decreases: we find $N = 12487$ for ET and $N = 9021$ for CE.

As shown in Fig. 3, different assumptions on the spins (referred to as "model I" and "model II" above) lead to significantly different mass-spin distributions of the BH remnants, with mean spin values clustering around $\chi_f \sim 0.75$ for model I and $\chi_f \sim 0.9$ for model II.

In Fig. 4 we show the distribution of the SNR for the two detectors and for the two spin models. Note that model I yields slightly larger values of the SNR for both ET and CE. In the frequency range which is most relevant for ringdown detection, CE and ET (in a single-detector configuration) have similar SNR (see the discussion in Appendix C). The difference in the number of events above the threshold is
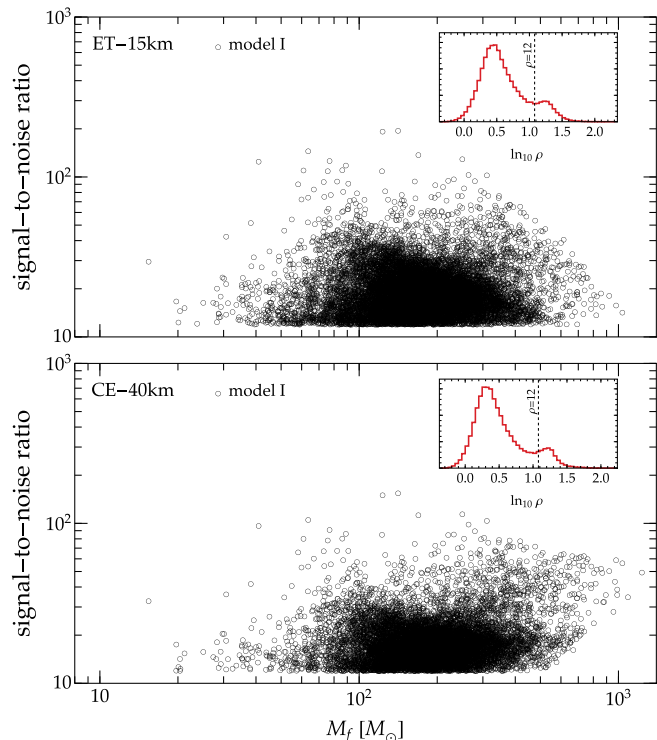

FIG. 2. SNR of the (22) mode as a function of the detector-frame remnant mass for detectable events in ET (top panel) and CE (bottom panel), assuming the model I population. "Detectable" events are those for which the SNR of the (22) QNM is larger than the detection threshold, $\rho \geq 12$. The inset in each panel shows the SNR distribution of the whole catalog, before applying the detection threshold cut. The SNR-M_f distribution for model II is similar.

mostly due to the fact that we consider a two-detector network configuration for ET. This boosts the ET SNR by a factor of $\sqrt{2}$. In Table I we summarize some salient features of these SNR distributions: for both detectors we estimate thousands of ringdown events with $\rho \geq 25$, but only a few events with $\rho \geq 100$.

The GW energy released in the different QNMs considered in our analysis is shown in Fig. 5. The component spins directly affect the spin of the remnant, leading to sensibly different results for the $\ell = m = 2$ mode amplitude in models I and II (top panel).

The dependence of the effective amplitude on the spin is quite mild for the (33) and the (21) QNM amplitudes, because our astrophysical population is dominated by comparable-mass binaries—cf. Eq. (35)—and odd-m multipoles of the radiation vanish by symmetry for equal-mass binaries [92,93]. Unlike the (33) mode, the (21) mode is quite sensitive to the binary component spins: see, e.g., Eq. (3) of Ref. [94]. This explains why the (21) mode amplitudes for models I and II, shown in the bottom panel of Fig. 5, are quite different. However, for our population model the (33) mode generally yields larger values of the ringdown efficiency ϵ_{RD}.
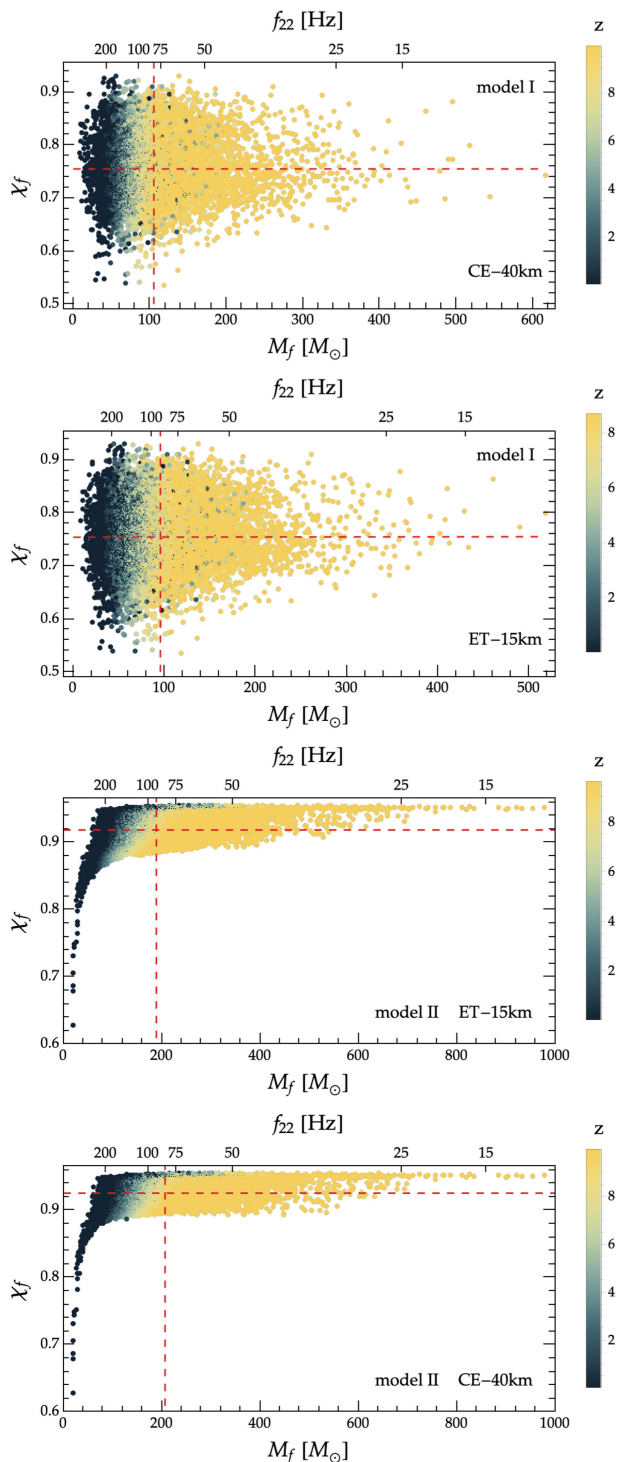
FIG. 3. Distribution of the remnant detector-frame masses and spins obtained from the binary population catalog for detectable events. The color code identifies the redshift associated to the binary, for a specific detector. Dashed lines correspond to the mean values of the distributions. For reference, the top x axis in each panel shows the frequency of the fundamental (22) mode for a remnant BH with dimensionless spin $\chi_f = 0.8$, and mass corresponding to the bottom x axis.
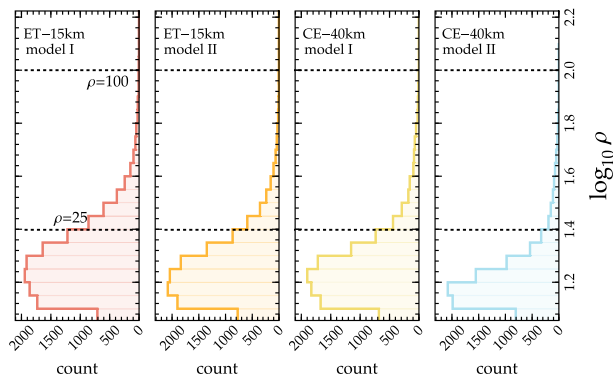

FIG. 4. Distribution of the SNR for the fundamental $\ell = m = 2$ mode, for the selected events observed by ET and CE above the detectability threshold $\rho = 12$.

The four panels of Fig. 6 show the percentage relative errors on the frequency and on the damping time of the fundamental mode with $\ell = m = 2$ inferred through the FIM approach described in Sec. II C, for all detectable ringdown events. For both ET and CE the mode frequency can be measured with an accuracy better than 1%, while the damping time is usually measured with less accuracy. Only few observations (those with $\rho \gtrsim 100$) have relative errors on τ_{22} at the level of 2%–3%.

In Fig. 7 we show the corresponding errors for the secondary modes. As expected from our discussion of the effective amplitudes, the error histograms are almost independent of the underlying astrophysical spin distribution, with the frequency and damping time of the (33) mode being better measured than the frequency and damping time of the (21) mode.

A. Agnostic inference

We start by exploring agnostic tests of GR, using simple templates at first and then progressively complicating the analysis. We follow the two strategies discussed in Sec. II A: (i) we assume that the (unperturbed) masses and spins of the sources, $(M_i, \chi_i)_{i=1\ldots N}$, are known *a priori*; or (ii) we extract these quantities from the ringdown signal by

TABLE I. Number of ringdown observations with signal-to-noise ratio of the fundamental mode larger than (12, 25, 100) for the two spin distributions we consider. The last column shows the maximum SNR of each population.

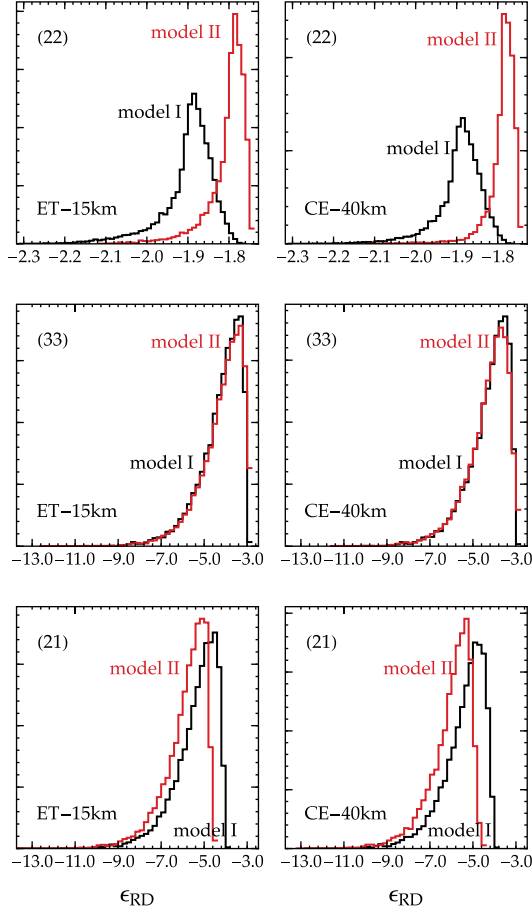|	$\rho \geq 12$	$\rho \geq 25$	$\rho \geq 100$	max ρ
ET—model I	13547	2554	14	196
CE—model I	11126	1454	6	155
ET—model II	12487	1662	7	164
CE—model II	9021	754	2	119

FIG. 5. Distribution of the energy released in the (22), (33), and (21) modes for the ringdown events with signal-to-noise ratio of the fundamental QNM larger than 12, as observed by ET and CE, assuming both population models described in Sec. II E.
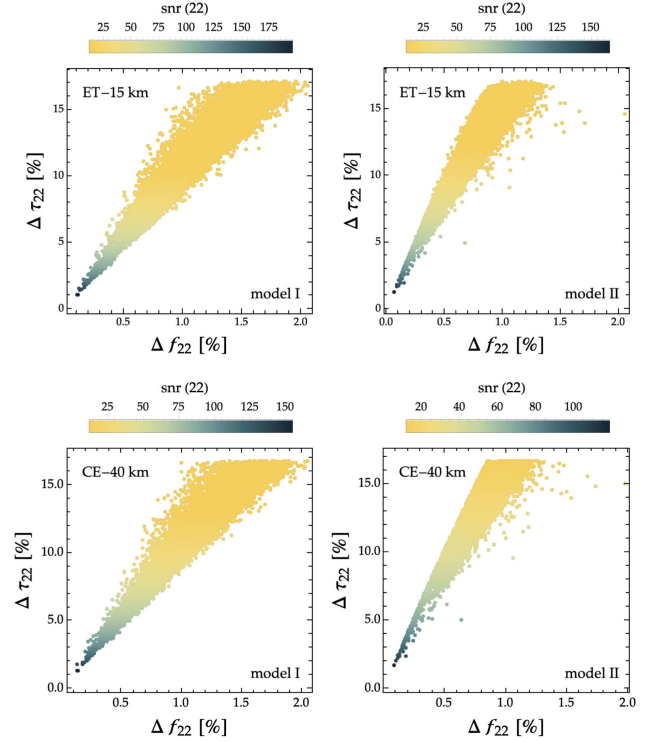

FIG. 6. Relative percentage errors on the frequency and the damping time of the fundamental QNM with $\ell = m = 2$ for the binary systems described in Sec. III C. The color scheme identifies the SNR of the (22) mode as detected by ET (top row) and CE (bottom row).

considering them as part of the parameter set to be determined, i.e., $\vec{\zeta} = (\delta\omega_{\ell m}^{(k_2)}, \delta\tau_{\ell m}^{(k_2)}, M_i, \chi_i)$.

1. Strategy (i): Mass and spin are known

No spin-dependent corrections ($n_2 = 0$), dimensionless coupling ($p = 0$). We consider QNM modifications at zeroth order in the spin ($n_2 = 0$), assume that the underlying theory has a dimensionless coupling ($p = 0$), and investigate the constraints on this class of beyond-GR modifications with the ParSpec template of Eqs. (1) and (2). We perform a single-mode analysis, i.e., we use only observations of the (22) mode, with both mass and spin of the remnant being set equal to their GR values (see the discussion in Sec. II A and in Ref. [67]).

In Fig. 8 we show how the posteriors of $\delta\omega_{22}^{(0)}$ and $\delta\tau_{22}^{(0)}$ get tighter by stacking an increasing number of sources. The probability distributions for beyond-GR corrections in both the frequency and damping time are nearly symmetrical around zero (the GR value). For a given number of observations, ET provides slightly stronger bounds

than CE, mainly because the ringdown SNR is typically larger. The posterior densities get narrower as the number of events, N, grows. For large N, the width of the distribution at 90% credible level is well approximated by a $\sim N^{-1/2}$ scaling for both spin models (models I and II) and for both detectors (ET and CE).

The joint posterior distributions (that we do not display for brevity) show that the ParSpec parameters $\delta\omega_{22}^{(0)}$ and $\delta\tau_{22}^{(0)}$ are mostly uncorrelated with each other, and that $\delta\omega_{22}^{(0)}$ is always better constrained than $\delta\tau_{22}^{(0)}$. The constraints are almost identical for models I and II, i.e., they are insensitive to the spin prescription. This will be discussed in greater detail below, when we will include spin-dependent corrections in the analysis. By stacking all of the events for model I we find $|\delta\omega_{22}^{(0)}| \lesssim 1.8 \times 10^{-4}$, $|\delta\tau_{22}^{(0)}| \lesssim 1.4 \times 10^{-3}$ for ET and $|\delta\omega_{22}^{(0)}| \lesssim 2.1 \times 10^{-4}$, $|\delta\tau_{22}^{(0)}| \lesssim 1.7 \times 10^{-3}$ for CE. For model II the bounds are $|\delta\omega_{22}^{(0)}| \lesssim 1.5 \times 10^{-4}$, $|\delta\tau_{22}^{(0)}| \lesssim 2.2 \times 10^{-3}$ for ET and $|\delta\omega_{22}^{(0)}| \lesssim 1.9 \times 10^{-4}$, $|\delta\tau_{22}^{(0)}| \lesssim 2.9 \times 10^{-3}$ for CE, respectively.

Spin-dependent corrections ($n_2 = 2$), dimensionless coupling ($p = 0$). We now augment the ParSpec template by including spin corrections up to second order in χ_f, i.e., we set $n_2 = 2$ in Eqs. (1) and (2). In Fig. 9 we show the
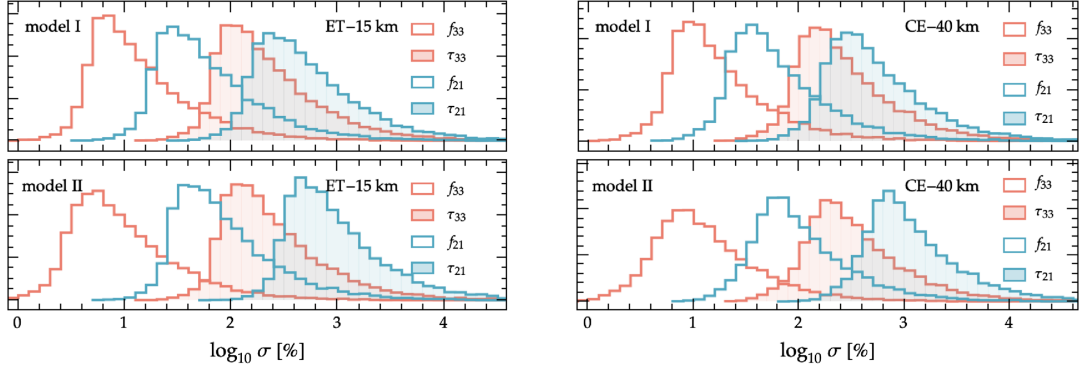
FIG. 7. Relative percentage errors on the frequency and the damping time of the secondary modes, $(\ell, m) = (33)$ and $(\ell, m) = (21)$, for ET and CE. The top and bottom panels refer to the two population models we considered.
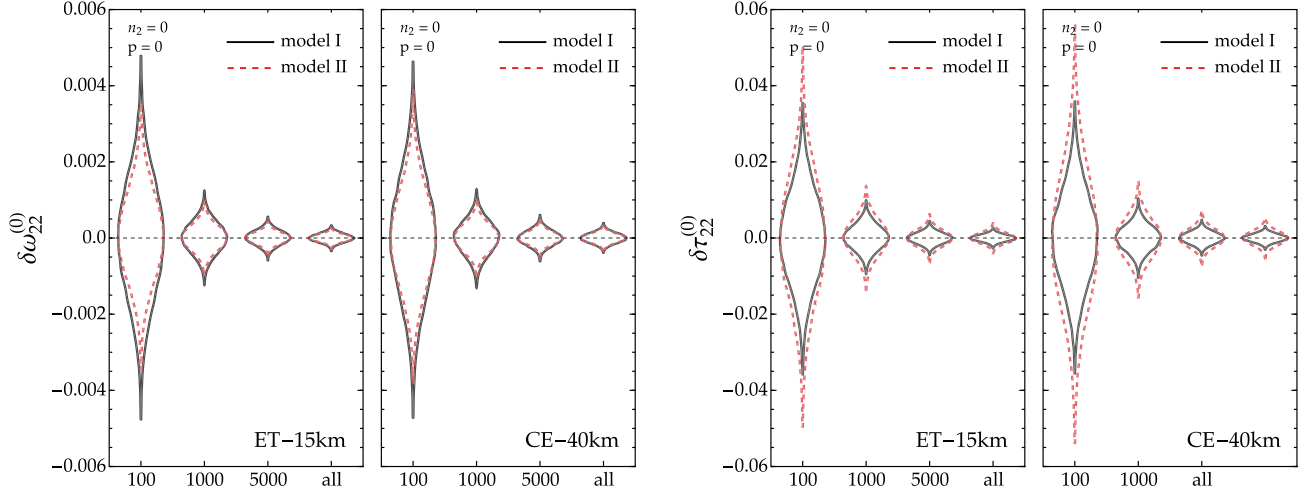

FIG. 8. Posterior probability densities for the nonrotating corrections $(\delta\omega_{22}^{(0)}, \delta\tau_{22}^{(0)})$, inferred in the agnostic analysis assuming strategy (i), as functions of the number of sources we analyzed. Results are derived following a single-mode analysis, in which only the fundamental (22) mode is taken into account. Solid and dashed contours refer to the two population models adopted in this paper.

68% and 90% credible regions of the joint posteriors of the beyond-GR parameters that are constrained by the data. For brevity we focus on ET, and we stack the full set of detectable BH events. The results in Fig. 9 confirm that there is a definite hierarchy among the ParSpec parameters: the posterior distributions of $\delta\omega_{22}^{(0)}, \delta\omega_{22}^{(1)}$, and $\delta\omega_{22}^{(2)}$ are informative with respect to the prior support, while all spinning corrections to the damping time (with the exception of $\delta\tau_{22}^{(0)}$) are unconstrained. The coefficients that control the frequency deviations are strongly correlated with each other, while they are almost uncorrelated with $\delta\tau_{22}^{(0)}$.

These broad conclusions hold regardless of the detector and population models, as shown by the first two columns of Table II, where we list the 90% credible intervals of the parameters for both ET and CE. Model II yields in general looser constraints on all the ParSpec coefficients, because the typical SNRs and the number of stacked events are lower

than those for model I. While the nonrotating corrections $(\delta\omega_{22}^{(0)}, \delta\tau_{22}^{(0)})$ are constrained with the tightest accuracy, their bounds increase by more than two orders of magnitude compared to the case shown in Fig. 8, due to correlations among the spin-dependent ParSpec coefficients.

We have also studied how the constraints change with the total number of stacked ringdown observations. In Table II we show the 90% credible levels for $N = 1000$ and $N = 500$ events. The bounds on the parameters broaden (as expected) as N decreases, and the posterior on some of the frequency modifications becomes uninformative, or almost as large as the prior.

Dimensionful coupling (p = 4). We now consider agnostic tests of GR for a coupling constant with dimensions (mass)^4, i.e., $p = 4$ in Eq. (3). In Fig. 10 we show the joint posterior credible intervals for the nonspinning ParSpec parameters $(\delta\omega_{22}^{(0)}, \delta\tau_{22}^{(0)})$, computed by assuming $p = 4$ in Eqs. (1) and (2). The results are rather insensitive to the
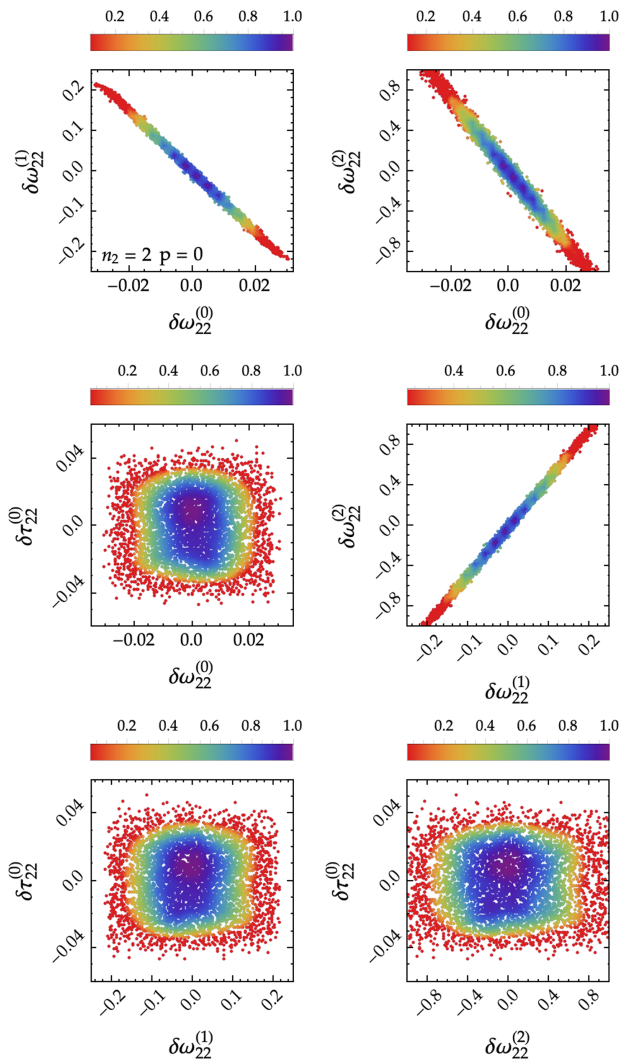
FIG. 9. Joint posterior samples for the ParSpec parameters in the agnostic analysis using strategy (i), including quadratic spin corrections ($n_2 = 2$) and dimensionless coupling ($p = 0$). The color scheme identifies the probability of the samples, with red (blue) regions corresponding to low (high) probabilities. We stack the full set of events observable by ET and assume the spin distribution of model I, but results for model II are qualitatively similar.

chosen spin distribution model, and they mostly depend on the mass distribution of the remnant. Since the QNM corrections are (to a first approximation) proportional to the coupling (3), which scales as $\sim M_f^{-4}$, they are strongly suppressed for heavy BHs, so the constraints shown in Fig. 10 are dominated by the ~ 400 (~ 200) ET (CE) events having $M_f \lesssim 50 M_\odot$. As a rough estimate, our results suggest that ET and CE may set bounds on the fundamental coupling of a putative beyond-GR gravity theory of the order $\alpha \sim [\delta\omega_{22}^{(0)}]^{1/4} \sim \mathcal{O}(10 \text{ km})$, at 90% credible level.

We have also studied how the constraints are affected by the inclusion of spin-dependent terms in the ParSpec

template. We find that the bounds on the spin-dependent corrections are not informative with respect to the prior, and that the credible intervals on the nonspinning terms $(\delta\omega_{22}^{(0)}, \delta\tau_{22}^{(0)})$ deteriorate by about an order of magnitude, possibly violating the small-coupling assumption implicit in the ParSpec approach.

2. Strategy (ii): Mass and spin are unknown

No spin-dependent corrections ($n_2 = 0$), dimensionless coupling ($p = 0$). So far we have investigated the ability of XG detectors to constrain beyond-Kerr deviations, assuming that the remnant BH mass and spin are known (e.g., because they can be inferred from the full inspiral-merger-ringdown waveform). This procedure [that we called "strategy (i)" in Sec. II A] drastically improves the bounds, because the observation of a single QNM—i.e., the dominant (22) mode—is sufficient to perform BH spectroscopy tests. We now turn to the more conservative and pessimistic approach [called "strategy (ii)" in Sec. II A] in which we use information from the ringdown only. We now need at least two modes in order to simultaneously constrain the ParSpec parameters, as well as the final BH mass and spin of each observed ringdown event.

For reasons that will be apparent soon, we perform the analysis only for dimensionless couplings ($p = 0$) and for nonrotating corrections ($n_2 = 0$). We analyze the (22) mode in combination with either a (33) or a (21) secondary QNM.

In Fig. 11 we show the posterior distributions of the ParSpec parameters as functions of the observed events for models I and II, respectively. Different histograms correspond to posteriors inferred by stacking a different number of observations. In the first two rows of Fig. 11 we assume that M_f and χ_f are fixed to their true values. While in this case the posteriors of $\delta\omega_{22}^{(0)}$ and $\delta\tau_{22}^{(0)}$ are correctly centered around zero, and become narrower as N increases, the panels in the third and fourth rows, where M_f and χ_f are allowed to vary, change dramatically, and we observe a significant bias in all parameters. In the majority of cases, the GR null hypothesis is excluded at more than 90% of the marginalized posterior distributions. Both the fundamental and the secondary mode are affected: their probability distribution can peak around the wrong value and rally against the prior. These results are independent of the choice of the secondary mode.

The systematic bias builds up due to correlations with M_f and χ_f, which are partially degenerate with the ParSpec parameters, and it can become more significant as N grows. Note that for practical purposes here we consider only 50 events: this already requires more than 100 parameters—i.e., 2×50 masses and spins—to be sampled along with the four ParSpec parameters for the fundamental and secondary QNM.

TABLE II. We list the 90% credible intervals for the marginalized distribution of the ParSpec parameters with quadratic spin corrections ($n_2 = 2$) and dimensionless coupling ($p = 0$). We stack a different number N of sources for the two models. Dotted entries correspond to parameters for which the posterior is not informative with respect to the prior.

		Model I All	Model II All	Model I $N = 1000$	Model II $N = 1000$	Model I $N = 500$	Model II $N = 500$
CE-40 km	$\delta\omega_{22}^{(0)}$	$[-0.0234, 0.0237]$	$[-0.0341, 0.0363]$	$[-0.0282, 0.0273]$	$[-0.0485, 0.0509]$	$[-0.0312, 0.0312]$	$[-0.0602, 0.0586]$
	$\delta\omega_{22}^{(1)}$	$[-0.178, 0.175]$	$[-0.222, 0.213]$	$[-0.193, 0.196]$	$[-0.257, 0.246]$	$[-0.204, 0.205]$	$[-0.276, 0.277]$
	$\delta\omega_{22}^{(2)}$	$[-0.853, 0.834]$	$[-0.978, 0.780]$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
	$\delta\tau_{22}^{(0)}$	$[-0.0343, 0.0337]$	$[-0.0667, 0.0709]$	$[-0.0534, 0.0529]$	$[-0.210, 0.201]$	$[-0.0725, 0.0746]$	$[-0.279, 0.277]$
ET-15 km	$\delta\omega_{22}^{(0)}$	$[-0.0217, 0.0225]$	$[-0.034, 0.0366]$	$[-0.0266, 0.0281]$	$[-0.0448, 0.0464]$	$[-0.0311, 0.0316]$	$[-0.0541, 0.057]$
	$\delta\omega_{22}^{(1)}$	$[-0.169, 0.161]$	$[-0.229, 0.212]$	$[-0.197, 0.188]$	$[-0.255, 0.243]$	$[-0.204, 0.203]$	$[-0.273, 0.258]$
	$\delta\omega_{22}^{(2)}$	$[-0.807, 0.771]$	$[-1.00, 0.741]$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
	$\delta\tau_{22}^{(0)}$	$[-0.0323, 0.0324]$	$[-0.0561, 0.0569]$	$[-0.0526, 0.0508]$	$[-0.169, 0.159]$	$[-0.0751, 0.0742]$	$[-0.25, 0.249]$

In order to clarify the impact of the remnant BH properties on the reconstruction of $\delta\omega_{\ell m}^{(0)}$ and $\delta\tau_{\ell m}^{(0)}$, we repeat the analysis, but this time we replace the "true" mass and spin by "mock" values that are shifted by some fixed amount:

$$\overline{M}_f = M_f + \Delta M_f, \qquad \bar{\chi}_f = \chi_f + \Delta\chi_f. \quad (43)$$

For simplicity, we assume $\Delta M_f/M_f = \Delta\chi_f/\chi_f = \epsilon$. We then sample over the ParSpec parameters only.

In Fig. 12 we show the posterior distributions inferred with the (22)–(33) multimode analysis for ET and model I. We stack $N = 50$ events and we consider values of ϵ ranging between 0.001 and 0.04. The probability densities drawn in the four panels demonstrate how the bias evolves as a function of the relative shift in mass and spin. An accuracy of $\sim 1\%$ in the masses and spins seems to be required in order for the GR limit to return within the posterior support.
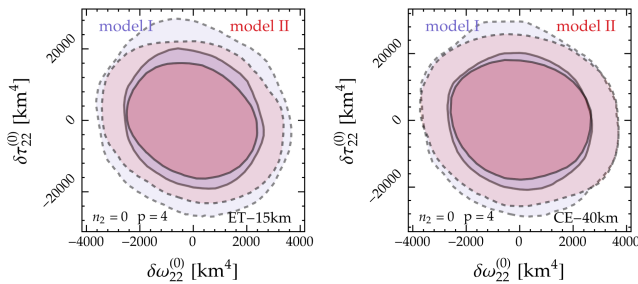

FIG. 10. Two-dimensional posterior distribution of the spin-independent ParSpec parameters $(\delta\omega_{22}^{(0)}, \delta\tau_{22}^{(0)})$, assuming that Kerr deviations are controlled by a dimensionful coupling with $p = 4$. We perform an agnostic analysis using strategy (i). Solid and dashed lines identify the 68% and 90% credible intervals, while light-purple and red regions correspond to different population models.

The masses and spins of the remnant inferred from XG ringdown observations alone are expected to have $\sim 10\%$ accuracy [95], about an order of magnitude worse than the $\sim 1\%$ accuracy estimated above. However, the uncertainty in the remnant's properties could be improved using different data analysis techniques (for example, by using the full inspiral-merger-ringdown information).

The origin of this bias can be traced back to the degeneracy between (M_f, χ_f) and the fundamental mode frequency and damping time [96]. In the case of the ParSpec expansion, for small changes ΔM_f, $\Delta\chi_f$, this degeneracy leads to a shift in the beyond-Kerr parameters of the following form:

$$\delta\omega_{\ell_j m_j}^{(k_2)} \to \delta\omega_{\ell_j m_j}^{(k_2)} - \Delta M_f + k_2\Delta\chi_f, \quad (44)$$

$$\delta\tau_{\ell_j m_j}^{(k_2)} \to \delta\tau_{\ell_j m_j}^{(k_2)} + \Delta M_f + k_2\Delta\chi_f, \quad (45)$$

Equations (44) and (45) suggest that, to avoid large systematics in the inference of $\delta\omega_{\ell_j m_j}^{(k_2)}$ and $\delta\tau_{\ell_j m_j}^{(k_2)}$, the deviations in the mass and spin should be small enough to satisfy the condition

$$\left|\delta\omega_{\ell_j m_j}^{(k_2)}\right| \gg |\Delta M_f - k_2\Delta\chi_f|, \quad (46)$$

$$\left|\delta\tau_{\ell_j m_j}^{(k_2)}\right| \gg |\Delta M_f + k_2\Delta\chi_f|. \quad (47)$$

Note that the two-mode analysis discussed above focuses on the $p = 0$ case. In general, we expect the bias to be worse for dimensionful corrections (e.g., $p = 4$). This is because the mismodeling of the mass would also affect the coupling parameter γ_i, which also controls deviations from the Kerr spectrum. For this reason we do not show results
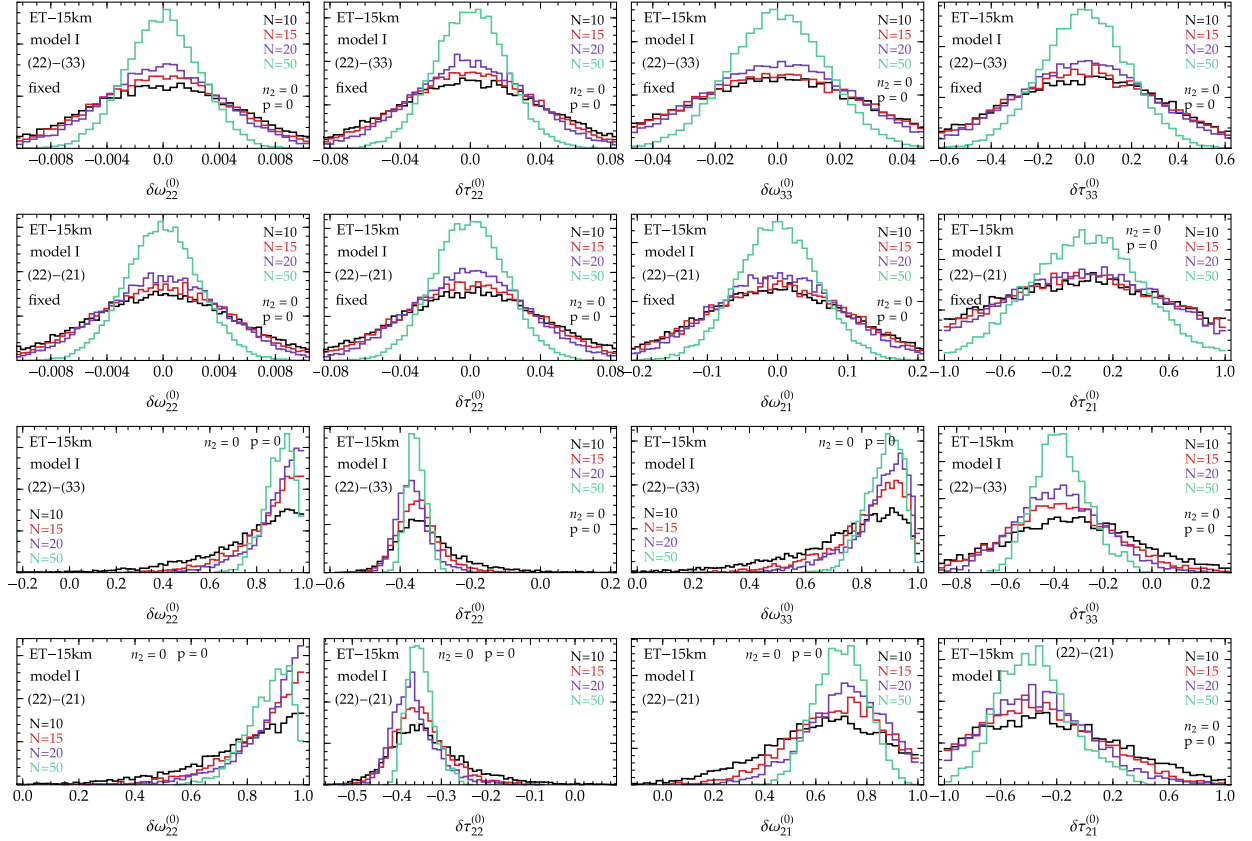
FIG. 11. Distribution of the ParSpec deviation parameters as a function of the number of observed events in the agnostic analysis. We assume a two-mode template that includes the (22) and either the (33) or the (21) QNM, and we use the model I population. The top two rows (labeled by "fixed") correspond to posteriors obtained by keeping the BH masses and spins fixed to their injected values in the MCMC simulations. The bottom rows are derived by sampling the ParSpec parameters as well as the mass and spin of the remnant.

for the case $p = 4$, and we do not include spin-dependent corrections ($n_2 > 0$) in this part of the analysis.

B. Theory-specific inference

We first focus on the quadratic theories of gravity discussed in Sec. III B. Both theories (EsGB and dCS) are characterized by a dimensionful coupling constant, bounded by astrophysical observations either in the gravitational or electromagnetic band. The maximum values currently allowed for the parameters $\alpha_{\rm GB,CS}$ are $\sqrt{\alpha_{\rm GB}} \simeq$ 6.3 km and $\sqrt{\alpha_{\rm CS}} \simeq 8.5$ km [97] (due to different normalizations, the EsGB coupling constant in this work is larger than the one defined in Ref. [97] by a multiplicative factor of $4\pi^{1/4}$). Note however for such values of the couplings, deviations from GR can become nonperturbative, with $\beta_{\rm GB,CS} = \alpha_{\rm GB,CS}/m_{1,2}^2 \sim 1$ for low-mass BHs of our catalogs. This is problematic because both EsGB and dCS gravity should be seen as effective field theory, valid in the limit $\beta_{\rm GB,CS} \ll 1$. To be consistent with the perturbative character of ParSpec, in this work we inject signals for both EsGB and dCS gravity assuming a coupling such that $\beta_{\rm GB,CS} = \alpha_{\rm GB,CS}/\min[m_1, m_2]^2 = 0.1$, with $\min[m_1, m_2]$

being the smallest mass of the binary BHs within our population models that yield an observable ringdown signal. For both models I and II, this leads to $\sqrt{\alpha_{\rm GB,CS}} \sim 2.5$ km.

Before assessing the capability of XG detectors to constrain the coupling parameter, it is instructive to estimate the order of magnitude of the expected deviations from the Kerr predictions for the frequencies and damping times. In Fig. 13 we show the cumulative distributions of the changes in the (22) mode predicted by EsGB gravity (top row) and dCS gravity (bottom row) relative to the (slowly rotating) Kerr predictions. These corrections are computed using the "beyond GR" part of the sum in Eqs. (1) and (2), truncated at the order in the slow-rotation expansion at which the corrections are currently known: second order for EsGB and first order for dCS (see Appendix A). The cumulative distribution function is computed by considering all of the events that have an observable ringdown signal. For completeness, in Figs. 28 and 30 of Appendix A we compare the relative changes in the dominant mode to the relative changes in the subdominant (33) and (21) mode frequencies.

For EsGB gravity, all systems lead to changes in the frequency smaller than ~0.01%, regardless of the adopted
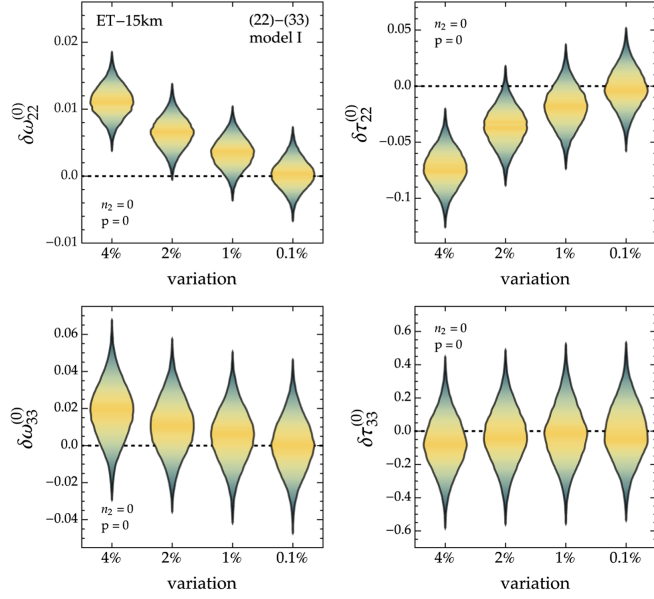
FIG. 12. Posterior density distributions of the ParSpec parameters for the $p = 0$ case, inferred in the agnostic analysis with strategy (ii), assuming that the BH masses and spins, fixed within the MCMC simulation, are shifted by a fixed amount (indicated on the x axis) with respect to their "true" values. Colors closer to yellow (green) correspond to higher (lower) probability. Results are obtained by stacking $N = 50$ events of the model I catalog.

spin distribution model. Events observed by CE have slightly smaller corrections compared to ET, because the catalog observed by CE includes a smaller number of light BHs. The maximum expected shifts in the damping times are of the same order of magnitude as the frequency shifts.

For dCS gravity, deviations with respect to Kerr are generally larger by one order of magnitude. Interestingly, the shifts in the damping time are bigger than the frequency shifts.

To understand how these estimates translate into actual constraints on the coupling constants, we first run our MCMC pipeline assuming that the mass and spin of the remnants are fixed to their injected values. We consider a single-mode and a two-mode analysis. In the two-mode analysis we only use the frequency (and not the damping time) of the secondary QNM. The main reason behind this choice is that the damping time of the secondary mode is never well measured.

1. Einstein-scalar-Gauss-Bonnet gravity

In Fig. 14 we plot the posterior of the EsGB coupling constant α_{GB} inferred by stacking the entire set of events observed by ET and CE. Unshaded and shaded curves correspond to the (22) and to the (22)–(33) templates, respectively. In both cases our results do no show any particular preference for the injected value of the coupling: the injected value does fall within the 90% credible intervals of the distributions, but the posteriors are fully
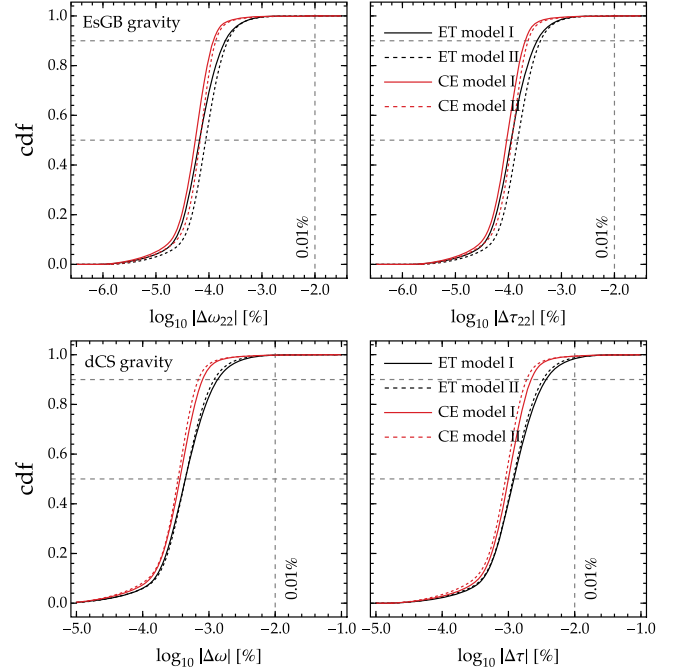

FIG. 13. Cumulative distribution for the changes in the frequencies and damping times of the (22) mode with respect to their Kerr values predicted by EsGB gravity (top row) and dCS gravity (bottom row), for detectable events. QNMs beyond GR are computed by assuming coupling constants $\alpha_{GB,CS}$ that saturate the current upper bounds from astrophysical observations [97]. The vertical lines mark 0.01% changes in frequency and damping time, while the horizontal lines mark the 50% and 90% confidence intervals of the distribution.
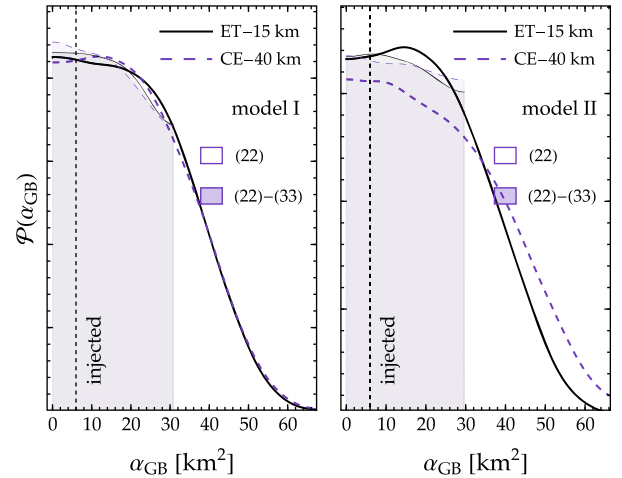

FIG. 14. Posterior distributions of the EsGB coupling constant α_{GB}. Bounds on α_{GB} are obtained by stacking the whole set of observations for ET and CE, and by fixing both the masses and the spins of the BH remnants. The vertical dashed lines mark the injected value of the coupling. Unshaded (shaded) regions are the bounds from a one-mode (two-mode) analysis, respectively.

consistent with the GR hypothesis, and only allow to infer an upper bound on α_{GB}. As expected, ET provides tighter constraints, mainly due to the larger number of observable ringdown events. The posteriors in the (22)–(33) case have a sharp cutoff around $\alpha_{GB} \sim 30$ km^2. This cutoff is imposed by the requirement that the ParSpec parameters must be perturbative in nature, i.e., $|\gamma_i \delta\omega_{\ell m}^{(k_2)}| \lesssim 1$ and $|\gamma_i \delta\tau_{\ell m}^{(k_2)}| \lesssim 1$ (see also the discussion in Appendix 28). This choice imposes a tight upper bound on the frequency of the (33) mode, because the coefficients of the dCS expansion (as shown in Appendix A) are such that deviations from Kerr are larger as the BH mass decreases. We have also run our Bayesian pipeline assuming the (22)–(21) combination (not shown here): in this case the results are very close to the single-mode analysis, confirming that most of the information comes from the dominant (22) multipole. The numerical values of the ParSpec coefficients for the (21) QNM make it possible to sample a wider range of values for α_{GB}, so the (22)–(21) posteriors do not have the sharp cutoff observed in the (22)–(33) case.

Our results are only mildly dependent on the chosen model for the binary spins. Focusing on the single-mode strategy, the 90% credible interval derived using ET is $\alpha_{GB} \sim [0, 40]$ km^2 and $\alpha_{GB} \sim [0, 41]$ km^2 for models I and II, respectively. The 90% credible interval derived using CE is $\alpha_{GB} \sim [0, 41]$ km^2 and $\alpha_{GB} \sim [0, 47]$ km^2 for models I and II, respectively. The top panel of Fig. 15 summarizes the 90% credible intervals on the EsGB coupling for different detector configurations and spin models. All constraints are close to the current upper bound for the EsGB coupling. Given the mass dependence of the EsGB corrections [see Eq. (42)] we expect our results
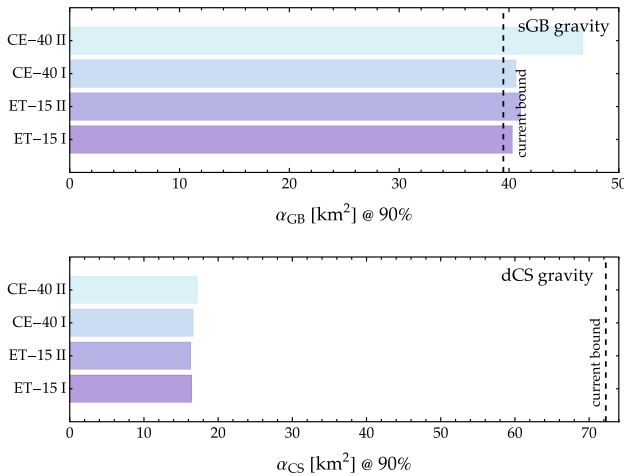

FIG. 15. Ninety percent credible intervals around the median for the coupling constants of EsGB and dCS gravity inferred using different combinations of detectors and population models. The bounds are obtained by stacking all events in a single-mode analysis, and assuming that the masses and spins of the remnants are fixed to their injected values. Vertical lines correspond to the current upper bound for $\alpha_{GB,CS}$.
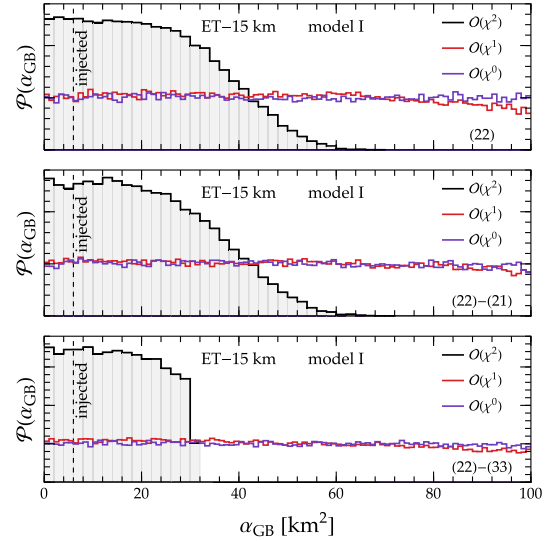

FIG. 16. Posterior distribution of the EsGB coupling constant inferred from ET observations with a single-mode analysis. Samples are obtained by injecting QNM signals in EsGB gravity that include second-order corrections in the spin. The black histogram corresponds to recovery templates which include $\mathcal{O}(\chi^2)$ spin terms in the beyond-Kerr parametrization. Red and purple histograms refer to a template that includes only non-rotating terms or first-order spin corrections, respectively.

to be dominated by low-mass systems. As a check, we have repeated the analysis by considering only BH remnants with $M_f \lesssim 50 M_\odot$, and the results are very similar to those shown in Fig. 14.

Spin-dependent corrections to the QNM frequencies in EsGB gravity are known up to quadratic order in the spin. In Fig. 16 we investigate how these corrections affect the posteriors on α_{GB}. For simplicity we focus on a single-mode analysis of ET observations. We show the posterior probabilities obtained by fixing $n_2 = 0$ (purple), $n_2 = 1$ (red), and $n_2 = 2$ (black) in the recovering templates of Eqs. (1) and (2), while the injected QNM signals always include all corrections in the spin (up to and including the second-order terms). Including only first-order or zeroth-order spin corrections in the template flattens the posterior distribution, which becomes uninformative with respect to the prior. Results for the two-mode analysis, assuming either the (33) or the (21) component as the secondary QNM, are qualitatively similar. These results suggest that spin corrections play a dominant role in the Bayesian inference. This can be understood by looking at the magnitude of quadratic terms in the frequency expansion, which are about one order of magnitude larger than the zeroth-order term for $\chi_f \sim 0.7$ (see Appendix A, and in particular Fig. 26). Therefore pushing the spin expansion to higher orders is very important to perform tests that are not dominated by systematic errors.

To further clarify the role of rotational corrections, we repeat the analysis by assuming that both the injected signal
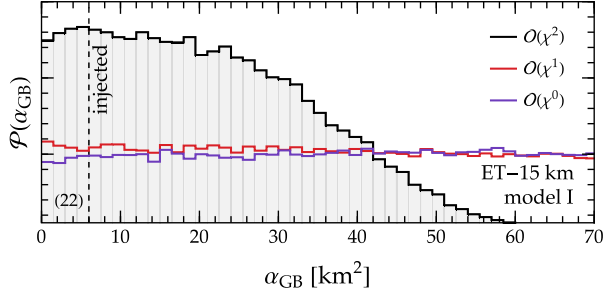
FIG. 17. Posterior of the EsGB coupling constant for ET observations of BHs with $M_f \leq 50 M_\odot$. Black histograms are obtained by assuming that both the injected frequencies and the recovery template include quadratic beyond-GR corrections in the spin. Red and purple histograms are obtained by MCMC simulations in which both the injected frequencies and recovery templates contain only nonrotating terms or only first-order spin corrections, respectively.
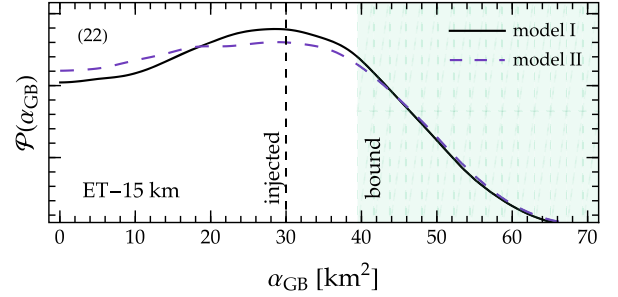

FIG. 18. Posterior distributions of the EsGB coupling parameter found by stacking all events observed by ET with a single-mode strategy. The injected signals assume a value of the coupling $\alpha_{\rm GB}/\min[m_1, m_2]^2 = 0.5$ that violates the perturbative assumption. The shaded region identifies the parameter space ruled out by astrophysical observations.

and the recovery template contain either nonrotating or linear-in-spin corrections beyond-Kerr corrections. We basically follow the procedure described in Sec. III C, but now we compute FIM errors by omitting higher-order in spin corrections to the EsGB frequencies and damping times. The results are shown in Fig. 17 and support our previous analysis. When we neglect quadratic terms in the spin the posterior distribution of the coupling is fully consistent with the uniform prior, suggesting that $\mathcal{O}(\chi^2)$ terms dominate the reconstruction of $\alpha_{\rm GB}$. It is possible, and even likely, that bounds inferred by truncating at quadratic order are affected by a bias due to ignoring higher-order coefficients in the QNM expansion, which are currently unknown. We remark that, even in the most optimistic case, the ringdown-only EsGB upper bounds of Fig. 14 are weaker than the bounds expected from XG observations of binary inspirals, which could be as small as $\alpha_{\rm GB} \sim 10^{-3}$ km^2 in the absence of systematic errors (see, e.g., Fig. 20 of Ref. [5]).

It is interesting to investigate what constraints can be placed on the EsGB coupling if we allow $\alpha_{\rm GB}$ to be nonperturbatively large. In Fig. 18 we show the posteriors of $\alpha_{\rm GB}$ found with ET by injecting signals such that $\beta_{\rm GB} = \alpha_{\rm GB}/\min[m_1, m_2]^2 = 0.5$, which yields $\alpha_{\rm GB} \simeq 30$ km^2. As expected, the detectability of beyond-Kerr deviations improves significantly: the distributions now have a clear peak around the injected value of the coupling. Note however that, even in this extreme case, the posteriors also support the GR hypothesis. Similar conclusions apply also to CE.

It is worth asking how the bounds would change if we allowed the BH masses and spins to be sampled together with $\alpha_{\rm GB}$. To limit the computational cost coming from the inclusion of a larger number of parameters in the MCMC simulations, we restrict our analysis to ringdown events such that (i) the remnants have masses $M_f \leq 50 M_\odot$ and

(ii) the relative error on the damping time of the secondary QNM is $\sigma_{\omega_{33}}/\omega_{33} \leq 50\%$ [and similarly for the (21) component]. For concreteness we focus on ET. Within all observable events in model I, we find 100 and 69 events which satisfy the two requirements above for the (22)–(33) and (22)–(21) templates, respectively (for model II, there are 92 and 50 events satisfying these conditions). Among these events, we select the 50 lightest-mass BHs, which are expected to yield the largest deviations from the Kerr spectrum.

In Fig. 19 we focus for simplicity on model II and on the (22)–(21) combination (other configurations lead to similar results). As expected, the ability of XG detectors to constrain $\alpha_{\rm GB}$ worsens dramatically. For comparison we also plot $\mathcal{P}(\alpha_{\rm GB})$ when M and χ are kept fixed. When the masses and spins are sampled together with $\alpha_{\rm GB}$ the posteriors become less informative. However the MCMC can still correctly infer the injected values of (M_f, χ_f), and
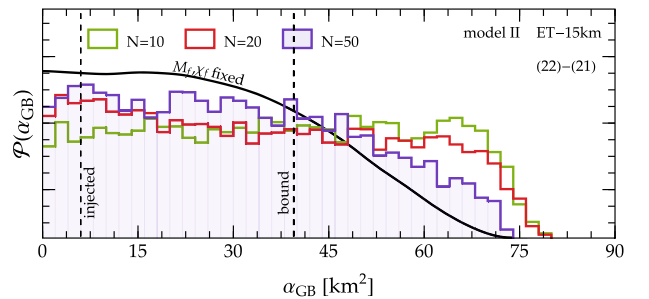

FIG. 19. Posterior histograms for the GB coupling parameter inferred by stacking different sets of events observed by ET. Here we use a two-mode analysis in which the mass and spin of the remnant are sampled together with $\alpha_{\rm GB}$. The vertical dashed lines identifies the injected value of the coupling, and the current upper bound. For sake of comparison we show as solid black curve the posterior distribution obtained by fixing mass and spin to their injected values, and stacking $N = 50$ events. The normalization constant of such distribution has been scaled for visualization purposes.
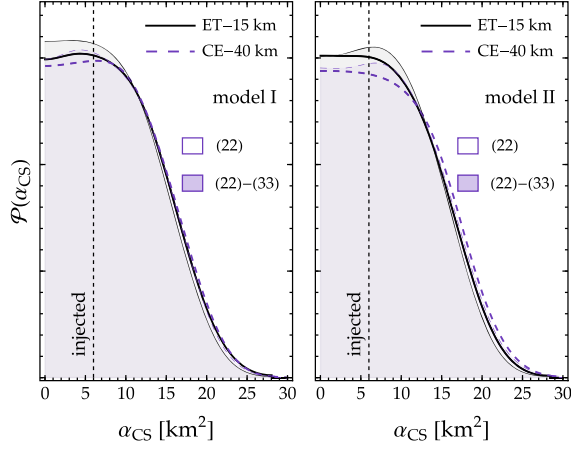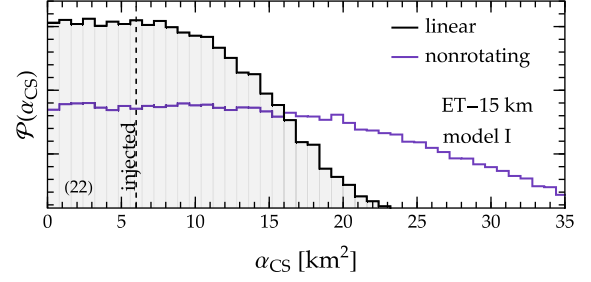
FIG. 20. Same as Fig. 14, but for dCS gravity.

FIG. 21. Posteriors of the dCS coupling constant inferred from ET observations by injecting single-mode QNM signals. Black histograms correspond to recovery templates which include $\mathcal{O}(\chi)$ spin terms, while purple histograms refer to templates that include only nonrotating terms in the beyond-Kerr parametrization.

2. Dynamical Chern-Simons gravity

The constraining power of XG detectors improves only mildly for dCS gravity compared to EsGB, as shown in the two panels of Fig. 20. Even in this case the posteriors of α_{CS}, obtained by stacking the full sets of BH remnants and keeping first-order spin corrections (the only corrections available so far), show full support for the GR value.

From a single-mode analysis based on the (22) dominant mode we infer that ET may be able to bound the dCS coupling to the range $\alpha_{CS} \sim [0, 16.4]$ km^2 and $\alpha_{CS} \sim [0, 16.3]$ km^2 (90% confidence level) for models I and II, respectively. For CE, the forecast values have a slightly larger spread: $\alpha_{CS} \sim [0, 16.6]$ km^2 and $\alpha_{CS} \sim [0, 17.2]$ km^2 for models I and II, respectively. For completeness, we have repeated the analysis with the (22)–(21) combination of QNMs. We find that the inferred bounds are only mildly dependent on the choice of the secondary mode.

A comparison of the 90% credible intervals on α_{CS} for different detector configurations and spin population models is shown in the bottom panels of Fig. 15. Unlike EsGB gravity, forecasts from both detectors are tighter than the current upper bound by a factor $\sim$5. As in the case of EsGB, the constraints on α_{CS} are dominated by observations of low-mass sources with $M_f \leq 50M_\odot$.

We have further assessed the relevance of spin corrections by studying how bounds on the dCS coupling change if we adopt a recovery template that only includes nonrotating coefficients in the beyond Kerr expansion. The results are shown in Fig. 21 for ET observations and model I. At variance with EsGB gravity, spin corrections seem to play a less dominant role for dCS: the posteriors obtained using zeroth-order terms in χ are roughly a factor of 2 broader, and still informative over the prior. Results for other detector configurations and BH populations are similar. While the analysis is limited by our current knowledge of higher-order spin terms in the dCS QNM spectrum, our findings seem to suggest that quadratic and higher-order corrections in χ may not dramatically affect our conclusions.

In EsGB gravity, including masses and spins within the sampling parameters yields uninformative posteriors on the coupling constants. The results for dCS shown in the four panels of Fig. 22 are more promising. Here we follow the same strategy discussed in Sec. IV B 1 in the context of EsGB gravity, and we focus on BHs with masses smaller than $50M_\odot$. Green, red, and purple histograms are the marginalized distributions of α_{CS}, obtained by stacking an ever larger set of events observed by ET. For comparison we also show (in black) the posterior obtained by fixing M_f and χ_f to their true values and assuming $N = 50$ events: this posterior is consistent with zero, but it usually peaks around the injected value of the dCS coupling (marked by a vertical dashed line).

The inferred value of α_{CS} is quite sensitive to the spin distribution of the population and only mildly affected by choice of the secondary mode, even if dCS corrections to the (33) Kerr frequencies and damping times are always one order of magnitude larger than the corresponding corrections to the (21) component (compare Fig. 30). This confirms that the bulk of information comes from the fundamental QNM, as we discussed earlier in the context of Fig. 20. For the model I population (left panels of Fig. 22) the posterior distributions shift to the wrong value of α_{CS}, peaking far from the injected value of the coupling as the number of observations grows. The origin of this systematic bias can be traced to the degeneracy between the source parameters (i.e., M_f, χ_f) and the beyond-GR deviation parameter (in this case, the dCS coupling) that we also observe in the agnostic two-mode analysis (cf. Sec. IV A 2). Indeed, the bias in Fig. 22 is slightly larger when we consider the (21) mode (which leads to smaller changes in the Kerr spectrum) as the secondary mode.

For the model II population (right panels of Fig. 22) the bias is less evident. In this case, the posterior become wider
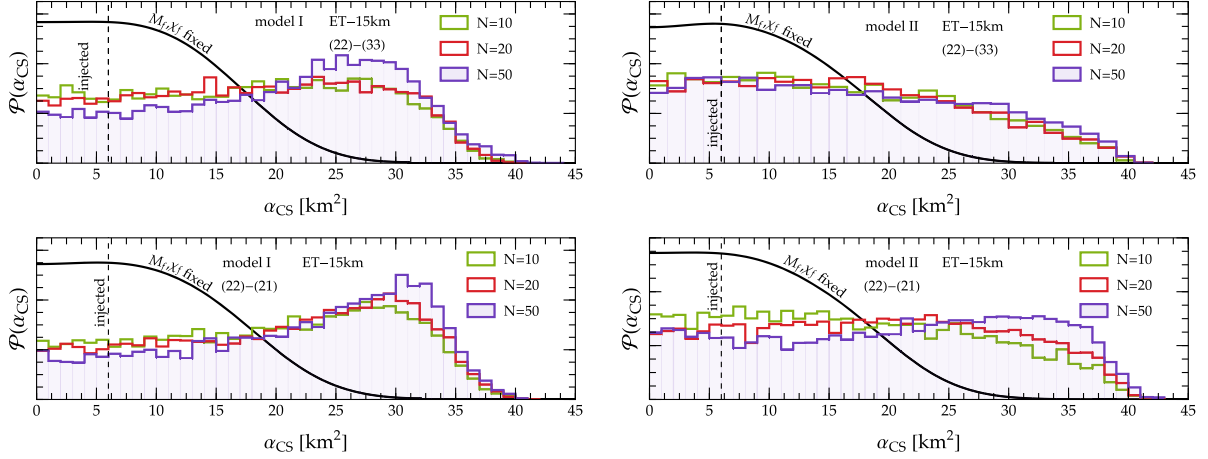
FIG. 22. Posterior distributions of the dCS coupling constant α_{CS} obtained by stacking different sets of events observed by ET in a two-mode analysis. The vertical dashed lines identifies the injected value of the coupling. The mass and spin of the remnants are sampled together with α_{CS}. For comparison we also show the posterior of the coupling obtained by stacking the largest set of events considered, and fixing M_f and χ_f to their true values, with $N = 50$ (solid black line). The left and right panels refer to different spin population models.

that those obtained by fixing the BH masses and spins, with the upper bound on α_{CS} being a factor two bigger. As for EsGB gravity, the remnant mass and spin posteriors are always correctly centered around the "true" injected values: see Fig. 32 in Appendix B. It is also clear from Fig. 32 that the mass and spin inference is generally better for large-SNR systems. These are usually large-mass systems: unfortunately, the degeneracy between mass, spin, and coupling constants is less relevant precisely for those systems for which beyond-GR corrections are suppressed.

3. Effective field theory

We now discuss how well ringdown observations by XG detectors could constrain the EFT gravity models of Ref. [66]. Once again, we first assume that BH masses and spins are fixed to their injected values, and only sample over the coupling constant α_q for each of the eight EFT models we consider. We inject signals with three different values of α_q such that $\gamma_q^{cubic} = \alpha_q^4/\min[m_1, m_2]^4 = (0.1, 0.2, 0.5)$ and $\gamma_q^{quartic} = \alpha_q^6/\min[m_1, m_2]^6 = (0.1, 0.2, 0.5)$, where $\min[m_1, m_2]$ is the smallest mass among all binary components that yield an observable ringdown signal within each of our population models.

The posterior distributions for both cubic and quartic couplings are shown in Fig. 23, where we stack all BH remnants observed by ET and we assume the model I spin distribution. Each panel shows results obtained by either a single-mode or a two-mode analysis. As we observed for EsGB gravity, the hard cut in the posteriors inferred from the (22)–(33) combination for some of the models corresponds to imposing the requirement that the ParSpec coefficients remain in the perturbative regime.

All of the EFT gravity theories we analyze display a common trend: for $\gamma_q \leq 0.2$ (first and second column) the inferred posteriors are poorly informative, yielding upper bounds on α_q of the order of $\mathcal{O}(10)$ km, which depend only mildly on the mass scaling of the coupling (either $\gamma_q \propto M^{-4}$ or $\gamma_q \propto M^{-6}$) and on the specific gravity theory. All distributions rail against zero, showing support for the GR hypothesis.

At the highest value of γ_q we consider (third column in Fig. 23), the posteriors for some of the models narrow, peaking around the injected value of α_q. Note however that all posterior distributions still rail against the GR value within the 90% credible interval. These results imply that ringdown observations may be able to mildly constrain the coupling constant of these EFT gravity models only when the coupling constants are very large, possibly violating the domain of validity of the theory. For signals with $\gamma_q^{cubic} = 0.5$ (top-right panels) we only show results for the single-mode inference: in the two-mode inference, the coupling (which is outside the EFT domain of validity) is such that we always find a BH mass in our catalog that violates the perturbative character of the ParSpec corrections. As usual, including the mass and spins of the remnant within the MCMC sampling introduces additional correlations among the parameters, thus widening the posteriors.

The high-order expansion of the QNM frequencies derived in Ref. [66] allows us to test the relevance of different contributions in the spin expansion to the QNM spectrum. To this end, in Fig. 24 we repeat the previous analysis with a ParSpec template that includes spin terms only up to a given order. For concreteness we show the constraints on the coupling parameter of the 1^+ model, and we inject signals corresponding to large coupling
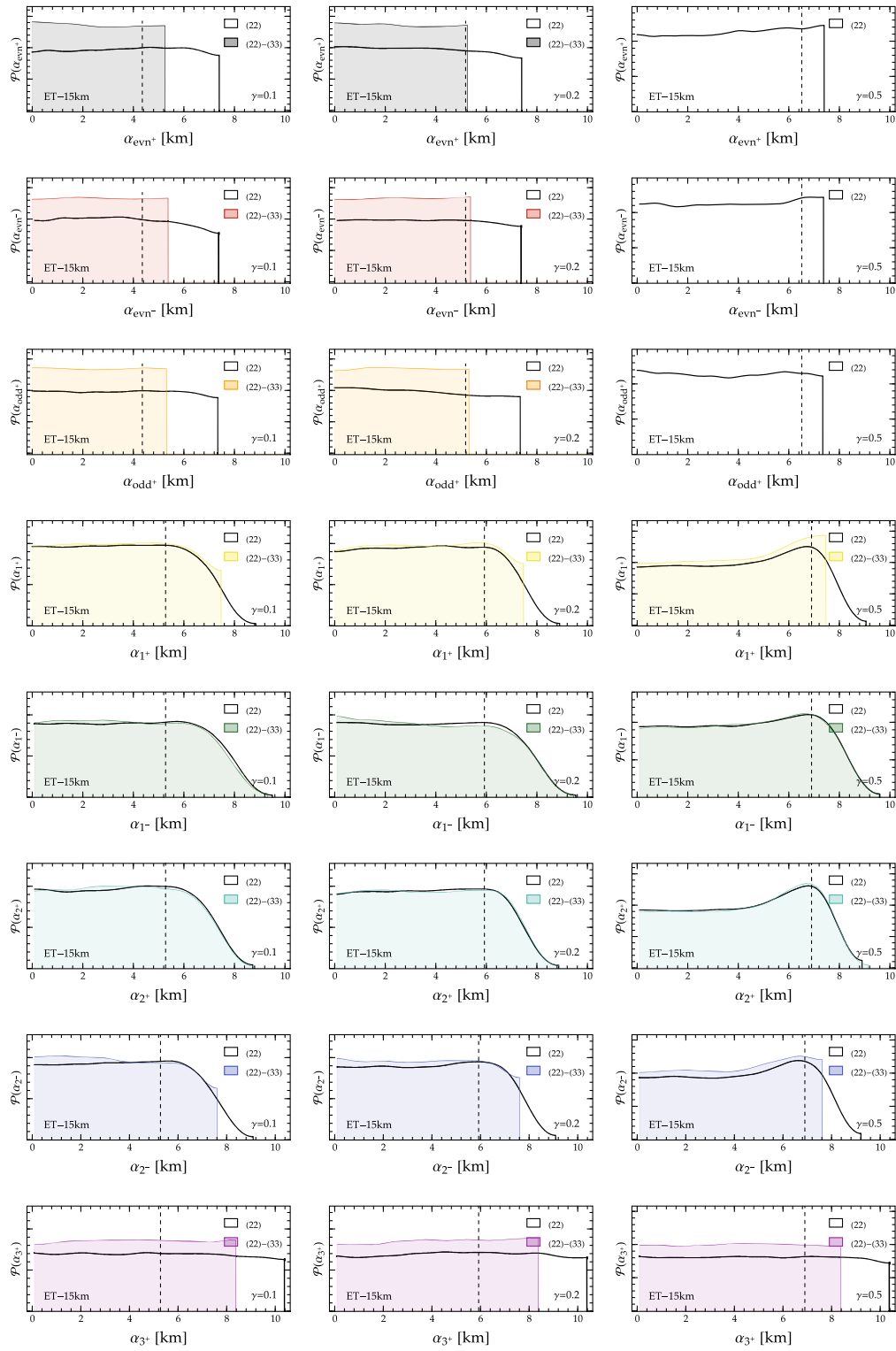
FIG. 23. Posterior probability distributions for the coupling parameters of the eight EFT models we consider. Bounds are obtained by stacking all of the model I signals observed by ET. Left, middle, and right panels refer to a different injected value of α_q, corresponding to $\gamma_q = (0.1, 0.2, 0.5)$ in each specific theory. The injected values are marked by vertical lines.
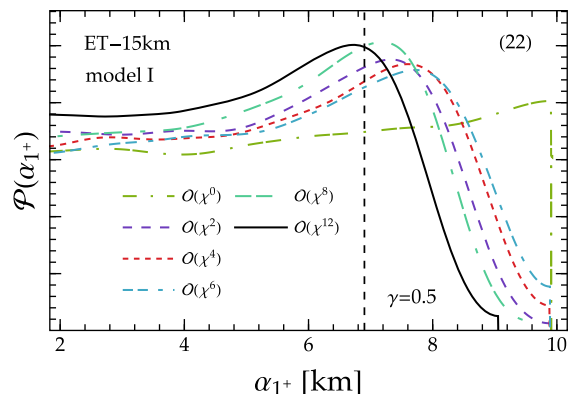
FIG. 24. Posterior distribution of the EFT coupling parameter α_{1^+}. Different colors correspond to recovery templates that include spin effects at different orders, as indicated in the legend. The vertical dashed line marks the injected value of the coupling.

$(\alpha_{1^+}^6 / \min[m_1, m_2]^6 = 0.5)$. While a zero-spin template introduces a significant bias in the posterior distribution, a sixth-order expansion in χ is sufficient to recover the "true" injected value of α_{1^+}. Note that the quadratic expansion is remarkably good in this case. However the convergence of the series may well be asymptotic, and this is probably just a coincidence. The results for the other families of EFT gravity models are similar: depending on the specific model, including between the sixth and the ninth order is typically sufficient to infer the injected coupling parameter without significant bias.

V. CONCLUSIONS

In this paper we employ the ParSpec parametrization of the QNM spectrum of spinning BHs beyond GR developed by some of us in previous work [67] to address the following question: can future ringdown observation with XG detectors detect (or constrain) deviations from the Kerr spectrum by stacking multiple observations of binary mergers? We use astrophysically motivated populations and address this question for two different families of spectroscopy tests: (i) agnostic (null) tests and (ii) theory-based tests, which make use of QNM frequency calculations in EsGB gravity, dCS gravity, and various EFT-based extensions of general relativity. When we inject beyond-GR signals in the data, we generally assume the coupling constant of specific beyond-GR theory to saturate current experimental bounds.

We find that robust inference of hypothetical corrections to GR requires pushing the slow-rotation expansion to high orders. Even when high-order expansions are available, ringdown observations alone may not be sufficient to measure deviations from the Kerr spectrum for theories with dimensionful coupling constants because the constraints are dominated by "light" black hole remnants, and only few of these light remnants have sufficiently high SNR in the ringdown.

For theory-agnostic tests with dimensionless couplings ($p = 0$) that include spin corrections to the QNM spectrum we find a hierarchy among the beyond-GR parameters, with the nonrotating terms providing the tightest constraints. This confirms some of the main results of Ref. [67]: spin corrections to the damping time may be unmeasurable, even when stacking thousands of events.

For theory-agnostic tests where couplings have dimensions of $(\text{mass})^4$ ($p = 4$), we can only constrain nonrotating corrections to the QNM spectrum. The inferred bounds on the ParSpec parameters imply bounds on the fundamental coupling of a putative beyond-GR theory of the order of $\mathcal{O}(10)$ km.

For EsGB gravity, the constraints inferred by XG detectors are strongly dependent on the accuracy of the beyond-GR spin expansion (cf. Figs. 16 and 17). When we include second-order terms in the small-spin expansion we can better constrain the EsGB coupling constant, but there is still support for the GR hypothesis ($\alpha_{\mathrm{GB}} = 0$). The analysis suggests that the constraints may change when we include the unknown $\mathcal{O}(\chi^3)$ terms in the beyond-GR spin expansion. Moreover, even the most optimistic constraints inferred with the inclusion of second-order spin terms are still weaker than current bounds from LIGO/Virgo/KAGRA inspiral observations (Fig. 15). All of the bounds discussed above are optimistic, because they are found assuming that the masses and spins of the remnants are known (e.g., from the full inspiral-merger-ringdown waveform) and fixed to their injected values. When we allow the mass and spin to vary in the MCMC simulations, the posteriors on α_{GB} become uninformative.

Our findings for dCS are slightly more optimistic: by stacking ringdown observations with ET and CE we may constrain the dCS coupling constant α_{CS} with an accuracy about five times better than current bounds (Fig. 15), at least when we assume that the masses and spins of the remnants are known. In any case, ringdown-only constraints on α_{CS} would still be about two orders of magnitude worse than those achievable by XG observations of post-Newtonian corrections in the inspiral [5].

All of the EFT models that we consider lead to uninformative posteriors for cubic and quartic GR modifications, even when we consider values of the coupling so large that they would violate the domain of validity of the EFT. Constraining quartic models is particularly difficult because of the $1/M_f^6$ suppression of the beyond-Kerr deviations. The QNMs computed in EFT gravity include high-order terms in the spin expansion up to $\mathcal{O}(\chi^{12})$. The inclusion of these contributions is useful to understand the convergence properties of the small-spin expansion, but it does not significantly improve the bounds.

In conclusion, it is difficult to set constraints on theories with a dimensionful coupling—like those modifying GR in the large curvature regime—using ringdown observations alone. This is because constraints on any dimensionful

coupling constants are dominated by low-mass black holes, i.e., by higher frequency signals, for which ET and CE are less sensitive (at least, with the currently planned configuration). For both ET and CE, we find that the bounds obtained by stacking ~ 100 events with $M_f \lesssim 50 M_\odot$ are comparable to the bounds found by stacking the full dataset.

It should be possible to achieve better constraints on theories with dimensionful coupling constants by exploiting the full inspiral-merger-ringdown waveform. This is a more ambitious, ongoing research program that will require a combination of numerical simulations (see, e.g., [98–111]), post-Newtonian and effective-one-body models [71,112–119], and parametrizations of the full inspiral-merger-ringdown waveform [109,120–123]. The constraining power of ringdown observations is better for theories with a dimensionless coupling, for which suppression of beyond Kerr deviations due some energy scale is not effective. In this case the analysis we carried out in the agnostic, dimensionless scenario suggests that we may be able to set bounds at least on the leading-order (nonrotating) terms of the beyond-GR contribution to the QNM frequencies.

APPENDIX A: QUASINORMAL MODE FREQUENCIES FOR MODIFIED THEORIES OF GRAVITY

In this appendix we map the frequencies and damping times computed in specific modified theories of gravity to the ParSpec parameters defined in Eqs. (1) and (2). As discussed in Sec. III, we consider two models of quadratic gravity for which QNM frequencies have been computed in a slow-rotation expansion (EsGB and dCS gravity) and on eight models of EFT gravity.

The expansion coefficients for a slowly-rotating Kerr background can be found (in principle, to arbitrary order in the small-spin expansion) by fitting the numerical values computed, e.g., in Ref. [15]. For consistency with the highest-order expansion computed in Ref. [66], we fit the frequencies and damping times by a 12th order polynomial in χ using the `NonlinearModelFit` routine in *Mathematica*:

$$M\omega_{\ell m}^{\mathrm{K,s}} = \sum_{k_1=0}^{12} \bar{\omega}_{\ell m}^{(k_1)} \chi^{k_1}, \tag{A1}$$

$$\tau_{\ell m}^{\mathrm{K,s}}/M = \sum_{k_1=0}^{12} \bar{\tau}_{\ell m}^{(k_1)} \chi^{k_1}. \tag{A2}$$

We further impose that $(\bar{\omega}_{\ell m}^{(0)}, \bar{\tau}_{\ell m}^{(0)})$ are fixed to match exactly their Schwarzschild value. The resulting fitting coefficients are listed in Table III.

It would be interesting to consider alternative fitting procedures proposed in the literature (see, e.g., Ref. [69], in particular their Fig. 5), but the Taylor series presented here is sufficiently accurate for our present purposes. As shown in Fig. 25, the expansion reproduces the numerical QNM spectrum with a relative accuracy better than 1% in both frequency and damping time, as long as $\chi \lesssim 0.95$.

1. Einstein-scalar-Gauss-Bonnet gravity

The EsGB QNM spectrum at second order in a small-spin expansion was computed in Ref. [49]. In terms of the dimensionless spin $\chi = J/M^2$ and of the coupling constant $\beta_{\mathrm{GB}} = \alpha_{\mathrm{GB}}/M^2$ it reads

$$M\Omega_{\ell m} = \left(\Omega_0^A + \beta_{\mathrm{GB}}^2 \Omega_0^B\right) + m\chi\left(\Omega_1^A + \beta_{\mathrm{GB}}^2 \Omega_1^B\right) \\ + \chi^2\left[\left(\Omega_{2a}^A + \beta_{\mathrm{GB}}^2 \Omega_{2a}^B\right) + m^2\left(\Omega_{2b}^A + \beta_{\mathrm{GB}}^2 \Omega_{2b}^B\right)\right], \tag{A3}$$

where $\Omega_{\ell m} = \omega_{\ell m} - i/\tau_{\ell m}$ and the subscripts "A" and "B" refer to the GR and beyond-GR contributions, respectively. We can recast this expression into the ParSpec framework by identifying $\gamma = \alpha_{\mathrm{GB}}^2/M^4 = \beta_{\mathrm{GB}}^2$, such that

$$\omega_{\ell m} = \frac{1}{M}\left[\bar{\omega}_{\ell m}^{(0)}\left(1 + \beta_{\mathrm{GB}}^2 \delta\omega_{\ell m}^{(0)}\right) + \chi\bar{\omega}_{\ell m}^{(1)}\right. \\ \times \left(1 + \beta_{\mathrm{GB}}^2 \delta\omega_{\ell m}^{(1)}\right) + \chi^2\bar{\omega}_{\ell m}^{(2)}\left(1 + \beta_{\mathrm{GB}}^2 \delta\omega_{\ell m}^{(2)}\right)\right] \\ \left. + \frac{1}{M}\sum_{k_1=3}^{12} \bar{\omega}_{\ell m}^{(k_1)}, \right. \tag{A4}$$

TABLE III. Values of the GR coefficients for the ParSpec spin expansion for the modes with $(\ell m) = (22), (33), (21)$ modes up to order 12 in the spin.

k_1	$\bar{\omega}_{22}^{(k_1)}$	$\bar{\tau}_{22}^{(k_1)}$	$\bar{\omega}_{33}^{(k_1)}$	$\bar{\tau}_{33}^{(k_1)}$	$\bar{\omega}_{21}^{(k_1)}$	$\bar{\tau}_{21}^{(k_1)}$
0	0.37368	11.241	0.59946	10.787	0.37367	11.241
1	0.12947	0.98133	0.2065	0.95996	0.062929	0.4055
2	−0.017526	−19.094	0.0046294	−19.021	0.049757	−6.8293
3	0.80811	184.11	0.91245	185.08	−0.057124	70.950
4	−2.8392	−766.01	−3.0061	−769.48	0.46321	−292.17
5	4.3651	1434.2	4.1145	1441.7	−1.1694	534.65
6	1.0487	−608.63	2.9075	−614.3	1.4399	−182.24
7	−8.1789	−1337.3	−10.535	−1340.6	−0.27864	−566.17
8	−0.21522	594.52	−1.51	599.45	−1.0989	190.92
9	11.546	1730.8	15.248	1734.2	0.72622	749.45
10	2.5717	−111.70	4.3535	−114.94	0.9947	4.3652
11	−18.135	−2473.0	−24.505	−2476.4	−1.6423	−1098.0
12	9.4186	1387.6	12.54	1390.2	0.71601	604.36

$$
\tau_{\ell m} = M \Big[\bar{\tau}_{\ell m}^{(0)} \Big(1 + \beta_{\mathrm{GB}}^2 \delta\tau_{\ell m}^{(0)} \Big) + \chi \bar{\tau}_{\ell m}^{(1)}
$$
$$
\times \Big(1 + \beta_{\mathrm{GB}}^2 \delta\tau_{\ell m}^{(1)} \Big) + \chi^2 \bar{\tau}_{\ell m}^{(2)} \Big(1 + \beta_{\mathrm{GB}}^2 \delta\tau_{\ell m}^{(2)} \Big) \Big]
$$
$$
+ M \sum_{k_1=3}^{12} \bar{\tau}_{\ell m}^{(k_1)}, \tag{A5}
$$

where the quantities with an overbar identify the GR coefficients up to 12th order in the spin. We determine the beyond-GR parameters in Eqs. (A4) and (A5) by fitting them against the analytical expressions provided in Ref. [49] up to $\mathcal{O}(\chi^2)$ and $\mathcal{O}(\alpha_{\mathrm{GB}}^6)$. We perform the fit in a range of spins and coupling constants where the QNM frequencies are well described by Taylor expansions according to Ref. [49], i.e., $\chi \in [0, 0.2]$, $\alpha_{\mathrm{GB}}/M^2 \in [0, 0.2]$.

The values of the EsGB fitting coefficients found in this way are listed in Table IV. In Fig. 26 we plot the linear and quadratic spin corrections to the frequency and damping time of the (22), (33), and (21) modes, normalized to the nonrotating term, as functions of χ. Bullets mark values of χ
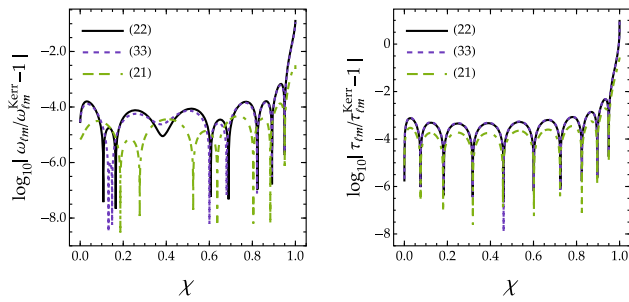
FIG. 25. Relative errors between the frequencies and damping times in GR computed by solving the Teukolsky equation and their slow-rotation approximations at twelfth-order in a Taylor expansion in the spin, with coefficients given in Table III.

TABLE IV. Values of the EsGB coefficients for the ParSpec spin expansion of the $(\ell m) = (22), (33), (21)$ modes.

k_2	$\delta\omega_{22}^{(k_2)}$	$\delta\tau_{22}^{(k_2)}$	$\delta\omega_{33}^{(k_2)}$	$\delta\tau_{33}^{(k_2)}$	$\delta\omega_{21}^{(k_2)}$	$\delta\tau_{21}^{(k_2)}$
0	−0.05062	−0.10606	−0.10432	−0.08524	−0.04227	−0.08179
1	0.51999	9.6119	0.22466	4.0986	0.11534	9.7028
2	47.388	2.1226	−250.52	0.97981	−5.4561	2.7290

for which the linear or quadratic correction becomes larger (in absolute value) than the nonrotating contribution. For example, the second-order corrections to the frequency and damping time of the (22) mode become larger than the nonrotating contribution for $\chi \gtrsim 0.2$. This highlights the importance of pushing the spin expansion to higher orders.

To further clarify the relevance of spin corrections, we study the mass dependence of the ParSpec coefficients $\gamma \delta\omega_{\ell m}^{k_2}$ and $\gamma \delta\tau_{\ell m}^{k_2}$ for EsGB gravity. In Fig. 27 we plot the nonrotating, linear, and quadratic spin terms as a function of M, assuming $\sqrt{\alpha_{\mathrm{GB}}} \simeq 2.5$ km, as discussed in Sec. IV B. Empty circles in each panel correspond to the minimum observable component mass among all binary BHs present
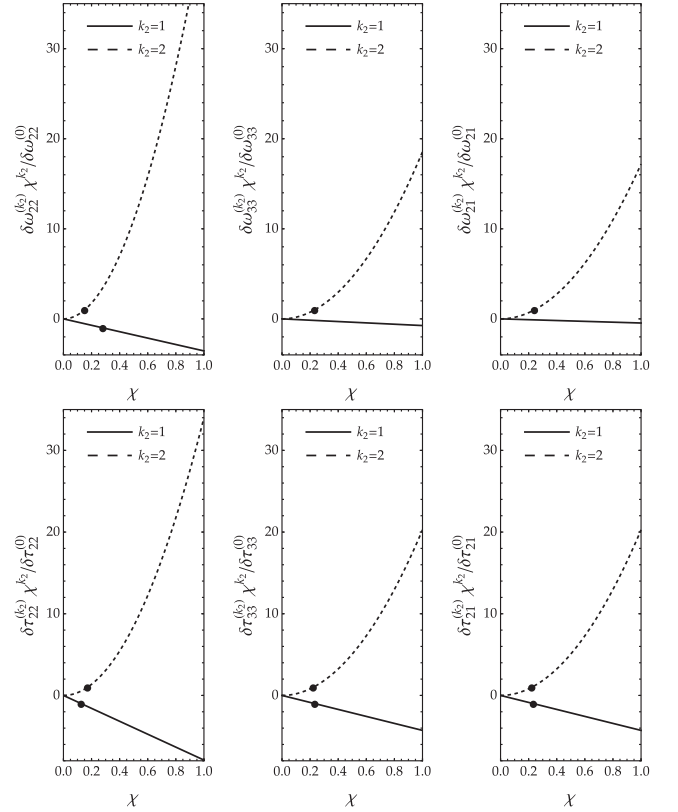
FIG. 26. Linear ($k_2 = 1$) and quadratic ($k_2 = 2$) corrections to the EsGB QNM frequency (top) and damping time (bottom) as a function of χ, normalized to the nonrotating term, for selected values of (ℓm). Bullets mark values of χ for which the linear or quadratic correction becomes larger (in absolute value) than the nonrotating contribution.
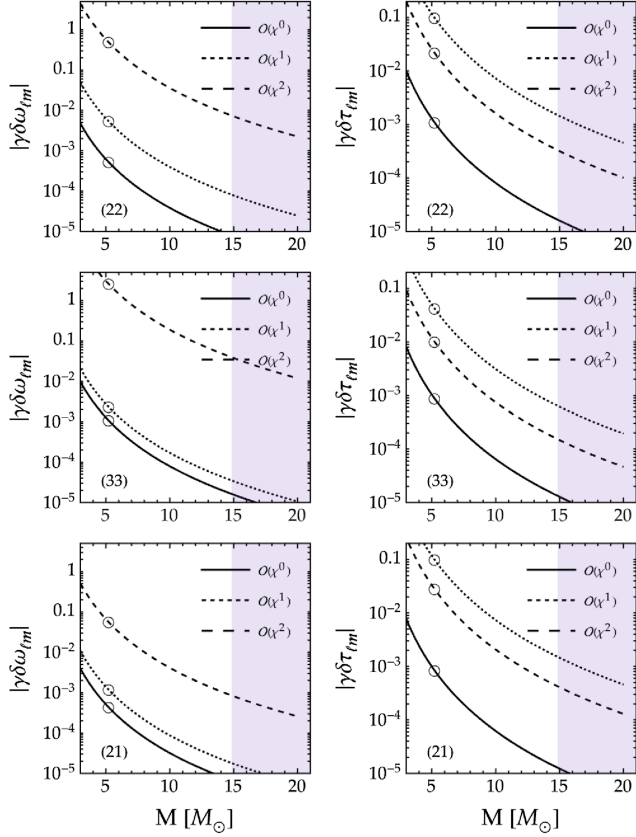
FIG. 27. ParSpec coefficients $\gamma(\delta\omega_{\ell m}^{k_2}, \delta\tau_{\ell m}^{k_2})$, with $\gamma = \beta_{GB}^2 = \alpha_{GB}^2/M^4$ for EsGB gravity, as a function of the BH mass. Solid, dashed and dotted curves refer to nonrotating, linear, and quadratic coefficients. Empty dots correspond to the smallest observable binary component mass in our population models. Vertical shaded regions identify the lightest remnant BH mass among all events with an observable ringdown. The top, center, and bottom rows show values of the coefficients for the (22), (33), and (21) mode, respectively.

in our catalogs. For EsGB gravity, both frequencies and damping time coefficients remain much smaller than one in the whole range of BH remnant masses.

In Fig. 28 we show the distributions of the relative change with respect to the slow-rotation expansion of the Kerr QNM frequencies and damping times $(\omega_{\ell m}^{K,s}, \tau_{\ell m}^{K,s})$, given by Eqs. (A1) and (A2), due to the EsGB corrections:

$$\Delta\omega_{\ell m}^{GB} = \frac{\beta_{GB}^2}{\omega_{\ell m}^{K,s}}\left[\bar{\omega}_{\ell m}^{(0)}\delta\omega_{\ell m}^{(0)} + \chi\bar{\omega}_{\ell m}^{(1)}\delta\omega_{\ell m}^{(1)} + \chi^2\bar{\omega}_{\ell m}^{(2)}\delta\omega_{\ell m}^{(2)}\right],$$

(A6)

$$\Delta\tau_{\ell m}^{GB} = \frac{\beta_{GB}^2}{\tau_{\ell m}^{K,s}}\left[\bar{\tau}_{\ell m}^{(0)}\delta\tau_{\ell m}^{(0)} + \chi\bar{\tau}_{\ell m}^{(1)}\delta\tau_{\ell m}^{(1)} + \chi^2\bar{\tau}_{\ell m}^{(2)}\delta\tau_{\ell m}^{(2)}\right], \quad \text{(A7)}$$

As discussed before we fix the EsGB coupling to $\sqrt{\alpha_{GB}} \sim$ 2.5 km and we evaluate $\Delta\omega_{\ell m}^{GB}$ and $\Delta\tau_{\ell m}^{GB}$ for all BHs of the two populations (models I and II) that have a detectable
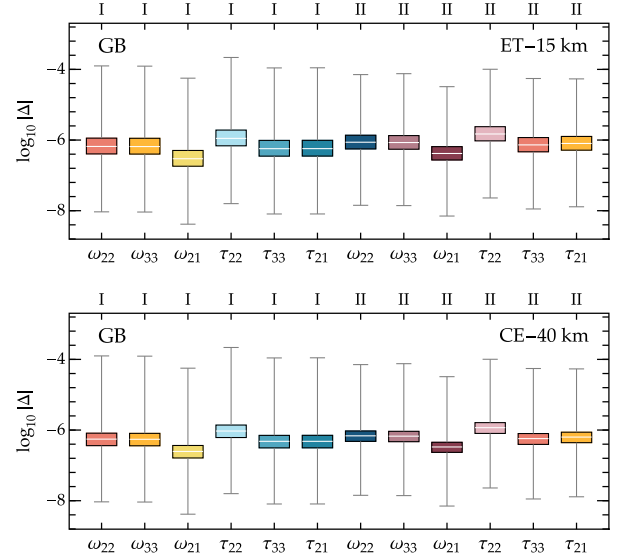

FIG. 28. Distributions of the relative change (A6) and (A7) in frequencies and damping times induced by EsGB gravity with respect to GR. White horizontal lines in each colored box mark the median over all events with ringdown observable by ET (top) and CE (bottom); the edges of the box identify the upper and lower quartiles, while the ends of the whiskers are the maximum and minimum value in the distribution. Labels in the top x axis of the panels refer to the spin model used in the astrophysical population (models I or II). The value of the EsGB coupling constant is fixed to $\sqrt{\alpha_{GB}} \simeq 2.5$ km [97].

ringdown in ET, i.e., those for which (22) mode SNR $\rho \geq 12$. The fundamental (22) mode leads to the largest deviations, but deviations in the (33) and (21) components are comparable in order of magnitude. Note however that the SNR, and therefore the uncertainties on frequency and damping times, depend on the energy released in each mode, ϵ_{RD} (or, equivalently, on the mode amplitudes). As discussed in Sec. IV (see in particular Fig. 5), the values of ϵ_{RD} for the (22) QNM are more than one order of magnitude larger than those of the (33) and (21) modes, and therefore the (22) mode is the most important for Bayesian inference.

2. Dynamical Chern-Simons gravity

For dCS gravity we use the analytical fits provided in Ref. [50], which are valid up to first order in the spin and up to second order in the coupling constant α_{CS}.[1] For simplicity we focus on the axial sector, which is more strongly affected by the dCS modification with respect to the polar sector.

---

[1]Reference [51] provides terms at leading order in the dCS coupling constant which were missing in Ref. [50]. However Ref. [51] does not list values for $l = 3$. Therefore, for consistency, we only use results from Ref. [50].

TABLE V. Values of the beyond-GR coefficients for the ParSpec spin expansion for the axial modes with $(\ell m) = (22), (33), (21)$ in dCS gravity, as computed from the fits provided in Ref. [50].

k_2	$\delta\omega_{22}^{(k_2)}$	$\delta\tau_{22}^{(k_2)}$	$\delta\omega_{33}^{(k_2)}$	$\delta\tau_{33}^{(k_2)}$	$\delta\omega_{21}^{(k_2)}$	$\delta\tau_{21}^{(k_2)}$
0	2.4182	4.8493	6.3807	−32.535	2.4569	4.8768
1	20.499	173.56	113.02	−3792.9	−22.421	−47.561

At this order, defining $\gamma = \alpha_{CS}^2/M^4 = \beta_{CS}^2$, the ParSpec mapping becomes

$$
\omega_{\ell m} = \frac{1}{M}\left[\bar{\omega}_{\ell m}^{(0)}\left(1 + \beta_{CS}^2 \delta\omega_{\ell m}^{(0)}\right)\right.
$$
$$
\left. + \chi\bar{\omega}_{\ell m}^{(1)}\left(1 + \beta_{CS}^2 \delta\omega_{\ell m}^{(1)}\right)\right] + M\sum_{k_1=2}^{12}\bar{\tau}_{\ell m}^{(k_1)}, \quad \text{(A8)}
$$

$$
\tau_{\ell m} = M\left[\bar{\tau}_{\ell m}^{(0)}\left(1 + \beta_{CS}^2 \delta\tau_{\ell m}^{(0)}\right)\right.
$$
$$
\left. + \chi\bar{\tau}_{\ell m}^{(1)}\left(1 + \beta_{CS}^2 \delta\tau_{\ell m}^{(1)}\right)\right] + M\sum_{k_1=2}^{12}\bar{\tau}_{\ell m}^{(k_1)}. \quad \text{(A9)}
$$

We find the beyond-GR parameters with a procedure analogous to the EsGB case. We fit Eqs. (A8) and (A9), with the GR coefficients fixed to the values listed in Table III, to the expressions provided in Ref. [50] in the range where the authors find their expansion to be reliable,
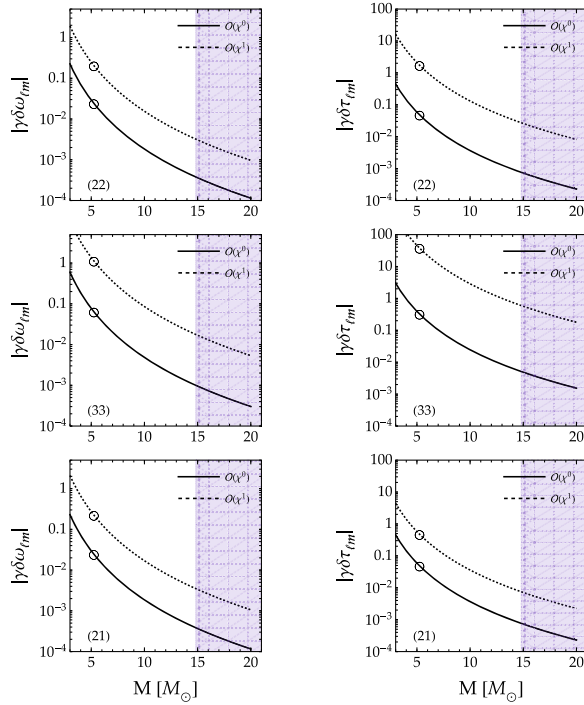
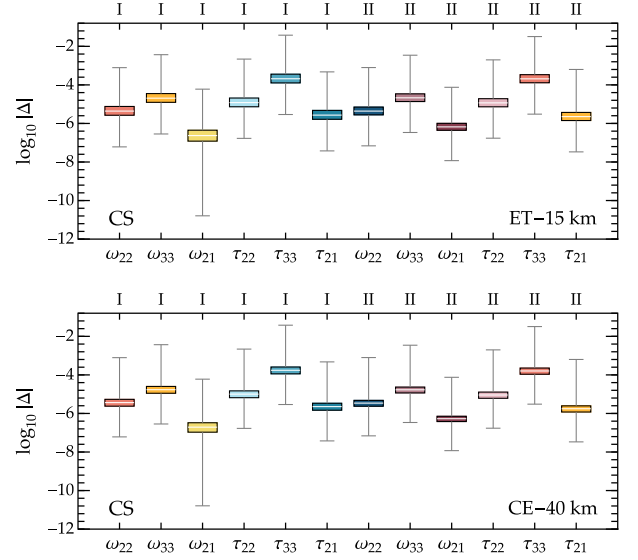FIG. 29.　Same as Fig. 27, but for dCS gravity.

FIG. 30.　Same as Fig. 28, but for dCS gravity. We assume the maximum value of the dCS coupling constant allowed by GW observations, $\sqrt{\alpha_{CS}} \simeq 2.5$ km [97].

i.e., $\chi \in [0, 0.04], \alpha_{CS}/M^2 \in [0, 0.1]$. The dCS axial coefficients of the ParSpec expansion determined in this way are listed (up to first order in the spin) in Table V.

In Fig. 29 we show the behavior of $\gamma\delta\omega_{\ell m}^{k_2}$ and $\gamma\delta\tau_{\ell m}^{k_2}$ as a function of the BH mass, for $\sqrt{\alpha_{CS}} \simeq 2.5$ km. Compared to EsGB gravity (Fig. 27) the values of the dCS parameters are, in general, larger.

Figure 30 is the dCS equivalent of Fig. 28 for EsGB gravity. Most of the considerations we made for EsGB gravity apply also in this case. Note, however, that deviations in the (33) component are now always larger than deviations in the (22) and (21) modes.

3. Effective field theory

For EFT gravity we use the numerical fits provided in Ref. [66], where beyond-Kerr QNM corrections are written in the form

$$
M\omega = M\omega_{\ell m}^{\text{kerr}} + \alpha_q \delta\omega^{\text{fit}}, \quad \text{(A10)}
$$

where

$$
\delta\omega_{\ell m}^{\text{fit}} = \sum_{n=0}^{12} c_n \chi^n. \quad \text{(A11)}
$$

The numerical coefficients c_n for the (22) and (33) QNMs—but not for the (21) QNM, which therefore is not included in this analysis—are given in Tables I and II of Ref. [66].

Equation (A10) can be straightforwardly mapped to the ParSpec coefficients in Eqs. (1) and (2) once we are given the GR part of the expansion, Eqs. (A1) and (A2).
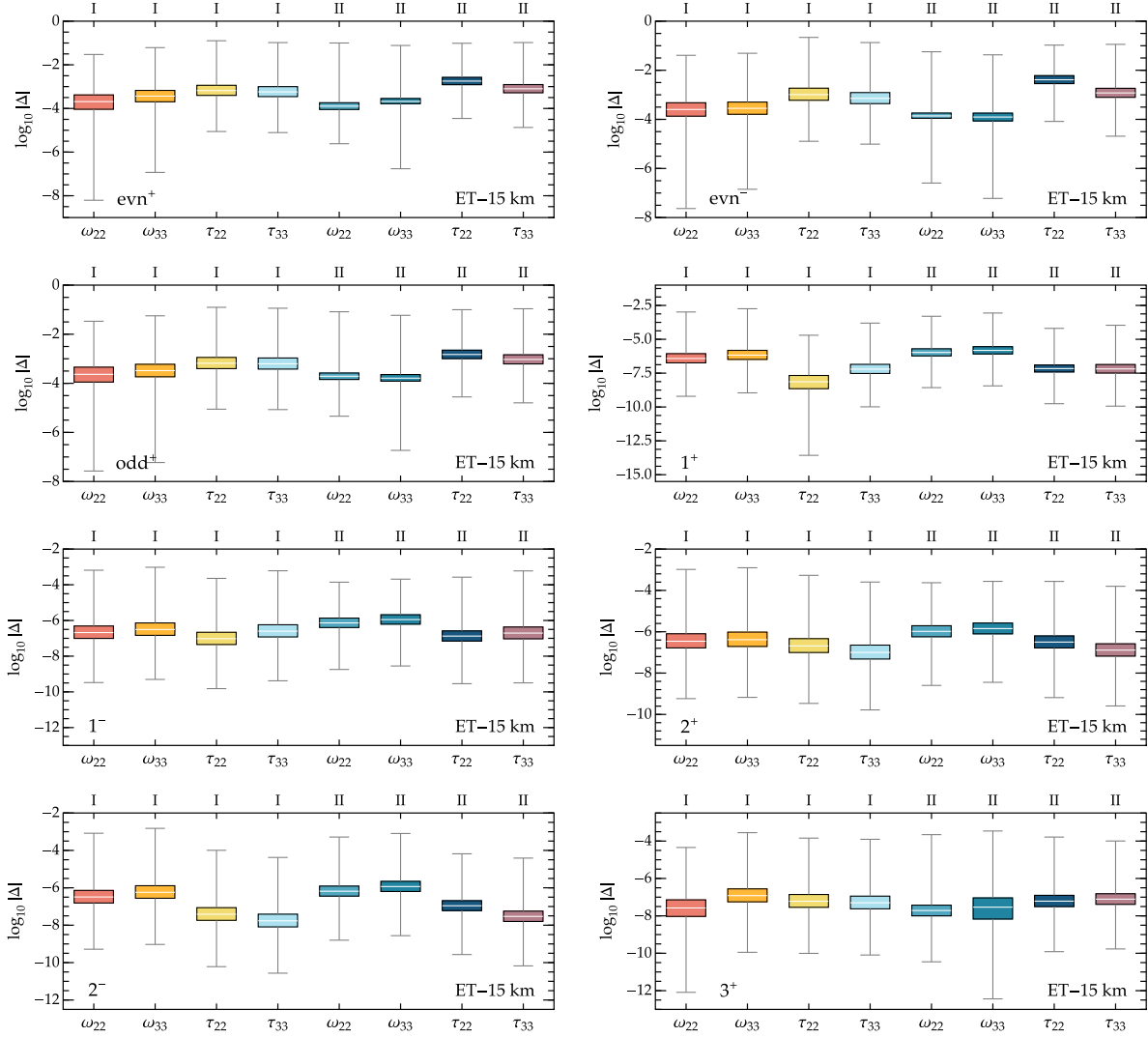
FIG. 31. Same as Figs. 28 and 30, but for the eight EFT gravity models of Ref. [66] considered in this work. We only show ringdown signals detectable by ET. The results for CE are qualitatively similar.

Note that for EFT gravity the beyond-GR component of the spin expansion is known to much higher order ($n_1 = n_2 = 12$) than for EsGB and dCS gravity. The distributions of deviations from the Kerr frequencies and damping times in the eight EFT models we consider are shown in Fig. 31. In all panels we assume a value of the coupling such that $\gamma_q = 0.1$, and (for brevity) we only consider ringdown observations with ET.

APPENDIX B: BLACK HOLE MASS AND SPIN MEASUREMENTS

The main purpose of this appendix is to show that the remnant BH's mass and spin are generally recovered with remarkable accuracy. For illustration, in Fig. 32 we plot the

probability distributions for the masses and spins inferred from a two-mode analysis for dCS gravity. Black and green contours correspond to constraints derived by stacking together two different sets of ringdown events (compare Fig. 22), while red dots identify the injected values of M_f and χ_f.

The constraints are strongly dependent on the SNR of the fundamental mode, shown on the top horizontal axis of each panel. The figure illustrates four key features common to the (22)–(33) and the (22)–(21) analysis: (i) the constraints on M_f and χ_f are almost independent of the number of events that we stack, (ii) masses are always better measured than spins, and (iii) measurement accuracies are slightly worse for model II, which exhibits long tails in the posterior distributions of the spins.
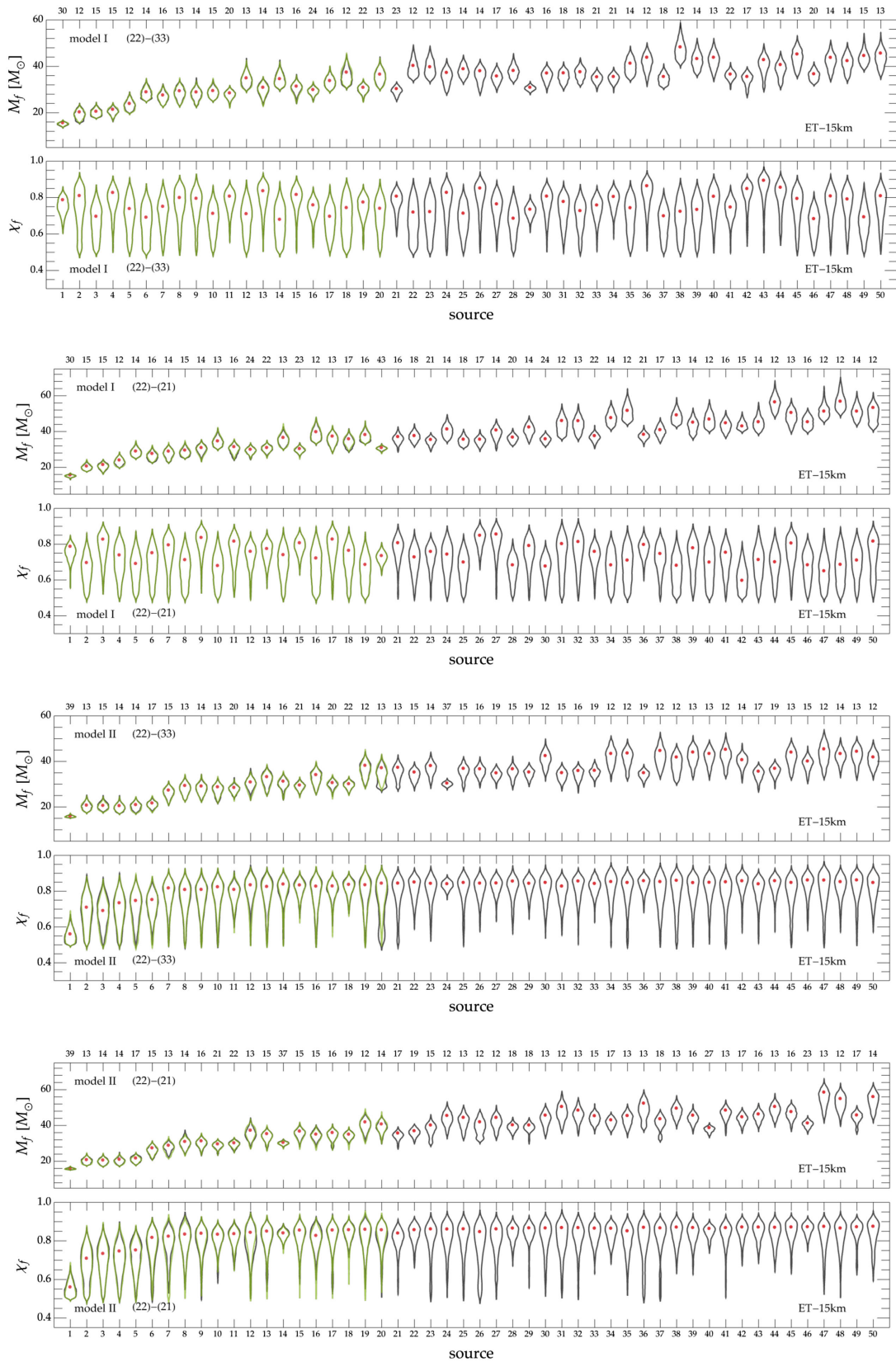
FIG. 32. Posterior densities for the BH masses and spins inferred from a (22)–(33) and (22)–(21) analysis for dCS gravity. All constraints are inferred from ringdown observations made by ET for BHs following either the models I or II spin distributions. Black (green) violins are obtained by stacking the largest (second largest) set of observations available in the population model (see also Fig. 22). Red bullets are the injected values. Labels on the top horizontal axes identify the values of the SNR of the fundamental mode.
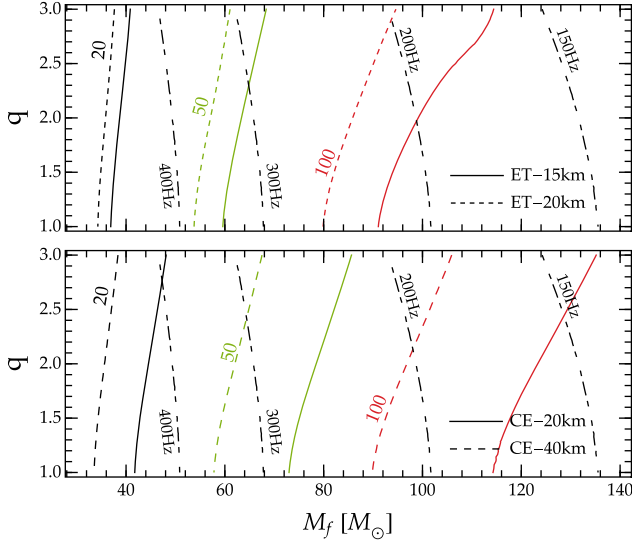
FIG. 33. Curves of constant SNR for the (22) mode as a function of the remnant BH mass M_f and of the binary progenitor's mass ratio q for ET (solid) and CE (dashed). Dot-dashed curves identify configurations with a specific (22) frequency.
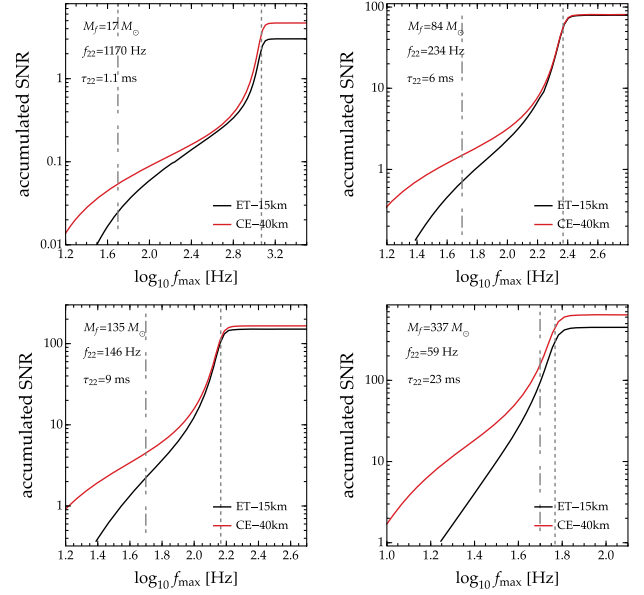

FIG. 34. Accumulated SNR as a function of the maximum integration frequency $f_{\max}$ for four different BH binaries, all at luminosity distance $d_L = 1$ Gpc (see text). The vertical dashed and dot-dashed lines in each panel mark the frequency of the fundamental (22) mode and a reference frequency of 50 Hz, respectively.

APPENDIX C: SNR AND DETECTOR CONFIGURATIONS

In this appendix we analyze the dependence of the ringdown SNR on the remnant BH mass M_f and on the progenitor binary mass ratio q. We also study how the SNR changes for ET and CE detectors with different armlengths. Note that in this case we consider a single ET, at variance with the network configuration of two L-shaped detectors used in the rest of the paper. To compute the SNR we uniformly sample the primary mass in the range $m_1 \in [10, 100] M_\odot$ and the mass ratio in the range $q = m_1/m_2 \in [1, 3]$. For simplicity we assume the component spins to be equal and aligned, with magnitude $\chi_1 = \chi_2 = 0.6$. This choice leads to BH remnants with spins in the range $\chi_f \sim [0.78, 0.86]$, but we have checked that our conclusions are only mildly dependent on the magnitude of the progenitor BH spins. For concreteness, and without loss of generality, we assume all sources to be located at luminosity distance $d_L = 1$ Gpc. We compute the amplitude of the (22) mode and the resulting SNR as described in Sec. III C.

In Fig. 33 we show constant-SNR contours for ET and CE computed in this way. The two configurations we considered in this paper, the 15 km ET and the 40 km CE, have similar performances, with the latter yielding slightly larger SNRs. A 20-km armlength CE would yield lower values of the ringdown SNR, with differences with respect to ET increasing as the remnant BH mass grows.

To further clarify what frequency range is most relevant for ringdown tests, we have computed the accumulated SNR for ET (single detector) and CE as a function of the upper integration limit $f_{\max}$ in Eq. (14). We focus on BH remnants from four selected binary systems with primary mass $m_1 = (10, 50, 80, 200) M_\odot$, mass ratio $q = 2$, aligned spins $\chi_1 = \chi_2 = 0.6$, and luminosity distance $d_L = 1$ Gpc.

In the four panels of Fig. 34 (one for each binary) we see some common trends: the SNR for CE is always larger than for ET when $f_{\mathrm{fmax}} \lesssim 50$ Hz, since the PSD of CE is better between 3 and 50 Hz. As $f_{\max}$ increases, ET "catches up" and yields an overall similar value of the SNR. Indeed, for $f \gtrsim 50$ Hz, the ratio of the two PSDs is similar, with CE being less than a factor of 2 better than ET. Note also that for the mass range of interest, i.e., for BHs with $M_f \sim 100 M_\odot$, the two interferometers yields almost identical results.

APPENDIX D: FISHER MATRIX COMPONENTS

Here we list explicit expressions for the components of the FIM (22), averaged over the source orientation in the sky. The FIM is computed from the ringdown waveform model of Eqs. (9)–(11). For a given multipolar component (ℓm) it depends on the parameters $\vec{\theta} = (\bar{\mathcal{A}}_{\ell m}, \phi_{\ell m}, \omega_{\ell m}, \tau_{\ell m})$, where $\bar{\mathcal{A}}_{\ell m} = M \mathcal{A}_{\ell m}/r$ is the effective amplitude. Denoting partial derivatives by commas, the ten independent components of the FIM are given by

$$\Gamma_{\bar{\mathcal{A}}_{\ell m} \bar{\mathcal{A}}_{\ell m}} = \frac{b_+^2 + b_-^2}{10\pi}, \qquad \Gamma_{\phi_{\ell m} \phi_{\ell m}} = \frac{\bar{\mathcal{A}}_{\ell m}^2}{10\pi}[b_+^2 + b_-^2],$$

$$\Gamma_{\omega_{\ell m} \omega_{\ell m}} = \frac{\bar{\mathcal{A}}_{\ell m}^2}{10\pi}[b_{+,\omega_{\ell m}}^2 + b_{-,\omega_{\ell m}}^2], \tag{D1}$$

$$\Gamma_{\tau_{\ell m}\tau_{\ell m}} = \frac{\bar{\mathcal{A}}_{\ell m}^2}{10\pi}[b_{+,\tau_{\ell m}}^2 + b_{-,\tau_{\ell m}}^2],$$

$$\Gamma_{\bar{\mathcal{A}}_{\ell m}\phi_{\ell m}} = \frac{\bar{\mathcal{A}}_{\ell m}}{10\pi}[-ib_+^2 + ib_-^2], \qquad (D2)$$

$$\Gamma_{\bar{\mathcal{A}}_{\ell m}\omega_{\ell m}} = \frac{\bar{\mathcal{A}}_{\ell m}}{10\pi}[b_+ b_{+,\omega_{\ell m}} + b_- b_{-,\omega_{\ell m}}],$$

$$\Gamma_{\bar{\mathcal{A}}_{\ell m}\tau_{\ell m}} = \frac{\bar{\mathcal{A}}_{\ell m}}{10\pi}[b_+ b_{+,\tau_{\ell m}} + b_- b_{-,\tau_{\ell m}}], \qquad (D3)$$

$$\Gamma_{\phi_{\ell m}\omega_{\ell m}} = \frac{i\bar{\mathcal{A}}_{\ell m}^2}{10\pi}[b_+ b_{+,\omega_{\ell m}} - b_- b_{-,\omega_{\ell m}}],$$

$$\Gamma_{\phi_{\ell m}\tau_{\ell m}} = \frac{i\bar{\mathcal{A}}_{\ell m}^2}{10\pi}[b_+ b_{+,\tau_{\ell m}} - b_- b_{-,\tau_{\ell m}}], \qquad (D4)$$

$$\Gamma_{\omega_{\ell m}\tau_{\ell m}} = \frac{\bar{\mathcal{A}}_{\ell m}^2}{10\pi}[b_{+,\omega_{\ell m}} b_{+,\tau_{\ell m}} + b_{-,\omega_{\ell m}} b_{-,\tau_{\ell m}}]. \qquad (D5)$$

[1] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), GWTC-3: Compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run, Phys. Rev. X **13**, 041039 (2023).

[2] M. Punturo *et al.*, The Einstein Telescope: A third-generation gravitational wave observatory, Classical Quantum Gravity **27**, 194002 (2010).

[3] Sheila Dwyer, Daniel Sigg, Stefan W. Ballmer, Lisa Barsotti, Nergis Mavalvala, and Matthew Evans, Gravitational wave detector with cosmological reach, Phys. Rev. D **91**, 082001 (2015).

[4] Michele Maggiore *et al.*, Science case for the Einstein Telescope, J. Cosmol. Astropart. Phys. 03 (2020) 050.

[5] Scott E. Perkins, Nicolás Yunes, and Emanuele Berti, Probing fundamental physics with gravitational waves: The next generation, Phys. Rev. D **103**, 044024 (2021).

[6] Marica Branchesi *et al.*, Science with the Einstein Telescope: A comparison of different designs, J. Cosmol. Astropart. Phys. 07 (2023) 068.

[7] Ssohrab Borhanian and B. S. Sathyaprakash, Listening to the universe with next generation ground-based gravitational-wave detectors, arXiv:2202.11048.

[8] Matthew Evans *et al.*, Cosmic Explorer: A submission to the NSF MPSAC ngGW subcommittee, arXiv:2306.13745.

[9] Ish Gupta *et al.*, Characterizing gravitational wave detector networks: From A# to Cosmic Explorer, arXiv:2307.10421.

[10] Kostas D. Kokkotas and Bernd G. Schmidt, Quasinormal modes of stars and black holes, Living Rev. Relativity **2**, 2 (1999).

[11] Valeria Ferrari and Leonardo Gualtieri, Quasi-normal modes and gravitational wave astronomy, Gen. Relativ. Gravit. **40**, 945 (2008).

[12] Emanuele Berti, Vitor Cardoso, and Andrei O. Starinets, Quasinormal modes of black holes and black branes, Classical Quantum Gravity **26**, 163001 (2009).

[13] Steven L. Detweiler, Black holes and gravitational waves. III. The resonant frequencies of rotating holes, Astrophys. J. **239**, 292 (1980).

[14] Olaf Dreyer, Bernard J. Kelly, Badri Krishnan, Lee Samuel Finn, David Garrison, and Ramon Lopez-Aleman, Black hole spectroscopy: Testing general relativity through gravitational wave observations, Classical Quantum Gravity **21**, 787 (2004).

[15] Emanuele Berti, Vitor Cardoso, and Clifford M. Will, On gravitational-wave spectroscopy of massive black holes with the space interferometer LISA, Phys. Rev. D **73**, 064030 (2006).

[16] S. Gossan, J. Veitch, and B. S. Sathyaprakash, Bayesian model selection for testing the no-hair theorem with black hole ringdowns, Phys. Rev. D **85**, 124056 (2012).

[17] J. Meidam, M. Agathos, C. Van Den Broeck, J. Veitch, and B. S. Sathyaprakash, Testing the no-hair theorem with black hole ringdowns using TIGER, Phys. Rev. D **90**, 064009 (2014).

[18] Emanuele Berti, Alberto Sesana, Enrico Barausse, Vitor Cardoso, and Krzysztof Belczynski, Spectroscopy of Kerr black holes with Earth- and space-based interferometers, Phys. Rev. Lett. **117**, 101102 (2016).

[19] Huan Yang, Kent Yagi, Jonathan Blackman, Luis Lehner, Vasileios Paschalidis, Frans Pretorius, and Nicolás Yunes, Black hole spectroscopy with coherent mode stacking, Phys. Rev. Lett. **118**, 161101 (2017).

[20] Andrea Maselli, Kostas Kokkotas, and Pablo Laguna, Observing binary black hole ringdowns by advanced gravitational wave detectors, Phys. Rev. D **95**, 104026 (2017).

[21] Vishal Baibhav, Emanuele Berti, Vitor Cardoso, and Gaurav Khanna, Black hole spectroscopy: Systematic errors and ringdown energy estimates, Phys. Rev. D **97**, 044048 (2018).

[22] Vishal Baibhav and Emanuele Berti, Multimode black hole spectroscopy, Phys. Rev. D **99**, 024005 (2019).

[23] Vishal Baibhav, Emanuele Berti, and Vitor Cardoso, LISA parameter estimation and source localization with higher harmonics of the ringdown, Phys. Rev. D **101**, 084053 (2020).

[24] Iara Ota and Cecilia Chirenti, Black hole spectroscopy horizons for current and future gravitational wave detectors, Phys. Rev. D **105**, 044015 (2022).

[25] Swetha Bhagwat, Costantino Pacilio, Enrico Barausse, and Paolo Pani, Landscape of massive black-hole spectroscopy with LISA and the Einstein Telescope, Phys. Rev. D **105**, 124063 (2022).

[26] Vishal Baibhav, Mark Ho-Yeuk Cheung, Emanuele Berti, Vitor Cardoso, Gregorio Carullo, Roberto Cotesta, Walter Del Pozzo, and Francisco 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).

[27] Clifford M. Will and Kenneth Nordtvedt, Jr., Conservation laws and preferred frames in relativistic gravity. I. Preferred-frame theories and an extended PPN formalism, Astrophys. J. **177,** 757 (1972).

[28] Clifford M. Will, The confrontation between general relativity and experiment, Living Rev. Relativity **17,** 4 (2014).

[29] Luc Blanchet and B. S. Sathyaprakash, Detecting the tail effect in gravitational wave experiments, Phys. Rev. Lett. **74,** 1067 (1995).

[30] K. G. Arun, Bala R. Iyer, M. S. S. Qusailah, and B. S. Sathyaprakash, Probing the non-linear structure of general relativity with black hole binaries, Phys. Rev. D **74,** 024006 (2006).

[31] K. G. Arun, Bala R. Iyer, M. S. S. Qusailah, and B. S. Sathyaprakash, Testing post-Newtonian theory with gravitational wave observations, Classical Quantum Gravity **23,** L37 (2006).

[32] Nicolas Yunes and Frans Pretorius, Fundamental theoretical bias in gravitational wave astrophysics and the parameterized post-Einsteinian framework, Phys. Rev. D **80,** 122003 (2009).

[33] Katerina Chatziioannou, Nicolas Yunes, and Neil Cornish, Model-independent test of general relativity: An extended post-Einsteinian framework with complete polarization content, Phys. Rev. D **86,** 022004 (2012); **95,** 129901(E) (2017).

[34] Katie Chamberlain and Nicolas Yunes, Theoretical physics implications of gravitational wave observation with future detectors, Phys. Rev. D **96,** 084039 (2017).

[35] Nicolás Yunes and Xavier Siemens, Gravitational-wave tests of general relativity with ground-based detectors and Pulsar Timing-Arrays, Living Rev. Relativity **16,** 9 (2013).

[36] Emanuele Berti, Kent Yagi, and Nicolás Yunes, Extreme gravity tests with gravitational waves from compact binary coalescences: (I) inspiral-merger, Gen. Relativ. Gravit. **50,** 46 (2018).

[37] Oliver J. Tattersall, Pedro G. Ferreira, and Macarena Lagos, General theories of linear gravitational perturbations to a Schwarzschild black hole, Phys. Rev. D **97,** 044021 (2018).

[38] Emanuele Berti, Kent Yagi, Huan Yang, and Nicolás Yunes, Extreme gravity tests with gravitational waves from compact binary coalescences: (II) ringdown, Gen. Relativ. Gravit. **50,** 49 (2018).

[39] Vitor Cardoso, Masashi Kimura, Andrea Maselli, Emanuele Berti, Caio F. B. Macedo, and Ryan McManus, Parametrized black hole quasinormal ringdown: Decoupled equations for nonrotating black holes, Phys. Rev. D **99,** 104077 (2019).

[40] Ryan McManus, Emanuele Berti, Caio F. B. Macedo, Masashi Kimura, Andrea Maselli, and Vitor Cardoso, Parametrized black hole quasinormal ringdown. II. Coupled equations and quadratic corrections for nonrotating black holes, Phys. Rev. D **100,** 044061 (2019).

[41] S. H. Völkel, N. Franchini, and E. Barausse, Theory-agnostic reconstruction of potential and couplings from quasinormal modes, Phys. Rev. D **105,** 084046 (2022).

[42] S. H. Völkel, N. Franchini, E. Barausse, and E. Berti, Constraining modifications of black hole perturbation potentials near the light ring with quasinormal modes, Phys. Rev. D **106,** 124036 (2022).

[43] N. Franchini and S. H. Völkel, Parametrized quasinormal mode framework for non-Schwarzschild metrics, Phys. Rev. D **107,** 124063 (2023).

[44] Saul A. Teukolsky, The Kerr metric, Classical Quantum Gravity **32,** 124006 (2015).

[45] Saul A. Teukolsky, Perturbations of a rotating black hole. 1. Fundamental equations for gravitational electromagnetic and neutrino field perturbations, Astrophys. J. **185,** 635 (1973).

[46] Paolo Pani, Advanced methods in black-hole perturbation theory, Int. J. Mod. Phys. A **28,** 1340018 (2013).

[47] Nicola Franchini, Slow rotation black hole perturbation theory, Phys. Rev. D **108,** 044079 (2023).

[48] Lorenzo Pierini and Leonardo Gualtieri, Quasi-normal modes of rotating black holes in Einstein-dilaton Gauss-Bonnet gravity: The first order in rotation, Phys. Rev. D **103,** 124017 (2021).

[49] Lorenzo Pierini and Leonardo Gualtieri, Quasinormal modes of rotating black holes in Einstein-dilaton Gauss-Bonnet gravity: The second order in rotation, Phys. Rev. D **106,** 104009 (2022).

[50] Pratik Wagle, Nicolas Yunes, and Hector O. Silva, Quasinormal modes of slowly-rotating black holes in dynamical Chern-Simons gravity, Phys. Rev. D **105,** 124003 (2022).

[51] Manu Srivastava, Yanbei Chen, and S. Shankaranarayanan, Analytical computation of quasinormal modes of slowly rotating black holes in dynamical Chern-Simons gravity, Phys. Rev. D **104,** 064034 (2021).

[52] Pratik Wagle, Dongjun Li, Yanbei Chen, and Nicolas Yunes, Perturbations of spinning black holes in dynamical Chern-Simons gravity I. Slow rotation equations, arXiv:2311.07706.

[53] Pablo A. Cano, Kwinten Fransen, and Thomas Hertog, Ringing of rotating black holes in higher-derivative gravity, Phys. Rev. D **102,** 044047 (2020).

[54] Pablo A. Cano, Kwinten Fransen, Thomas Hertog, and Simon Maenaut, Gravitational ringing of rotating black holes in higher-derivative gravity, Phys. Rev. D **105,** 024064 (2022).

[55] Oscar J. C. Dias, Mahdi Godazgar, and Jorge E. Santos, Linear mode stability of the Kerr-Newman black hole and its quasinormal modes, Phys. Rev. Lett. **114,** 151101 (2015).

[56] Oscar J. C. Dias, Mahdi Godazgar, Jorge E. Santos, Gregorio Carullo, Walter Del Pozzo, and Danny Laghi, Eigenvalue repulsions in the quasinormal spectra of the Kerr-Newman black hole, Phys. Rev. D **105,** 084044 (2022).

[57] Oscar J. C. Dias, Mahdi Godazgar, and Jorge E. Santos, Eigenvalue repulsions and quasinormal mode spectra of Kerr-Newman: An extended study, J. High Energy Phys. 07 (2022) 076.

[58] Alex Davey, Oscar J. C. Dias, and Jorge E. Santos, Scalar QNM spectra of Kerr and Reissner-Nordström revealed by eigenvalue repulsions in Kerr-Newman, J. High Energy Phys. 12 (2023) 101.

[59] Adrian Ka-Wai Chung, Pratik Wagle, and Nicolas Yunes, Spectral method for the gravitational perturbations of black holes: Schwarzschild background case, Phys. Rev. D **107,** 124032 (2023).

[60] Dongjun Li, Pratik Wagle, Yanbei Chen, and Nicolás Yunes, Perturbations of spinning black holes beyond general relativity: Modified Teukolsky equation, Phys. Rev. X **13,** 021029 (2023).

[61] Asad Hussain and Aaron Zimmerman, Approach to computing spectral shifts for black holes beyond Kerr, Phys. Rev. D **106,** 104018 (2022).

[62] Pablo A. Cano, Kwinten Fransen, Thomas Hertog, and Simon Maenaut, Universal Teukolsky equations and black hole perturbations in higher-derivative gravity, Phys. Rev. D **108,** 024040 (2023).

[63] Dongjun Li, Asad Hussain, Pratik Wagle, Yanbei Chen, Nicolás Yunes, and Aaron Zimmerman, Isospectrality breaking in the Teukolsky formalism, arXiv:2310.06033.

[64] K. Dey, E. Barausse, and S. Basak, Measuring deviations from the Kerr geometry with black hole ringdown, Phys. Rev. D **108,** 024064 (2023).

[65] R. Ghosh, N. Franchini, S. H. Völkel, and E. Barausse, Quasinormal modes of nonseparable perturbation equations: The scalar non-Kerr case, Phys. Rev. D **108,** 024038 (2023).

[66] Pablo A. Cano, Kwinten Fransen, Thomas Hertog, and Simon Maenaut, Quasinormal modes of rotating black holes in higher-derivative gravity, Phys. Rev. D **108,** 124032 (2023).

[67] Andrea Maselli, Paolo Pani, Leonardo Gualtieri, and Emanuele Berti, Parametrized ringdown spin expansion coefficients: A data-analysis framework for black-hole spectroscopy with multiple events, Phys. Rev. D **101,** 024043 (2020).

[68] Gregorio Carullo, Enhancing modified gravity detection from gravitational-wave observations using the parametrized ringdown spin expansion coeffcients formalism, Phys. Rev. D **103,** 124043 (2021).

[69] Gregorio Carullo, Danny Laghi, Nathan K. Johnson-McDaniel, Walter Del Pozzo, Oscar J. C. Dias, Mahdi Godazgar, and Jorge E. Santos, Constraints on Kerr-Newman black holes from merger-ringdown gravitational-wave observations, Phys. Rev. D **105,** 062009 (2022).

[70] Vitor Cardoso and Paolo Pani, Testing the nature of dark compact objects: A status report, Living Rev. Relativity **22,** 4 (2019).

[71] Félix-Louis Julié, Hector O. Silva, Emanuele Berti, and Nicolás Yunes, Black hole sensitivities in Einstein-scalar-Gauss-Bonnet gravity, Phys. Rev. D **105,** 124031 (2022).

[72] Abhishek Hegade K. R., Justin L. Ripley, and Nicolás Yunes, Where and why does Einstein-scalar-Gauss-Bonnet theory break down?, Phys. Rev. D **107,** 044044 (2023).

[73] Stephon Alexander and Nicolas Yunes, Chern-Simons modified general relativity, Phys. Rep. **480,** 1 (2009).

[74] Emanuele Berti *et al.*, Testing general relativity with present and future astrophysical observations, Classical Quantum Gravity **32,** 243001 (2015).

[75] James Healy, Carlos O. Lousto, and Yosef Zlochower, Remnant mass, spin, and recoil from spin aligned black-hole binaries, Phys. Rev. D **90,** 104004 (2014).

[76] Eanna E. Flanagan and Scott A. Hughes, Measuring gravitational waves from binary black hole coalescences: 1. Signal-to-noise for inspiral, merger, and ringdown, Phys. Rev. D **57,** 4535 (1998).

[77] B. S. Sathyaprakash and B. F. Schutz, Physics, astrophysics and cosmology with gravitational waves, Living Rev. Relativity **12,** 2 (2009).

[78] Michele Vallisneri, Use and abuse of the Fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects, Phys. Rev. D **77,** 042001 (2008).

[79] Curt Cutler and Eanna E. Flanagan, Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral wave form?, Phys. Rev. D **49,** 2658 (1994).

[80] 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).

[81] Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman, EMCEE: The MCMC Hammer, Publ. Astron. Soc. Pac. **125,** 306 (2013).

[82] R. Abbott *et al.* (KAGRA, Virgo, and LIGO Scientific Collaborations), Population of merging compact binaries inferred using gravitational waves through GWTC-3, Phys. Rev. X **13,** 011048 (2023).

[83] Piero Madau and Mark Dickinson, Cosmic star formation history, Annu. Rev. Astron. Astrophys. **52,** 415 (2014).

[84] Ken K. Y. Ng, Salvatore Vitale, Will M. Farr, and Carl L. Rodriguez, Probing multiple populations of compact binaries with third-generation gravitational-wave detectors, Astrophys. J. Lett. **913,** L5 (2021).

[85] Enrico Barausse and Thomas P. Sotiriou, Black holes in Lorentz-violating gravity theories, Classical Quantum Gravity **30,** 244010 (2013).

[86] P. Kanti, N. E. Mavromatos, J. Rizos, K. Tamvakis, and E. Winstanley, Dilatonic black holes in higher curvature string gravity, Phys. Rev. D **54,** 5049 (1996).

[87] Paolo Pani and Vitor Cardoso, Are black holes in alternative theories serious astrophysical candidates? The case for Einstein-dilaton-Gauss-Bonnet black holes, Phys. Rev. D **79,** 084031 (2009).

[88] Jose Luis Blázquez-Salcedo, Caio F. B. Macedo, Vitor Cardoso, Valeria Ferrari, Leonardo Gualtieri, Fech Scen Khoo, Jutta Kunz, and Paolo Pani, Perturbed black holes in Einstein-dilaton-Gauss-Bonnet gravity: Stability, ringdown, and gravitational-wave emission, Phys. Rev. D **94,** 104024 (2016).

[89] Nicolas Yunes and Frans Pretorius, Dynamical Chern-Simons modified gravity. I. Spinning black holes in the slow-rotation approximation, Phys. Rev. D **79,** 084043 (2009).

[90] C. Molina, Paolo Pani, Vitor Cardoso, and Leonardo Gualtieri, Gravitational signature of Schwarzschild black

holes in dynamical Chern-Simons gravity, Phys. Rev. D **81,** 124021 (2010).

[91] Jose Luis Blázquez-Salcedo, Caio F. B. Macedo, Vitor Cardoso, Valeria Ferrari, Leonardo Gualtieri, Fech Scen Khoo, Jutta Kunz, and Paolo Pani, Perturbed black holes in Einstein-dilaton-Gauss-Bonnet gravity: Stability, ring-down, and gravitational-wave emission, Phys. Rev. D **94,** 104024 (2016).

[92] Emanuele Berti, Vitor Cardoso, Jose A. Gonzalez, Ulrich Sperhake, Mark Hannam, Sascha Husa, and Bernd Bruegmann, Inspiral, merger and ringdown of unequal mass black hole binaries: A multipolar analysis, Phys. Rev. D **76,** 064034 (2007).

[93] Lawrence E. Kidder, Using full information when comput-ing modes of post-Newtonian waveforms from inspiralling compact binaries in circular orbit, Phys. Rev. D **77,** 044016 (2008).

[94] E. Berti, V. Cardoso, J. A. Gonzalez, U. Sperhake, and Bernd Bruegmann, Multipolar analysis of spinning bina-ries, Classical Quantum Gravity **25,** 114035 (2008).

[95] Swetha Bhagwat, Costantino Pacilio, Paolo Pani, and Michela Mapelli, Landscape of stellar-mass black-hole spectroscopy with third-generation gravitational-wave detectors, Phys. Rev. D **108,** 043019 (2023).

[96] Costantino Pacilio and Swetha Bhagwat, Identifying modified theories of gravity using binary black-hole ring-downs, Phys. Rev. D **107,** 083021 (2023).

[97] Zhenwei Lyu, Nan Jiang, and Kent Yagi, Constraints on Einstein-dilation-Gauss-Bonnet gravity from black hole-neutron star gravitational wave events, Phys. Rev. D **105,** 064001 (2022); **106,** 069901(E) (2022).

[98] Maria Okounkova, Leo C. Stein, Mark A. Scheel, and Daniel A. Hemberger, Numerical binary black hole merg-ers in dynamical Chern-Simons gravity: Scalar field, Phys. Rev. D **96,** 044020 (2017).

[99] Helvi Witek, Leonardo Gualtieri, Paolo Pani, and Thomas P. Sotiriou, Black holes and binary mergers in scalar Gauss-Bonnet gravity: Scalar field dynamics, Phys. Rev. D **99,** 064035 (2019).

[100] Maria Okounkova, Leo C. Stein, Mark A. Scheel, and Saul A. Teukolsky, Numerical binary black hole collisions in dynamical Chern-Simons gravity, Phys. Rev. D **100,** 104026 (2019).

[101] Maria Okounkova, Leo C. Stein, Jordan Moxon, Mark A. Scheel, and Saul A. Teukolsky, Numerical relativity simulation of GW150914 beyond general relativity, Phys. Rev. D **101,** 104016 (2020).

[102] Justin L. Ripley and Frans Pretorius, Scalarized black hole dynamics in Einstein dilaton Gauss-Bonnet gravity, Phys. Rev. D **101,** 044015 (2020).

[103] Félix-Louis Julié and Emanuele Berti, $d+1$ formalism in Einstein-scalar-Gauss-Bonnet gravity, Phys. Rev. D **101,** 124045 (2020).

[104] Helvi Witek, Leonardo Gualtieri, and Paolo Pani, Towards numerical relativity in scalar Gauss-Bonnet gravity: $3+1$ decomposition beyond the small-coupling limit, Phys. Rev. D **101,** 124055 (2020).

[105] William E. East and Justin L. Ripley, Evolution of Einstein-scalar-Gauss-Bonnet gravity using a modified harmonic formulation, Phys. Rev. D **103,** 044040 (2021).

[106] Hector O. Silva, Helvi Witek, Matthew Elley, and Nicolás Yunes, Dynamical descalarization in binary black hole mergers, Phys. Rev. Lett. **127,** 031101 (2021).

[107] William E. East and Justin L. Ripley, Dynamics of spontaneous black hole scalarization and mergers in Einstein-Scalar-Gauss-Bonnet gravity, Phys. Rev. Lett. **127,** 101102 (2021).

[108] Matthew Elley, Hector O. Silva, Helvi Witek, and Nicolás Yunes, Spin-induced dynamical scalarization, descalariza-tion, and stealthness in scalar-Gauss-Bonnet gravity during a black hole coalescence, Phys. Rev. D **106,** 044018 (2022).

[109] Justin L. Ripley, Numerical relativity for Horndeski gravity, Int. J. Mod. Phys. D **31,** 2230017 (2022).

[110] Tamara Evstafyeva, Michalis Agathos, and Justin L. Ripley, Measuring the ringdown scalar polarization of gravitational waves in Einstein-scalar-Gauss-Bonnet grav-ity, Phys. Rev. D **107,** 124010 (2023).

[111] Maxence Corman, Justin L. Ripley, and William E. East, Nonlinear studies of binary black hole mergers in Einstein-scalar-Gauss-Bonnet gravity, Phys. Rev. D **107,** 024014 (2023).

[112] Félix-Louis Julié and Nathalie Deruelle, Two-body prob-lem in scalar-tensor theories as a deformation of general relativity: An effective-one-body approach, Phys. Rev. D **95,** 124054 (2017).

[113] Félix-Louis Julié, On the motion of hairy black holes in Einstein-Maxwell-dilaton theories, J. Cosmol. Astropart. Phys. 01 (2018) 026.

[114] Félix-Louis Julié, Gravitational radiation from compact binary systems in Einstein-Maxwell-dilaton theories, J. Cosmol. Astropart. Phys. 10 (2018) 033.

[115] Félix-Louis Julié and Emanuele Berti, Post-Newtonian dynamics and black hole thermodynamics in Einstein-scalar-Gauss-Bonnet gravity, Phys. Rev. D **100,** 104061 (2019).

[116] Banafsheh Shiralilou, Tanja Hinderer, Samaya Nissanke, Néstor Ortiz, and Helvi Witek, Nonlinear curvature effects in gravitational waves from inspiralling black hole bina-ries, Phys. Rev. D **103,** L121503 (2021).

[117] Banafsheh Shiralilou, Tanja Hinderer, Samaya M. Nissanke, Néstor Ortiz, and Helvi Witek, Post-Newtonian gravitational and scalar waves in scalar-Gauss–Bonnet gravity, Classical Quantum Gravity **39,** 035002 (2022).

[118] Félix-Louis Julié, Vishal Baibhav, Emanuele Berti, and Alessandra Buonanno, Third post-Newtonian effective-one-body Hamiltonian in scalar-tensor and Einstein-scalar-Gauss-Bonnet gravity, Phys. Rev. D **107,** 104044 (2023).

[119] Iris van Gemeren, Banafsheh Shiralilou, and Tanja Hinderer, Dipolar tidal effects in gravitational waves from scalarized black hole binary inspirals in quadratic gravity, Phys. Rev. D **108,** 024026 (2023).

[120] Richard Brito, Alessandra Buonanno, and Vivien Raymond, Black-hole spectroscopy by making full use of gravitational-wave modeling, Phys. Rev. D **98,** 084038 (2018).

[121] Abhirup Ghosh, Richard Brito, and Alessandra Buonanno, Constraints on quasinormal-mode frequencies with

LIGO-Virgo binary–black-hole observations, Phys. Rev. D **103**, 124041 (2021).

[122] Hector O. Silva, Abhirup Ghosh, and Alessandra Buonanno, Black-hole ringdown as a probe of higher-curvature gravity theories, Phys. Rev. D **107**, 044030 (2023).

[123] Elisa Maggio, Hector O. Silva, Alessandra Buonanno, and Abhirup Ghosh, Tests of general relativity in the nonlinear regime: A parametrized plunge-merger-ringdown gravitational waveform model, Phys. Rev. D **108**, 024043 (2023).

[124] http://rockfish.jhu.edu.