

Black hole spectroscopy horizons for current and future gravitational wave detectors

Iara Ota^{1,*} and Cecilia Chirenti^{2,3,4,5,†}

¹*Centro de Ciências Naturais e Humanas, UFABC, Santo André, SP 09210-170, Brazil*

²*Department of Astronomy, University of Maryland, College Park, Maryland 20742, USA*

³*Astroparticle Physics Laboratory NASA/GSFC, Greenbelt, Maryland 20771, USA*

⁴*Center for Research and Exploration in Space Science and Technology, NASA/GSFC, Greenbelt, Maryland 20771, USA*

⁵*Center for Mathematics, Computation and Cognition, UFABC, Santo André - SP, 09210-580, Brazil*

Black hole spectroscopy is the proposal to observe multiple quasinormal modes in the ringdown of a binary black hole merger. In addition to the fundamental quadrupolar mode, overtones and higher harmonics may be present and detectable in the gravitational wave signal, allowing for tests of the no-hair theorem. We analyze in detail the strengths and weaknesses of the standard Rayleigh criterion supplied with a Fisher matrix error estimation, and we find that the criterion is useful, but too restrictive. Therefore we motivate the use of a conservative high Bayes factor threshold to obtain the black hole spectroscopy horizons of current and future detectors, i.e., the distance (averaged in sky location and binary inclination) up to which one or more additional modes can be detected and confidently distinguished from each other. We set up all of our searches for additional modes starting at $t = 10(M_1 + M_2)$ after the peak amplitude in simulated signals of circular nonspinning binaries. An agnostic multimode analysis allows us to rank the subdominant modes: for nearly equal mass binaries we find $(\ell, m, n) = (2, 2, 1)$ and $(3, 3, 0)$ and, for very asymmetric binaries, $(3, 3, 0)$ and $(4, 4, 0)$, for the secondary and tertiary modes, respectively. At the current estimated rates for heavy stellar mass binary black hole mergers, with primary masses between 45 and 100 solar masses, we expect an event rate of mergers within our conservative estimate for the $(2, 2, 1)$ spectroscopy horizon of $0.03 - 0.10 \text{ yr}^{-1}$ for LIGO at design sensitivity and $(0.6 - 2.4) \times 10^3 \text{ yr}^{-1}$ for the future third generation ground-based detector Cosmic Explorer.

I. INTRODUCTION

The increasing number of detected binary black hole (BH) mergers has turned gravitational wave (GW) astronomy into a multi-faceted new field, with important ramifications for astrophysics and cosmology, among other disciplines. In the future, low-frequency ($\lesssim 1\text{Hz}$) and high-frequency ($\gtrsim 1\text{kHz}$) GW astronomers may become as specialized as their current radio and X-ray counterparts. However, some physical features will be the same over the entire frequency spectrum and remain relevant for all future GW astronomers. According to Einstein's theory of General Relativity (GR), BHs are *scale-invariant* (although corrections to GR need not be [1]): the properties of Kerr spacetime are the same for the lightest stellar mass BH [2] up to the heaviest behemoth imaginable [3].

These properties are encoded in the BH spectrum of discrete quasinormal mode (QNM) complex frequencies $\omega_{\ell mn} = \omega_{\ell mn}^r + i\omega_{\ell mn}^i$ [4–6], where the numbers (ℓ, m) identify the angular symmetry of each mode and the number n is the tone. The no-hair theorem guarantees that astrophysical BHs are described exclusively by their mass M and spin parameter a ; there is no additional fine print in GR. As a consequence, the detection of a single QNM is sufficient to provide two equations

$$M = M(\omega_{\ell mn}^r, \omega_{\ell mn}^i) \quad \text{and} \quad a = a(\omega_{\ell mn}^r, \omega_{\ell mn}^i) \quad (1)$$

to completely determine the BH spacetime. At least in the classical theory, BHs in a vacuum hold no secrets. Such measurements are already possible with the LIGO/Virgo observations [7, 8]. We can highlight the detections of GW150914 [9] (the first and still strongest binary BH event to date) and GW190521 [10] (which featured the heaviest merging BHs observed so far). The fundamental quadrupolar QNM, i.e. the mode labeled with $(\ell, m, n) = (2, 2, 0)$, was measured in the post-merger ringdown of both events. The properties of the merger remnant obtained in this way can be compared with the results from the analysis of the inspiral part of the GW signal, providing a consistency check of GR [11–13].

But the GW community is eager to probe deeper into the nature of BHs. The confident detection of a secondary, or subdominant, QNM would provide additional information to allow an independent test of the no-hair theorem through BH spectroscopy [14–18]. Higher harmonic modes with $(\ell, m) \neq (2, 2)$ become more relevant for more unequal mass binaries, with large mass ratio $q = M_1/M_2 \geq 1$. Moreover, overtone modes ($n > 0$) decay much faster than fundamental modes ($n = 0$). Therefore, most studies focused on performing spectroscopy add only fundamental higher harmonic modes, i.e., $(\ell, m, 0) \neq (2, 2, 0)$ [19–29], even though the importance of overtones has been known for years [30, 31] and a few more recent works already indicated the relevance of the first overtone of the quadrupolar mode, labeled $(2, 2, 1)$ [32–34].

The recent studies of the importance of higher overtones in the ringdown modeling [34, 35] and some sug-

* iara.ota@ufabc.edu.br

† chirenti@umd.edu

gestive evidence for the presence of the (2, 2, 1) mode in GW150914 [36] increased interest in the contribution of overtones [37–46]. More recently, some works extended the ringdown modeling analysis to include mirror modes [39, 43, 47, 48].

It is natural to ask when we will be confidently able to do BH spectroscopy. In particular, will high-precision BH spectroscopy be attainable with LIGO? Motivated by the ringdown horizons computed by [23], we present here the first detailed study of the *BH spectroscopy horizon*, which measures how far an event can be, as a function of the remnant mass, binary mass ratio and the detector being considered, for two or more QNMs to be detectable in the GW ringdown.¹ It is important to stress that we are not performing tests of GR in our work. Rather, we are assessing the detectability of a secondary mode, assuming that GR is the correct theory of gravity. All of our analyses start at a conservative time $t = 10(M_1 + M_2)$ after the signal peak.

We probe this topic with two different approaches. First, we use a Rayleigh criterion [15, 49] coupled with a Fisher matrix analysis [50] to estimate the spectroscopy horizon for LIGO at design sensitivity [51], future space detector LISA [52] and proposed third generation ground-based detectors Einstein Telescope (ET) [53] and Cosmic Explorer (CE) [54] (see Section II). Second, we use a fully Bayesian analysis to verify our results for LIGO and CE, using a well-motivated and conservative Bayes factor threshold (Section III). We confirm that the Rayleigh criterion is too restrictive (see Section IV), and find promising rates of detection with third generation detectors (and LISA). The LIGO rates are small, but not prohibitively so. We present our conclusion in Section V and also explore other prescriptions for mode resolvability in the Appendix A.

II. BH SPECTROSCOPY HORIZONS FROM THE RAYLEIGH CRITERION

Our goal is to determine whether two modes are detectable in the ringdown of a binary BH coalescence at a given luminosity distance D_L , with binary component masses $M_1 > M_2$ resulting in a merger remnant black hole of mass M and spin parameter a .

The ringdown gravitational waveform can be written

as a sum of QNMs

$$h_+ + ih_\times = \frac{M}{D_L} \sum_{\ell mn} A_{\ell mn} e^{i[\omega_{\ell mn}(t-t_0) - \phi_{\ell mn}]} \times {}_{-2}S_{\ell m}(a\omega_{\ell mn}, \iota, \beta), \quad (2)$$

for $t > t_0$, a time close but not necessarily equal to the merger time of the binary. Here h_+ and h_\times are the two polarizations of the wave, $A_{\ell mn}$ and $\phi_{\ell mn}$ are the initial amplitude and phase of each mode, respectively, $\omega_{\ell mn} = \omega_{\ell mn}^r + i\omega_{\ell mn}^i$ are the complex frequencies of the modes, ${}_{-2}S_{\ell m}(a\omega_{\ell mn}, \iota, \beta)$ are the spin-weighted spheroidal harmonics with spin $s = -2$, and (ι, β) are the inclination and azimuth angles of the binary relative to the detector. The frequencies of oscillation and damping times of the modes are given by

$$f_{\ell mn} = \frac{\omega_{\ell mn}^r}{2\pi} \quad \text{and} \quad \tau_{\ell mn} = \frac{1}{\omega_{\ell mn}^i}. \quad (3)$$

The physical quantities shown in eqs. (2) and (3) are measured in the source frame. The corresponding values measured in the detector frame are rescaled by the cosmological redshift z : $M^d = M(1+z)$, $f_{\ell mn}^d = f_{\ell mn}/(1+z)$ and $\tau_{\ell mn}^d = \tau_{\ell mn}(1+z)$, where the superscript “d” indicates that the quantities are measured in the detector frame. Whenever required, we use values of the cosmological parameters given in [55].

The waveform measured by the detector is given by [56]

$$h(t) = F_+(\theta, \phi, \psi)h_+(t) + F_\times(\theta, \phi, \psi)h_\times(t), \quad (4)$$

where $h_{+,\times}(t)$ are the wave polarizations defined in equation (2) with quantities measured in the detector frame, $F_{+,\times}(\theta, \phi, \psi)$ are the detector’s antenna patterns, (θ, ϕ) indicate the source sky location and ψ is the polarization angle. We use the full-sky angular average of the spin-weighted spheroidal harmonics $\langle |{}_{-2}S_{\ell m}| \rangle = 1/\sqrt{4\pi}$ and of the antenna patterns $\langle (F_{+,\times}^2)^{\frac{1}{2}} \rangle$, which are equal to $1/\sqrt{5}$ for L shaped detectors (LIGO and CE) [56] and $3/(2\sqrt{5})$ for ET [57] and does not have a closed analytical form for LISA [58].

We also need to define the single detector signal-to-noise ratio (SNR), given by

$$\rho^2 = 4 \int_0^\infty \frac{|\tilde{h}(f)|^2}{S_n(f)} df, \quad (5)$$

where the symbol $(\tilde{})$ denotes the Fourier transform of the function and $S_n(f)$ is the noise spectral density of the detector. In Figure 1 we show $S_n(f)$ for LIGO (at design sensitivity) and future detectors CE, ET and LISA.

For our spectroscopic analysis, we require that each mode (ℓ, m, n) has a single detector SNR $\rho_{\ell mn} > 8$ and appraise their detectability using a Rayleigh criterion (Section II A) with errors estimated by a Fisher matrix analysis (Section II B). The spectroscopy horizons are presented in Section II C.

¹ We define the BH spectroscopy horizon with the angular average of sky location and binary inclination (see Section II). In the LIGO literature, a “horizon” is usually defined as the maximum distance obtained with optimal values of the angular parameters, while a “range” refers to the angular average. Therefore our horizon definition is equivalent to the range definition used by LIGO.

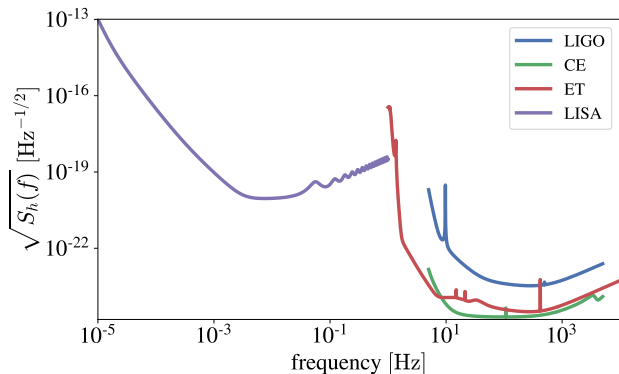


Figure 1. Noise spectral density for current and future GW detectors used in our study: LIGO (at design sensitivity) [51], proposed third generation ground-based detectors Cosmic Explorer (CE) [59] and Einstein Telescope (ET) [53, 60] and space-based detector LISA [52] (planned launch in mid 2030’s).

A. Rayleigh criterion and caveats

The original Rayleigh criterion serves the purpose of establishing whether two close stars are resolvable in an optical telescope. In direct analogy, Berti et al. [15] proposed an adaptation of the criterion to determine whether two quasinormal modes (ℓ, m, n) and (ℓ', m', n') can be resolved in a GW detection. The required conditions are:

$$\Delta f_{\ell mn, \ell' m' n'} > \max(\sigma_{f_{\ell mn}}, \sigma_{f_{\ell' m' n'}}), \quad (6a)$$

$$\Delta \tau_{\ell mn, \ell' m' n'} > \max(\sigma_{\tau_{\ell mn}}, \sigma_{\tau_{\ell' m' n'}}), \quad (6b)$$

where we use the definitions

$$\Delta f_{\ell mn, \ell' m' n'} \equiv |f_{\ell mn} - f_{\ell' m' n'}|, \quad (7a)$$

$$\Delta \tau_{\ell mn, \ell' m' n'} \equiv |\tau_{\ell mn} - \tau_{\ell' m' n'}|, \quad (7b)$$

and we assume that each frequency and damping time are reported from the observations as $f_{\ell mn} \pm \sigma_{f_{\ell mn}}$ and $\tau_{\ell mn} \pm \sigma_{\tau_{\ell mn}}$, respectively. This criterion has been used in different works to estimate the resolvability of two QNMs [37–39, 49].

We choose $(\ell, m, n) = (2, 2, 0)$ (the fundamental mode), and let (ℓ', m', n') be a subdominant mode, that is, either the first overtone $(2, 2, 1)$ or one of the fundamental higher harmonics. If *both* conditions (6) are satisfied, then we can say that both pairs of frequencies and damping times are independently determined. This allows a test of the no-hair theorem using eq. (1) for each mode.

As we can see from Fig. 2, the Rayleigh conditions (6) pose different challenges for the overtone (condition (6b) is easily satisfied) and the higher harmonics with $\ell \neq 2$ (condition (6a) is easily satisfied). There is also a dependence on the mass ratio of the progenitor BH binary: for nonspinning circular binaries with a fixed total mass, the

final spin a of the merger remnant will depend only on the binary mass ratio, due to its relation to the orbital angular momentum. Therefore the final spin is higher for equal mass binaries.

For modes with $\ell = m$, both $\Delta f_{\ell mn, \ell' m' n'}$ and $\Delta \tau_{\ell mn, \ell' m' n'}$ increase with q , which helps satisfy the Rayleigh criterion (6). But binaries with higher mass ratio excite the dominant mode with lower amplitude and emit less gravitational energy for a fixed total mass (see Table I), leading to lower SNR and larger measurement errors $\sigma_{f_{\ell mn}}$ and $\sigma_{\tau_{\ell mn}}$. The combination of these effects results in a *smaller* spectroscopy horizon of the modes with $\ell = 2$ for higher mass ratios. For the higher harmonics with $\ell = m \neq 2$ the spectroscopy horizon is approximately the same for low and high mass ratio, as we show in Section II C.

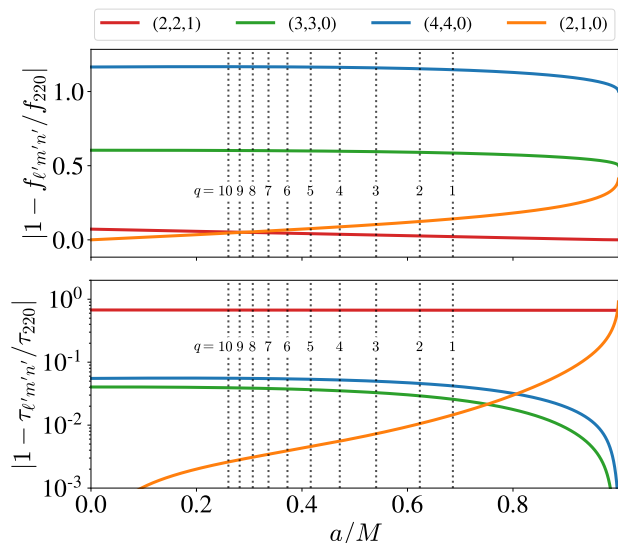


Figure 2. Relative differences of frequencies (top) and damping times (bottom) between the dominant mode $(2, 2, 0)$ and the most relevant subdominant modes as a function of the final black hole dimensionless spin a/M , obtained with linear perturbation theory [15, 61]. The vertical dotted lines indicate the final dimensionless spin of the remnant of nonspinning circular binaries with mass ratios q ranging from 1 (highest spin) to 10 (lowest spin). The modes with $\ell = 2$ have frequencies very similar to the $(2, 2, 0)$ frequency, whereas modes with $\ell \neq 2$ have more distinct (higher) frequencies, which are more easily resolvable with eq. (6a). The overtone has a lower damping time that is more easily resolvable with eq. (6b) than any of the fundamental modes, which have similar damping times to the $(2, 2, 0)$ mode.

Two caveats must be mentioned about the Rayleigh criterion. First, the standard implementation we present here gives an either-or answer (the two modes are either resolvable or not) and permits no gradation. (See also a discussion of variations of the Rayleigh criterion that are equivalent to different confidence levels in Appendix A.) Second, the criterion could be inherently *too restrictive*

for predicting the detectability of a secondary mode, i.e., it could be a stronger condition than what is actually needed for identifying a secondary mode in the data with a reasonable level of confidence (see Sec. IV).² Previous work showed that the minimum ringdown SNR needed to resolve both Rayleigh conditions is very high [15, 37, 38] (SNR ≈ 100). On the other hand, the minimum SNR for satisfying just one condition of the Rayleigh criterion may not be restrictive enough for a confident detection, and usually needs to be complemented with additional conditions, such as a minimum SNR threshold [15, 39, 62].

B. Fisher matrix analysis

For a high SNR measurement, the probability distribution of each parameter ϑ^a of the waveform will be a Gaussian centered at the real value $\vartheta_{\text{real}}^a$. Consequently, the estimated value for parameter ϑ^a will be close to the real value $\vartheta_{\text{real}}^a$, that is, $\vartheta^a = \vartheta_{\text{real}}^a + \delta\vartheta^a$ and the probability density distribution for $\delta\vartheta^a$ is proportional to the multidimensional Gaussian $\exp[-(1/2)\Gamma_{ab}\vartheta^a\vartheta^b]$. The statistical error in the determination of ϑ^a is

$$\sigma_{\vartheta^a} = \sqrt{(\Gamma^{-1})^{aa}}, \quad (8)$$

where Γ^{-1} is the inverse of the Fisher matrix [50], given by

$$\Gamma_{ab} \equiv \left\langle \frac{\partial h}{\partial \vartheta^a} \middle| \frac{\partial h}{\partial \vartheta^b} \right\rangle, \quad (9)$$

with the noise-weighted inner product defined as

$$\langle h_1 | h_2 \rangle \equiv 2 \int_0^\infty \frac{\tilde{h}_1^*(f)\tilde{h}_2(f) + \tilde{h}_1(f)\tilde{h}_2^*(f)}{S_n(f)} df, \quad (10)$$

and the symbol (*) denotes the complex conjugate of the function.

The errors $\sigma_{f_{\ell mn}}$ and $\sigma_{\tau_{\ell mn}}$ needed for the Rayleigh conditions (6) can be estimated in this formalism. We choose as our relevant parameters the QNM intrinsic parameters

$$\vartheta^a = \{f_{\ell mn}, \tau_{\ell mn}, A_{\ell mn}, \phi_{\ell mn}, f_{\ell' m' n'}, \tau_{\ell' m' n'}, R, \phi_{\ell' m' n'}\},$$

where $R = A_{\ell' m' n'}/A_{\ell mn}$ is the amplitude ratio between the modes, which gives us 8 parameters for a 2-mode analysis. We take the full-sky angular average for the

sky localization and binary inclination angle relative to the detector (see discussion following eq. (4)).

Following [15], we use the doubling prescription by Flanagan and Hughes [63] to compute the Fourier transform of $h(t)$. That is, we first obtain $h(t)$ for all t by reflecting the QNM waveform at t_0 ; this procedure avoids the spectral leakage at high frequencies that would be caused by a discontinuity in the waveform. The partial derivatives in equation (9) are obtained analytically and the noise-weighted inner product (10) and statistical errors given by (8) are computed numerically.

We use numerical relativity simulations of binary BH mergers from the Simulating eXtreme Spacetimes project (SXS) [64, 65] to fit the QNM amplitudes and phases for each waveform analyzed here. The procedure for obtaining these fits is detailed in [38], where we examined their dependence on the binary mass ratio. In short, the fundamental modes $(\ell, m, 0)$ are fitted starting at the earliest time when the contribution of overtones and non-linearities can be neglected; the overtones are fitted at the time when the waveform is best described by the fundamental mode and the first overtone. The initial amplitudes and phases are then rescaled to the time $t = t_{\text{peak}} + 10(M_1 + M_2)$, where t_{peak} is the time of maximum amplitude of the quadrupolar mode $(\ell, m) = (2, 2)$. In Table I we show the parameters of the QNMs obtained from simulations SXS:BBH:0593 ($q = 1.5$) and SXS:BBH:1107 ($q = 10$) [65], which we analyze in more detail in the following sections.

Table I. QNM amplitudes, phases, and frequencies (defined in eq. (2)) for nonspinning circular binaries with mass ratio $q = 1.5$ and $q = 10$. The amplitudes and phases are reported at time $t = t_{\text{peak}} + 10(M_1 + M_2)$. The values for the amplitude and phase of the $(2, 2, 1)$ mode are obtained with method II as described in [38].

(ℓ, m, n)	$A_{\ell mn}$	$\phi_{\ell mn}$ [rad]	$\omega_{\ell mn}^r MG/c^3$	$\omega_{\ell mn}^i MG/c^3$
$q = 1.5$				
(2, 2, 0)	0.40	0.41	0.517	0.082
(2, 2, 1)	0.28	4.59	0.505	0.248
(3, 3, 0)	0.05	6.16	0.821	0.084
(4, 4, 0)	0.01	5.46	1.112	0.086
(2, 1, 0)	0.03	5.26	0.448	0.083
$q = 10$				
(2, 2, 0)	0.14	3.56	0.412	0.088
(2, 2, 1)	0.07	1.17	0.390	0.269
(3, 3, 0)	0.05	3.32	0.661	0.092
(4, 4, 0)	0.02	3.32	0.894	0.093
(2, 1, 0)	0.05	5.57	0.394	0.088

² Both caveats were also addressed by [49], where a generalized likelihood ratio test (GLRT) was proposed to determine whether a secondary mode is present or not in a signal, finding a SNR higher than the SNR needed to resolve a single Rayleigh condition but smaller than the SNR needed to resolve both conditions, see their Appendix B. This is similar to the approach we use in Section III with a Bayes factor threshold.

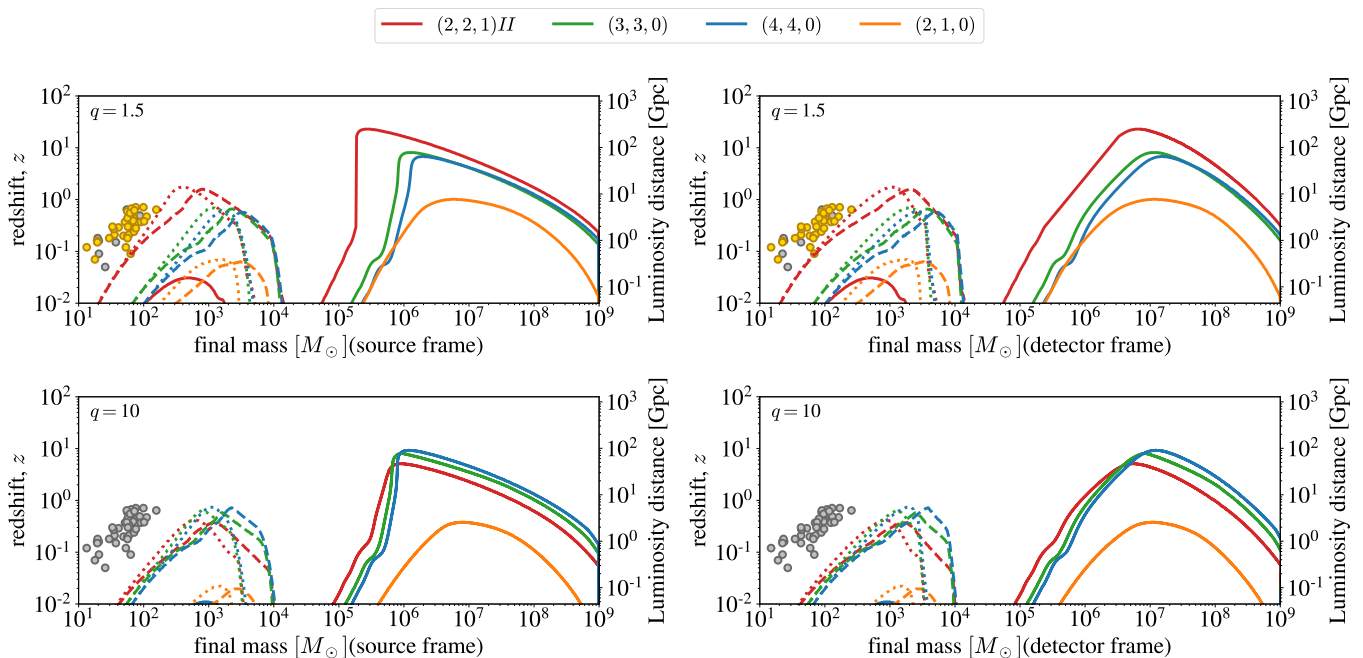


Figure 3. BH Spectroscopy horizons obtained with the Rayleigh criterion (6) and Fisher matrix error estimation with 8 parameters. Top (bottom) panels show results for a BH merger remnant from a binary with a low (high) mass ratio $q = 1.5$ ($q = 10$). The left (right) panels show the BH spectroscopy horizons as a function of the BH remnant mass in the source (detector) frame. Solid lines indicate the horizons for LIGO (lower masses) and LISA (higher masses); dashed (dotted) lines indicate the ET (CE) horizons. Circles show detections from GWTC-1 and GWTC-2; yellow circles are compatible with the mass ratio in each case, whereas gray circles are not. Ringdown modes $(2, 2, 0) + (\ell, m, n)$ will be detectable and resolvable for events that happen below the corresponding $(2, 2, 0) + (\ell, m, n)$ spectroscopy horizon curve.

C. BH spectroscopy horizons

Here we present the BH spectroscopy horizon, $z_{\ell mn}^{\text{spec,R}}$, which is the maximum distance, averaged over sky location and binary inclination, up to which a secondary mode (ℓ, m, n) can be detected and resolved in the signal, using the Rayleigh conditions (6) supplied with errors given by the Fisher matrix formalism (8). The spectroscopy horizon provides a figure of merit which allows us to compare results between detectors and across several orders of magnitude in mass range.

We propose the use of the source distance as a more universal measure for assessing the prospects of BH spectroscopy than the SNR needed for detection, which depends on the detector and the source.³ In short, the SNR is a good measure to determine detectability thresholds, but it is not the best option to use when comparing ringdown signals with different subdominant modes. If all else is kept equal (binary mass ratio, remnant mass, distance, etc) and we inform the properties of the modes (initial amplitudes and phases) with NR simulations,

then we find that two signals with different subdominant modes will have different SNRs at the same distance. Therefore, a fixed SNR threshold is not a good equalizer: it requires different source distances for different subdominant modes, leading to different detection rates. This problem is solved by quoting directly the source distance, which we calculate as detailed below to find the BH spectroscopy horizon.

Figure 3 shows the spectroscopy horizons for distinguishing the dominant mode $(2, 2, 0)$ from the most relevant subdominant modes as a function of the remnant BH final mass for nonspinning circular binaries with low and high mass ratios. In the area inside the curves, each of the two modes has SNR greater than 8 and they satisfy the Rayleigh criterion (6). The circles show the confirmed binary BH events to date [7, 8]. In figure 4 we have the spectroscopy horizons just for LIGO, which were mostly not visible in Figure 3. The LIGO horizons have similar trends as the proposed third generation gravitational wave detectors ET and CE.

Some general features can be seen for all of the detectors. The spectroscopy horizon is small for low masses due to the small amplitude of the QNMs in these systems, as the amplitude is proportional to the BH mass. Systems with larger mass have larger amplitudes, however, as the frequencies are proportional to the inverse of the mass, the SNR decreases as the mode frequencies

³ In [37], for instance, it was shown that the ringdown SNR of a signal containing the fundamental mode plus the overtone could be lower than the SNR of the fundamental mode alone, depending on the relative phases of the modes.

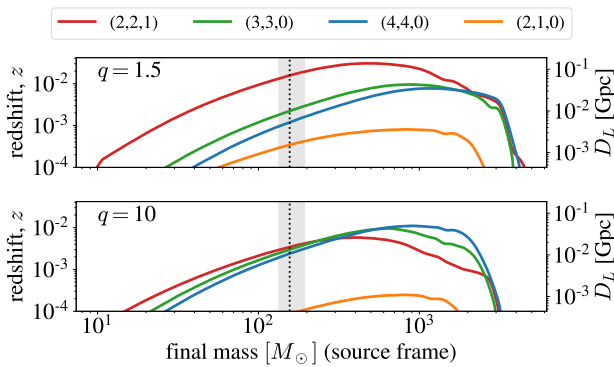


Figure 4. Same as Figure 3, but for the LIGO spectroscopy horizons (at lower distances). The gray band corresponds to the estimated remnant mass of GW190521. The largest corresponding spectroscopy horizon, however, is approximately 10 times closer than the estimated distance to the event.

approach the detector’s low frequency sensitivity limit. The SNR of the subdominant modes is larger than approximately 30 in all cases, therefore automatically satisfying our detectability criterion ($\text{SNR} > 8$). Although we do not impose a separate “measurability” criterion, see [39], we verify that $\sigma_{f_{\ell mn}}/f_{\ell mn} \lesssim 0.1$ in all cases. The damping time is typically much more uncertain, and $\sigma_{\tau_{\ell mn}}/\tau_{\ell mn} \lesssim 0.4$.

For the ground-based detectors, the spectroscopy horizon of the overtone (2,2,1) is up to ~ 10 -80 times larger than the spectroscopy horizon of the higher harmonics, for lower masses and low mass ratio. This result is compatible with the minimum SNR analysis of an event similar to GW150914 in [38]. For low mass ratio and higher masses, the horizons become comparable (except for the (2,1,0), which is still lower). This is again the case for the high mass ratio case, although we can see that the (2,2,1) spectroscopy horizon is slightly larger for lower masses, while the spectroscopy horizons of the higher harmonics (4,4,0) and (3,3,0) are larger for larger masses. The larger horizons for higher harmonics with $\ell \neq 2$ in the large mass range are expected, as these modes have larger frequencies than the modes with $\ell = 2$ (see Table I).

It is important to notice that the most massive event detected so far is GW190521 [10], with final mass $M = 156.3^{+36.8}_{-22.4} M_{\odot}$ [8]. This is also the most massive stellar-mass black hole detected to date. As the existence of even more massive stellar-mass black holes is still uncertain, the Rayleigh criterion analysis with Fisher matrix errors shown in Figures 3 and 4 indicates that, for sources compatible with the current detections and ground-based detectors, it will always be easier to distinguish the overtone from the dominant mode than any of the higher harmonics, even for higher mass ratios. This result is strongly influenced by the requirement that both Rayleigh conditions should be satisfied, which poses a considerable penalty on the higher harmonics, due to the condition on the damping times. (See Section III B, for the results

of our complementary Bayesian analysis.) For masses larger than the largest stellar-mass black hole detected to date, the higher harmonics (3,3,0) and (4,4,0) have larger spectroscopy horizons than the overtone.

For LISA the situation is still uncertain, as the properties of the population of supermassive binaries are largely unknown. Therefore we cannot safely estimate which subdominant mode will be more easily detected (although the overtone has a larger spectroscopy horizon in more cases). But the LISA spectroscopy horizons are very large and it is probable that more than one extra mode will be detected in a supermassive binary BH merger. There are two noticeable differences between the general trends of the spectroscopy horizons for LISA when compared with the ground-based detectors. For low mass ratios, the (2,2,1) has the largest spectroscopy horizon for all masses, as a result of the expected flatter behavior of the LISA noise curve for high frequencies (see Figure 1). Additionally, we can see the spectroscopy horizons grow almost vertically with the distance for lower masses in the source frame. This results from the interplay between the cosmological redshift and the detector’s colored noise: systems with lower masses have higher frequencies, which are redshifted back to the detectors’ most sensitive band. This does not happen in the detector frame because there the masses are also rescaled with the cosmological redshift.

Variations of the Rayleigh criteria we have used in this section have been proposed in other works. For example, [39] proposed the use of the quality factor $Q_{\ell mn} = \pi f_{\ell mn} \tau_{\ell mn}$ instead of the damping time $\tau_{\ell mn}$ in the Rayleigh criterion. Requiring only one of the two conditions of the Rayleigh criterion was suggested by other works such as [39, 62]. More recently, [45] also reexamined the Rayleigh criterion and put forward a less restrictive condition based on the distinguishability of the 2-dimensional posteriors for $f_{\ell mn}$ and $\tau_{\ell mn}$ for two modes. In Appendix A we present the spectroscopy horizons that result from these various criteria.

III. BAYESIAN INFERENCE AND MODEL COMPARISON

In this Section we use a fully Bayesian statistical approach to obtain the BH spectroscopy horizons, as an alternative to the Rayleigh criterion used in Section II. As we briefly explain in the next subsection, our method uses a threshold on the Bayes factor to confirm the presence of one or more secondary modes in the ringdown signal.

A. Bayesian formalism and tests

The posterior probability distribution for a set of parameters ϑ of a model \mathcal{M} , given the detected or assumed

data d , is

$$p(\vartheta|d, \mathcal{M}) = \frac{\mathcal{L}(d|\vartheta; \mathcal{M})\pi(\vartheta; \mathcal{M})}{\mathcal{Z}_{\mathcal{M}}}, \quad (11)$$

where $\pi(\vartheta; \mathcal{M})$ is the prior probability distribution for the parameters ϑ which encodes previously known information about the parameters, $\mathcal{L}(d|\vartheta; \mathcal{M})$ is the likelihood of observing the data d given the parameters ϑ of the model \mathcal{M} and $\mathcal{Z}_{\mathcal{M}}$ is the evidence, which is the normalization factor of the posterior, given by

$$\mathcal{Z}_{\mathcal{M}} = \int \mathcal{L}(d|\vartheta; \mathcal{M})\pi(\vartheta; \mathcal{M})d\vartheta. \quad (12)$$

To quantify the preference for one model \mathcal{M}_A over another model \mathcal{M}_B , we use the Bayes factor defined as

$$\mathcal{B}_B^A = \frac{\mathcal{Z}_{\mathcal{M}_A}}{\mathcal{Z}_{\mathcal{M}_B}}. \quad (13)$$

Assuming GW detectors with stationary Gaussian noise, the likelihood is given by

$$\mathcal{L}(d|\vartheta; \mathcal{M}) \propto \exp\left[-\frac{1}{2}\langle d - h(\vartheta; \mathcal{M})|d - h(\vartheta; \mathcal{M})\rangle\right], \quad (14)$$

where $h(\vartheta, \mathcal{M})$ is the waveform of the model \mathcal{M} with parameters ϑ .

If we assume that the data d contain the detector noise n and two modes, the dominant quadrupolar mode $(2, 2, 0)$ and a subdominant mode (ℓ, m, n) , we have $d = d_2$ given by

$$d_2 = n + h_{220} + h_{\ell mn}. \quad (15)$$

We choose not to use numerical simulation data because we wish to avoid any non-physical numerical noise and the spectral leakage at high frequencies introduced by taking the Fourier transform of the time-domain data with a rectangular window. The spectral leakage can be suppressed by windowing the signal with a smooth ‘‘turn on’’, but the QNMs decay very quickly (especially the overtone) and a smooth window could remove the most relevant part of the signal, which is not desirable. For real data, a time domain analysis may be preferable over the frequency domain analysis (see [24, 29, 44, 45]), but it is more computationally expensive and the analytical QNMs are a good approximation of the numerical simulation data for our work.

Following Sec. II B, we compute the analytical Fourier transform of each QNM using the Flanagan and Hughes prescription [63]. We use the noise spectral density $S_n(f)$ to generate independent realizations of the noise $n(f)$. In this Section we focus on the case of the LIGO and CE detectors only, due to computational constraints. Out of the remaining future GW detectors considered in this work, we can see from Figure 3 that ET and CE will have very similar horizon curves (ET’s extends slightly to higher masses, or lower frequencies, see Figures 1 and 3) and that LISA’s horizon will extend far enough for multiple modes to be detected in the ringdown of a supermassive BH binary.

Ringdown models The models considered in our analysis, which we suppose to be valid starting at $t = t_{\text{peak}} + 10(M_1 + M_2)$, are

- \mathcal{M}_1 , a *single-mode* model with 4 parameters $\vartheta_1 = \{A, \phi_{220}, f_{220}, \tau_{220}\}$,
- \mathcal{M}_2 , a *two-mode* model with 8 parameters $\vartheta_2 = \{A, \phi_{220}, f_{220}, \tau_{220}, R_{\ell mn}, \phi_{\ell mn}, f_{\ell mn}, \tau_{\ell mn}\}$,

where $R_{\ell mn} = A_{\ell mn}/A$ is the amplitude ratio between a specified subdominant mode (ℓ, m, n) and the dominant mode $(2, 2, 0)$. The global amplitude parameter A is proportional to M/D_L (see eq. (2)) and all model parameters are defined in the *detector frame*. An agnostic multimode analysis with two and three modes is presented in Sec. III C.

Prior distributions As the spectroscopic analysis will look for small contributions to an already detected signal, we consider prior distributions informed by the properties of the binary estimated by the complete inspiral-merger-ringdown (IMR) analysis. The ringdown parameters in the detector frame depend strongly on the final mass M and redshift z , which we assume to be estimated with $\pm 50\%$ errors at 90% credibility, that is, within $[M_{\min}, M_{\max}] = [0.5M, 1.5M]$ and $[z_{\min}, z_{\max}] = [0.5z, 1.5z]$ for each event.⁴ There is also a weaker dependence on the final spin a , or, equivalently, on the binary mass ratio q if we assume circular nonspinning binaries (the only parameter that depends strongly on q is $R_{\ell mn}$). For the cases we consider ($q = 1.5$ and $q = 10$) we fix the corresponding value of a (see Figure 2 and Table I). Consequently, we choose conservative prior distributions for the model parameters using their expected GR values in the ranges described below:

- $\pi(A; \mathcal{M}_{1,2})$ is log-uniform in $[A_{\min}, A_{\max}]$ with $A_{\min} = A(M_{\min}, z_{\max}, a)/10$ and $A_{\max} = 10 \times A(M_{\max}, z_{\min}, a)$,
- $\pi(\phi_{220}; \mathcal{M}_{1,2})$ is uniform in $[0, 2\pi]$,
- $\pi(f_{220}; \mathcal{M}_{1,2})$ is log-uniform in $[f_{220}^{\min}, f_{220}^{\max}]$ with $f_{220}^{\min} = f_{220}(M_{\max}, z_{\max}, a)$ and $f_{220}^{\max} = f_{220}(M_{\min}, z_{\min}, a)$,
- $\pi(\tau_{220}; \mathcal{M}_{1,2})$ is uniform in $[\tau_{220}^{\min}, \tau_{220}^{\max}]$ with $\tau_{220}^{\min} = \tau_{220}(M_{\min}, z_{\min}, a)$ and $\tau_{220}^{\max} = \tau_{220}(M_{\max}, z_{\max}, a)$,

for the dominant mode $(2, 2, 0)$ (used in both models \mathcal{M}_1 and \mathcal{M}_2), and

⁴ Note that this interval is not too narrow, as most of the confirmed LIGO events satisfy these conditions for the final mass and redshift [7, 8]. The exceptions are GW190909_114149 (mass and redshift), and 11 other events of the O3a catalogue (redshift only). All the events used in the LIGO ringdown analysis have parameters estimated by the IMR analysis inside the interval considered here [13].

- $\pi(R_{\ell mn}; \mathcal{M}_2)$ is uniform in $[0, 0.9]$,
- $\pi(\phi_{\ell mn}; \mathcal{M}_2)$ is uniform in $[0, 2\pi]$,
- $\pi(f_{\ell mn}; \mathcal{M}_2)$ and $\pi(\tau_{\ell mn}; \mathcal{M}_2)$ are defined in the same way as $\pi(f_{220}; \mathcal{M}_{1,2})$ and $\pi(\tau_{220}; \mathcal{M}_{1,2})$, but for a *specified* secondary mode (ℓ, m, n) in model \mathcal{M}_2 ,

where log-uniform priors (i.e., scale-invariant priors) are chosen for the global amplitude and the frequencies.

In order to determine how well models \mathcal{M}_1 and \mathcal{M}_2 describe the data $d_2 = n + h_{220} + h_{\ell mn}$, we can calculate the deviation between the injected parameters $\vartheta_{\text{inj}}^\alpha$ and the posteriors of the estimated parameters ϑ^α obtained with each model, analogously to a simulation-based calibration procedure [45, 66, 67]. This procedure helps identifying inaccurate computations and/or inconsistencies in model implementation. We present this deviation in units of “ σ ” by calculating the quantile function $Q(x)$ (associated with the normal distribution) of the integral

$$x = \int_{-\infty}^{\vartheta_{\text{inj}}^\alpha} p(\vartheta^\alpha | d_2, \mathcal{M}_{1,2}) d\vartheta^\alpha.$$

For example: if $x \approx 0.159$, then we have $Q(x; p) \approx -1$ (or -1σ deviation), if $x \approx 0.977$, then we have $Q(x; p) \approx 2$ (or $+2 \sigma$ deviation), etc.

In Figure 5 we show these results in the violin plots. For each case we generated 500 signals d_2 , with independent noise realizations and 50 final masses \times 10 redshifts distributed around the spectroscopy horizon for each subdominant mode (see Figure 8). The left-sided violins (in gray) show the distributions of deviations for each injected parameter of the dominant mode with the posteriors estimated with the single-mode model \mathcal{M}_1 . The colored right-sided violins show the same distributions for all of the parameters using the correct two-mode model \mathcal{M}_2 . Model \mathcal{M}_1 estimates the parameters of the fundamental mode with larger bias and a greater spread of deviations to compensate for neglecting the subdominant mode. The bias is not substantial because the subdominant mode can be neglected in half of the signals, which are outside the spectroscopy horizon.⁵ Therefore, both models produce posteriors compatible with the injected parameters in this distribution, but model \mathcal{M}_1 will have significant bias when the secondary mode is strong.

We use the Bayes factor \mathcal{B}_1^2 to quantify the preference for model \mathcal{M}_2 over model \mathcal{M}_1 in a given signal. Additionally, we choose a detectability threshold of $\ln \mathcal{B}_1^2 > 8$, that is, the evidence \mathcal{Z}_2 of the two-mode model must be approximately 3000 greater than the evidence \mathcal{Z}_1 of the

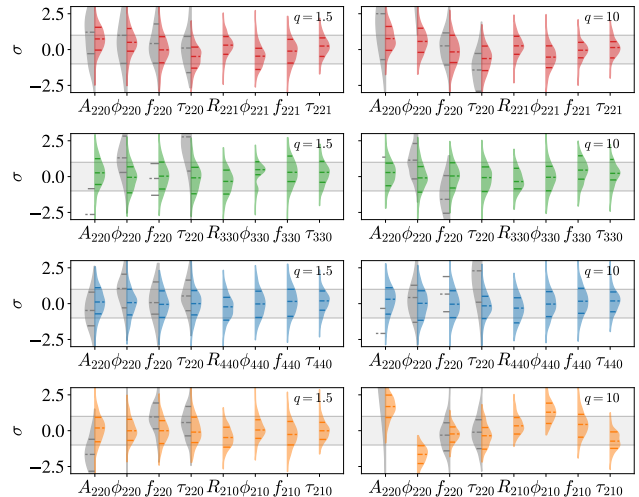


Figure 5. *Parameter estimation bias: Fitting data containing two modes with single-mode model vs. two-mode model.* The violin plots show the distribution of the deviation between the injected parameter $\vartheta_{\text{inj}}^\alpha$ and the posterior probability of ϑ^α from the data $d_2 = n + h_{220} + h_{\ell mn}$, see main text for more details. The results of each panel were obtained with 500 signals d_2 with different parameters and independent noise realizations (see main text for details). The left-sided violins (gray) show the deviation of each parameter of the dominant mode estimated with the single-mode model \mathcal{M}_1 and the colored right-sided violins show the deviation of all the parameters of the correct two-mode model \mathcal{M}_2 , considering the subdominant modes (2, 2, 1) (red), (3, 3, 0) (green), (4, 4, 0) (blue) and (2, 1, 0) (orange). Model \mathcal{M}_1 neglects the subdominant mode and produces results that are more biased and have a larger spread of deviations.

single-mode model for a confident claim of detection of a secondary mode.

This choice is supported by the results presented in Figure 6, where we compute the Bayes factor \mathcal{B}_1^2 using data that contain only the dominant mode $d_1 = n + h_{220}$ (to verify that $\ln \mathcal{B}_1^2$ is never greater than 8 in this case). As expected, most events have $\ln \mathcal{B}_1^2 \lesssim 0$ (or $\mathcal{B}_1^2 \lesssim 1$). We do not find any cases with $\ln \mathcal{B}_1^2 > 4$ (or $\mathcal{B}_1^2 \gtrsim 55$), but we caution that 5 out of the 2000 simulated signals ($\sim 0.3\%$ of all cases) had $10 \lesssim \mathcal{B}_1^2 \lesssim 55$, which is considered to be “strong” evidence according to the widely used scale by Kass and Raftery [68]. This indicates that such generic scales may not be suitable for different types of data (and noise), and should preferentially be avoided in the analysis of GW data if a more detailed analysis is possible.

In this and the following Sections, the evidences were calculated using PYMULTINEST [69–71] with 500 live points in each computation. A spot check with 1000 live points showed that the evidence changed by approximately 6%.

⁵ The exception are the grey violin plots for ϕ_{220} associated with (2,2,0)+(2,2,1) data for $q = 10$ and with (2,2,0)+(2,1,0) data for both $q = 1.5$ and $q = 10$, which are outside the axis limits and represent the largest bias ($> 6\sigma$).

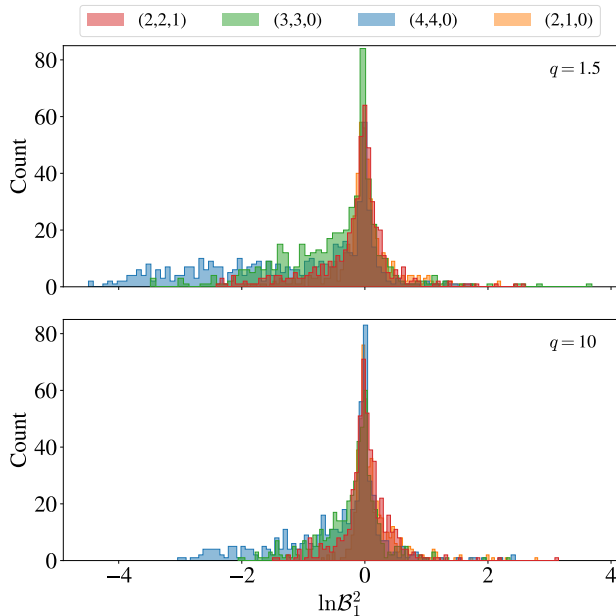


Figure 6. *Bayes factor threshold for confident detections: Looking for two modes in a single-mode signal.* The histograms show the natural logarithm of the Bayes factor \mathcal{B}_1^2 in favor of the two-mode model \mathcal{M}_2 over the single-mode model \mathcal{M}_1 , for 2000 signals containing single-mode data, $d_1 = n + h_{220}$ with LIGO noise realizations. The colors indicate the prior of the subdominant mode in each case: (2, 2, 1) (red), (3, 3, 0) (green), (4, 4, 0) (blue) and (2, 1, 0) (orange), see main text for more details. For each prior we generated 500 signals with $M \in [1, 5 \times 10^3]M_\odot$ and $z \in [10^{-2}, 1]$. For most signals the single-mode model is favored ($\mathcal{B}_1^2 \lesssim 1$), as expected, but we find 5 outliers with $10 \lesssim \mathcal{B}_1^2 \lesssim 55$. We require $\ln \mathcal{B}_1^2 > 8$ ($\mathcal{B}_1^2 \gtrsim 3000$) for claiming a confident detection of a secondary mode in a ringdown signal.

B. Bayesian model comparison: looking for a secondary mode

Recent multi-mode analyses of the ringdown of some GW detections found that the inclusion of subdominant modes decreases the errors in the inferred parameters of the final BH [10, 13, 36, 45]. However, the Bayes factors in favor of models with additional modes are close to 1.

In [42] the authors proposed the use of the ringdown part of the waveform approximants to probe the Kerr nature and found a high statistical evidence of $\ln \mathcal{B} \sim 6.5$ in favor of their model with respect to their best QNM model for GW150914. However, their model does not provide evidence for the individual detection of multiple QNMs. More recently, a new analysis of the ringdown of the most massive event GW190521 [29] had $\mathcal{B}_1^2 \sim 40$ ($\ln \mathcal{B}_1^2 \sim 3.7$). This is the highest evidence for a secondary mode presented so far.

We emphasize that the Bayes factor depends on the choice of priors and results from different works should not be directly compared without this context. In our

analysis we use Figure 6 to determine whether a given value for the Bayes factor provides “strong” evidence for each set of priors. We choose here the detectability threshold $\ln \mathcal{B}_1^2 = 8$; lower thresholds would result in larger spectroscopy horizons, albeit with lower significance.

The calculation of Bayes factors is very computationally expensive, and the time to find the root $\ln \mathcal{B}_1^2 - 8 = 0$ with independent noise realizations can be very long if one uses standard root-finding algorithms. However, for a fixed mass, the Bayes factor is approximately monotonic as a function of redshift. In Figure 7 we show an example of $\ln \mathcal{B}_1^2$ as a function of the redshift for many simulated signals with data $d_2 = h_{220} + h_{221} + n$ for events with fixed final mass and mass ratio at different distances. We find that $\ln \mathcal{B}_1^2$ is well approximated by a Laurent polynomial fit, which allows us to find the redshift at which $\ln \mathcal{B}_1^2 - 8 = 0$ for the fixed mass. This will be the redshift $z_{\ell mn}^{\text{spec,B}}$ of the BH spectroscopy horizon for that mass.

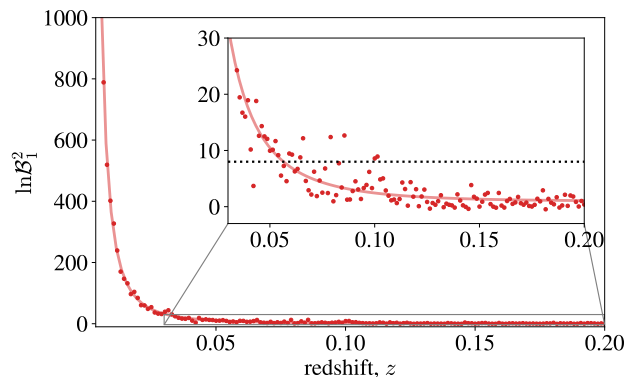


Figure 7. Bayes factor of the two-mode model \mathcal{M}_2 , here with the overtone (2, 2, 1) prior for the subdominant mode, over the single-mode model \mathcal{M}_1 as a function of redshift. There are 150 simulated signals with data $d_2 = h_{220} + h_{221} + n$, with fixed $M = 156M_\odot$, mass ratio $q = 1.5$ and independent LIGO noise realizations. The errors in the Bayes factors, as estimated by MultiNest, are smaller than the points. The scattering of the points is due to the independent noise realizations, but it is averaged out in the computation of the BH spectroscopy horizon, see Figure 8. The continuous curve is a Laurent polynomial $c_1 + c_2 z^{-1} + c_3 z^{-2}$ fitted to the points. The redshift at the detectability threshold $\ln \mathcal{B}_1^2 = 8$ is determined using the fitted curve.

To find the BH spectroscopy horizon using the Bayes factor threshold for a given GW detector, we choose several masses spaced log-uniformly in the range of interest. For each mass, initial samplings of $\ln \mathcal{B}_1^2$ at different distances provide a rough estimate of $z_{\ell mn}^{\text{spec,B}}$ for which $\ln \mathcal{B}_1^2 = 8$. This result is then refined using a Laurent polynomial fit with 10 (30) points chosen within $\pm 50\%$ of the rough estimate for LIGO (CE). Different thresholds could also be used with our method.

In Figure 8 we show the LIGO and CE BH spec-

troscopy horizons up to which $\ln \mathcal{B}_1^2 > 8$ for signals with data d_2 containing two modes. We fit the horizons with 3rd to 7th degree polynomials; all coefficients are given in Tables II (LIGO) and III (CE). In the $q = 1.5$ case, for lower masses, the first overtone (2, 2, 1) has the largest horizon, and the second largest is the (3, 3, 0) mode followed by the (2, 1, 0). For higher masses, the horizon of the (4, 4, 0) is slightly larger. For $q = 10$ the (3, 3, 0) and (4, 4, 0) modes have larger horizons than the (2, 2, 1) mode and the (2, 1, 0) mode has consistently the smallest horizon. The qualitative features are similar to the BH spectroscopy horizons obtained with the Rayleigh criterion (see Figure 3).

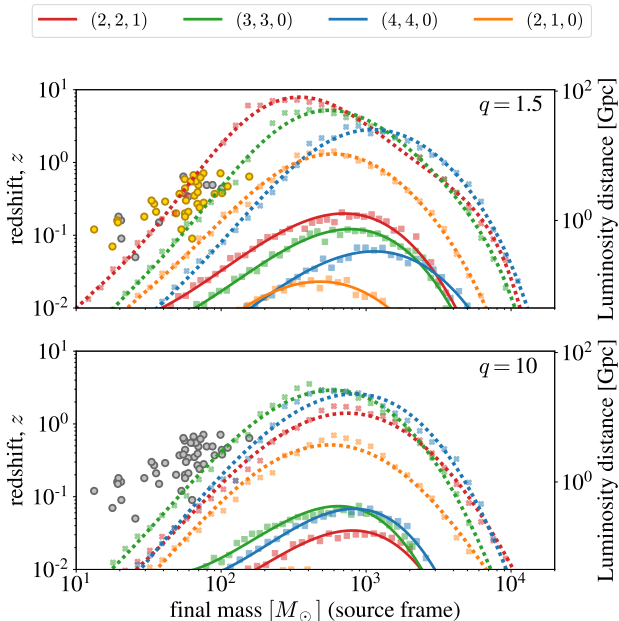


Figure 8. CE (crosses and dotted curves) and LIGO (squares and solid curves) BH spectroscopy horizons obtained with signals containing two modes, $d_2 = n + h_{220} + h_{\ell mn}$, by requiring Bayes factors $\ln \mathcal{B}_1^2 > 8$ for the subdominant modes (2, 2, 1) (red), (3, 3, 0) (green), (4, 4, 0) (blue) and (2, 1, 0) (orange). The curves are polynomials fitted to the points, with coefficients given in Tables II and III. Circles show detections from GWTC-1 and GWTC-2; yellow circles are compatible with the mass ratio in each case, whereas gray circles are not.

The BH spectroscopy horizon distance depends on the amplitude, the frequency and the damping time of the modes. As we can see in Table I for $q = 1.5$, the overtone has a much higher amplitude (at least a factor 5) than the higher harmonics and therefore it has the largest horizon. For $q = 10$, the amplitudes are much closer, and the overtone horizon is smaller than the (3, 3, 0) and (4, 4, 0) horizons, even though the amplitude of the overtone is still slightly larger than the amplitude of the higher harmonics. This shows that the difference between the frequency of the dominant and subdominant modes is more relevant than the difference between decay times, as the (3, 3, 0) and (4, 4, 0) modes have damping times very close

to the damping time of the (2, 2, 0) mode and very different frequencies, and there is an opposite relation between the (2, 2, 1) and (2, 2, 0) modes. Furthermore, for $q = 10$ the amplitude of the (4, 4, 0) mode is less than half of the amplitude of the (3, 3, 0) mode, but their horizons peak at approximately the same distance. This happens because the higher frequency of the (4, 4, 0) mode remains in the detector band for more massive BHs.

Moreover, the (2, 1, 0) mode always has a higher amplitude than the (4, 4, 0) mode and for $q = 10$ it has an amplitude comparable with the (2, 2, 1) and (3, 3, 0) modes, but its horizon is always much smaller. This is because the frequency *and* damping time of the (2, 1, 0) mode are very close to the dominant mode’s values. Therefore, the difference in damping times is also very important, as the overtone horizon is much larger than the (2, 1, 0) horizon.

These results confirm the importance of *both* $\Delta f_{220, \ell mn}$ and $\Delta \tau_{220, \ell mn}$ used in the Rayleigh criterion (6), however the horizons obtained using the Bayes factor threshold are much larger from the ones obtained using the Fisher matrix analysis and the Rayleigh criterion. For each secondary mode, the maximum horizon distance calculated using the Bayes factor threshold is at least a few times larger than the maximum horizon distance obtained with the Rayleigh criterion. The comparison between the methods is discussed in Sec. IV.

On the horizon $z_{\ell mn}^{\text{spec}, B}$, the SNR of the subdominant modes is larger than approximately 5 in all cases. The posterior probability distributions for all parameters are informed by the data and exclude zero amplitude of the secondary mode at the 90% credibility level (we examine one example in detail in Figure 13 below).

We can use our results together with the current rates for binary BH mergers to estimate the rate of events expected within the BH spectroscopy horizons. Recently, [72] quoted the rate of binary mergers with primary mass $45M_\odot < M_1 < 100M_\odot$ as $0.70^{+0.65}_{-0.35} \text{ Gpc}^{-3} \text{ yr}^{-1}$. Using the comoving volume at $z_{221}^{\text{spec}, B}$ for $q = 1.5$ and final mass $M = 156.3M_\odot$ (similar to GW190521), we find event rates $0.03 - 0.10 \text{ yr}^{-1}$ for LIGO and $(0.6 - 2.4) \times 10^3 \text{ yr}^{-1}$ for CE.

The horizons are smaller for lower masses and for other subdominant modes; optimal inclinations and sky locations could increase the horizons by less than a factor 2. We will have to be lucky to see such an event with LIGO, but there should be no lack of detections when the 3G detectors are operational.

C. Multimode LIGO BH spectroscopy horizons

The ringdown of a Kerr black hole is actually a superposition of infinite modes. The two-mode approximation is valid when the tertiary mode (and all of the higher

Table II. Coefficients used in the LIGO BH spectroscopy horizon fits shown in Figure 8, obtained with the Bayes factor condition $\ln \mathcal{B}_1^2 = 8$. The fits have the form $\ln z_{\ell mn}^{\text{spec,B}} = a_0 + a_1 \ln M + a_2 (\ln M)^2 + a_3 (\ln M)^3$, where $z_{\ell mn}^{\text{spec,B}}$ is the redshift at the horizon distance and M the black hole final mass. These fits are only valid for the mass intervals shown in Figure 8.

coefficient	(2, 2, 1)	(3, 3, 0)	(4, 4, 0)	(2, 1, 0)
$q = 1.5$				
a_0	-0.6685	-0.7298	-0.5887	-0.0022
a_1	4.0504	4.6084	3.8802	-1.3164
a_2	-6.8045	-8.3415	-7.1914	7.0597
a_3	1.2680	2.3177	1.3209	-11.0706
$q = 10$				
a_0	-0.8621	-0.9526	-1.1328	
a_1	5.7614	6.1710	7.8589	
a_2	-11.6508	-12.1380	-16.9297	
a_3	4.8845	5.3732	9.4642	

Table III. Same as Table II, but for the CE detector.

coefficient	(2, 2, 1)	(3, 3, 0)	(4, 4, 0)	(2, 1, 0)
$q = 1.5$				
a_0	-0.2331	-0.2081	-0.1342	-0.1491
a_1	3.3135	3.1694	2.1851	2.3180
a_2	-18.5618	-19.3023	-14.4192	-14.5241
a_3	51.7728	59.5351	48.7304	46.3409
a_4	-75.6249	-98.0044	-88.7223	-79.3446
a_5	57.1219	84.0752	84.6849	71.1739
a_6	-19.7984	-32.0157	-36.0500	-28.7501
$q = 10$				
a_0	0.1135	0.0647	-0.0107	0.1703
a_1	-1.5090	-1.0693	-0.3054	-1.9806
a_2	5.8306	4.1535	1.4965	7.0495
a_3	-6.9124	-3.9086	0.1276	-8.0004
a_4	-0.07017	-1.6608	-4.3371	-0.1510

modes) are too weak to be detected in the signal.⁶

In this section we consider the constraints imposed by a more realistic signal, where we assume that the data contains the noise, the dominant mode and the four most relevant subdominant modes [22], with their respective amplitudes informed by numerical relativity simulations. That is, we have $d = d_5$ given by

$$d_5 = n + h_{220} + h_{221} + h_{330} + h_{440} + h_{210}. \quad (16)$$

⁶ Additionally, the $(\ell, m) \neq (2, 2)$ harmonics cannot be detected in events viewed face-on ($\iota = 0$) and harmonics with odd m are not excited in non-spinning circular binaries with mass ratio $q = 1$.

The models considered in the multimode analysis are

- \mathcal{M}_1 , the same *single*-mode model defined in Section III A,
- \mathcal{M}_2 , a *two*-mode model with one unspecified mode (ℓ_1, m_1, n_1) and 8 parameters $\vartheta_2 = \{A, \phi_{220}, f_{220}, \tau_{220}, R_{\ell_1 m_1 n_1}, \phi_{\ell_1 m_1 n_1}, f_{\ell_1 m_1 n_1}, \tau_{\ell_1 m_1 n_1}\}$,
- \mathcal{M}_3 , a *three*-mode model with two unspecified modes (ℓ_1, m_1, n_1) and (ℓ_2, m_2, n_2) and 12 parameters $\vartheta_3 = \{A, \phi_{220}, f_{220}, \tau_{220}, R_{\ell_1 m_1 n_1}, \phi_{\ell_1 m_1 n_1}, f_{\ell_1 m_1 n_1}, \tau_{\ell_1 m_1 n_1}, R_{\ell_2 m_2 n_2}, \phi_{\ell_2 m_2 n_2}, f_{\ell_2 m_2 n_2}, \tau_{\ell_2 m_2 n_2}\}$.

The two-mode model \mathcal{M}_2 considered in this Section contains the same number of parameters as the model \mathcal{M}_2 defined in Section III A, but here it has a much broader prior. We are now looking for an unspecified subdominant mode and the priors need to allow the secondary mode to be any of the most relevant subdominant modes (2, 2, 1), (3, 3, 0), (4, 4, 0) and (2, 1, 0) (following the prescription described in Section III B). The same restrictions are applied to the priors of the secondary and tertiary modes in the three-mode model \mathcal{M}_3 .

In Figure 9 we show the multimode BH spectroscopy horizons for LIGO obtained with signals d_5 containing five modes, by requiring Bayes factors $\mathcal{B}_{n-1}^n > 8$, for $n = 3$ and $n = 2$ modes, where the Bayes factors \mathcal{B}_{n-1}^n are relative to the models \mathcal{M}_n and \mathcal{M}_{n-1} . The dashed curves are a 3rd degree polynomial fitted to the horizon points, with coefficients given in Table IV. As expected, the two-mode horizons are larger than the three-mode horizons. The $q = 1.5$ horizons are also larger than the $q = 10$ horizons, as expected. This happens because, for the same final mass, nonspinning binaries with more asymmetric initial masses emit less energy in the form of gravitational waves [20, 22, 33], and the mode amplitudes are smaller than the amplitudes of more symmetric systems (see Table I).

Figure 10 compares the LIGO spectroscopy horizons computed in Section III B with data containing only two modes (shown in Figure 8), and the LIGO multimode horizons presented in Figure 9. For $q = 1.5$, the two-mode horizon is equivalent to the (2, 2, 1) horizon. However, the three-mode horizon is slightly smaller than the (3, 3, 0) horizon, which is the second largest. This may be due to the penalization of a high number of parameters in the Bayes factor. For $q = 10$ the two-mode horizon is compatible with the (3, 3, 0) and the (4, 4, 0) horizons, all of which peak at approximately the same distance. Again, the three-mode horizon is smaller than the second largest horizon and it even seems to be compatible with the (2, 2, 1) horizon.

To determine which are the secondary and tertiary modes detected on the two- and three-mode horizons we compute the deviation of each estimated parameter relative to all of the injected modes (following the same procedure detailed in Section III A). In Figure 11 we show the deviations calculated on the two-mode horizon, that

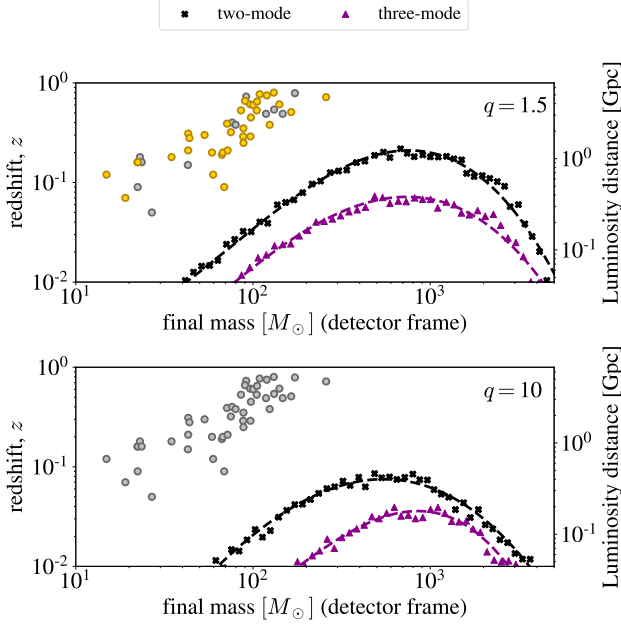


Figure 9. Multimode BH spectroscopy horizons for LIGO, using signals containing five modes, $d_5 = n + h_{220} + h_{221} + h_{330} + h_{440} + h_{210}$, by requiring Bayes factors $\mathcal{B}_{n-1}^n > 8$ for $n = 2$ (black) and $n = 3$ (purple). Here we are looking for the most relevant, but unspecified, secondary and tertiary modes in the five-mode signal. The dashed curves are 3rd degree polynomials fitted to the horizons.

Table IV. LIGO Bayes factor $\ln \mathcal{B}_1^2 = 8$ spectroscopy horizons fitting coefficients, $\ln z_n^{\text{spec,B}} = a_0 + a_1 \ln M + a_2 (\ln M)^2 + a_3 (\ln M)^3$, where $z_n^{\text{spec,B}}$ is the redshift at the horizon and M the black hole final mass. These fits are only valid for the mass intervals of the points in the Figure 9.

coefficient	2-mode model	3-mode model
$q = 1.5$		
a_0	-0.4238	-0.3022
a_1	2.2212	1.3808
a_2	-2.3470	-0.4939
a_3	-2.2267	-3.9584
$q = 10$		
a_0	-0.1121	-0.7089
a_1	-0.2280	4.5011
a_2	3.7164	-8.1516
a_3	-7.2915	1.6289

is, for each mass the redshift is fixed to the redshift of the horizon, $z_{n=2}^{\text{spec,B}}$. We consider the mass ranges where $z_{n=2}^{\text{spec,B}} > 10^{-2}$, as in Figure 9. For each horizon we select 10 final masses spaced log-uniformly, and for each mass we generate 100 signals with independent noise realizations. The colored bands show the $\pm 1\sigma$ highest probability density region of the deviations.

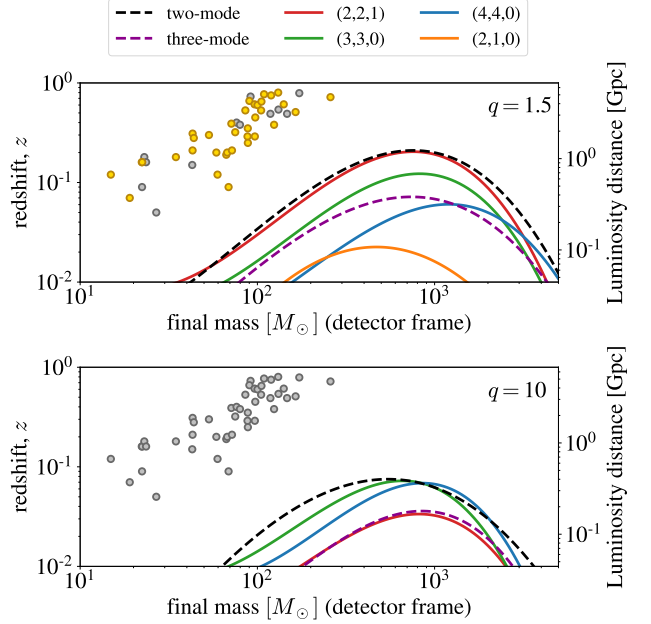


Figure 10. Comparison between the LIGO spectroscopy horizons presented in Figure 8, shown with solid lines, and the LIGO multimode spectroscopy horizons presented in Figure 9, shown with dashed lines. The (unspecified) two-mode horizon follows the largest (specified) horizon for each mass; the three-mode horizon is smaller than the second largest horizon.

As the secondary mode is unspecified, and the priors are broad enough to accommodate any of the four most relevant subdominant modes, the secondary mode identification can change over the two-mode horizon. For $q = 1.5$, the secondary mode is compatible with the $(2, 2, 1)$ mode in almost the entire mass range, as expected from Figure 10. For the highest masses $M \gtrsim 10^3 M_\odot$ (lowest frequencies), the secondary mode becomes compatible with the $(3, 3, 0)$ mode, which has a higher frequency and therefore remains longer in the detector band. This trend is especially clear from the deviations in the amplitude ratio $R_{\ell mn}$ and damping time $\tau_{\ell mn}$. For $q = 10$, Figure 10 is not enough to identify the secondary mode, but Figure 11 allows us to identify it as the $(3, 3, 0)$ mode for most of the mass range and the $(4, 4, 0)$ mode for the highest masses, similarly to the transition in the low mass ratio case.

In Figure 12 we show a similar analysis for the three-mode horizon, but only for the modes we identified as the secondary and tertiary modes in the three-mode model \mathcal{M}_3 . For $q = 1.5$, the modes are identified as the $(2, 2, 1)$ and the $(3, 3, 0)$, whereas for $q = 10$ we have the $(3, 3, 0)$ and the $(4, 4, 0)$. The larger offsets that are more clearly visible in the deviations of some parameters of the tertiary mode are a result of the “contamination” from the next subdominant modes in the signal.

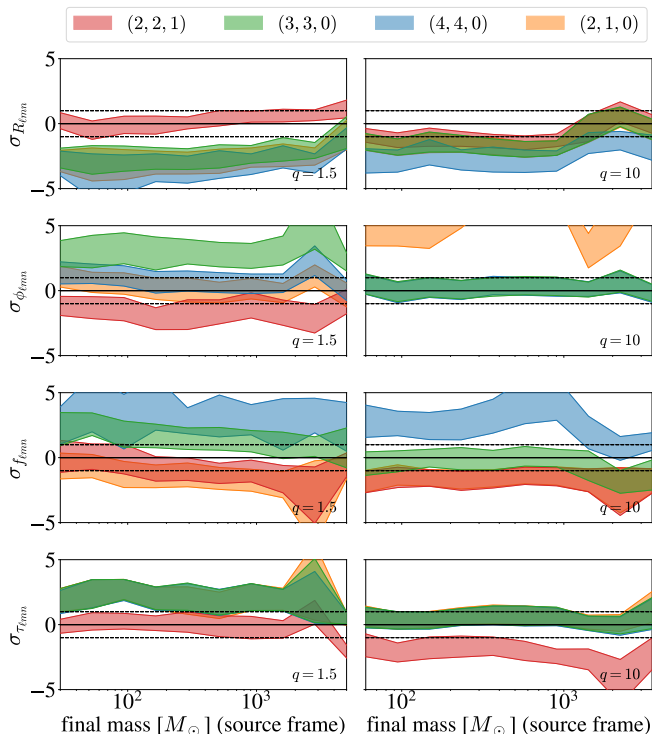


Figure 11. *Identifying the secondary mode:* deviations between the posteriors of parameters ϑ^α of the unspecified secondary mode in model \mathcal{M}_2 and the injected parameters $\vartheta_{\text{inj}}^\alpha$ of each of the four subdominant modes in the data $d_5 = n + h_{220} + h_{221} + h_{330} + h_{440} + h_{210}$. The colored bands show the $\pm 1\sigma$ highest probability density region of the distributions of the deviations in each case. For $q = 1.5$ (left), the (2, 2, 1) is the secondary mode detected; for $q = 10$, it is the (3, 3, 0) mode.

IV. DISCUSSION: RAYLEIGH VERSUS BAYES

Comparing Figures 4 and 8, we can see that the BH spectroscopy horizons obtained with the Rayleigh criterion and with the Bayes factor threshold have similar trends. However, $z_{\ell mn}^{\text{spec,R}}$ is consistently smaller than $z_{\ell mn}^{\text{spec,B}}$. This can be due to two reasons: the Fisher matrix approximation is not valid at these large distances and/or the Rayleigh criterion is inherently too restrictive.

The Fisher matrix analysis is only valid for high SNR signals [73], therefore we must be careful when using it for obtaining the BH spectroscopy horizons. Figure 13 shows the posterior distributions for the eight parameters of a ringdown signal with modes (2, 2, 0) and (2, 2, 1). We place the source at three distances: inside, outside and on the horizon $z_{221}^{\text{spec,B}}$, in order to compare the example posteriors (grey curves) and the Fisher error estimation (black Gaussian) in each case. The agreement between the grey curves (within the variation provided by the noise) shows that the parameter estimation has converged. Additionally, the agreement of the grey curves with the black curve (which peaks at the injected value)

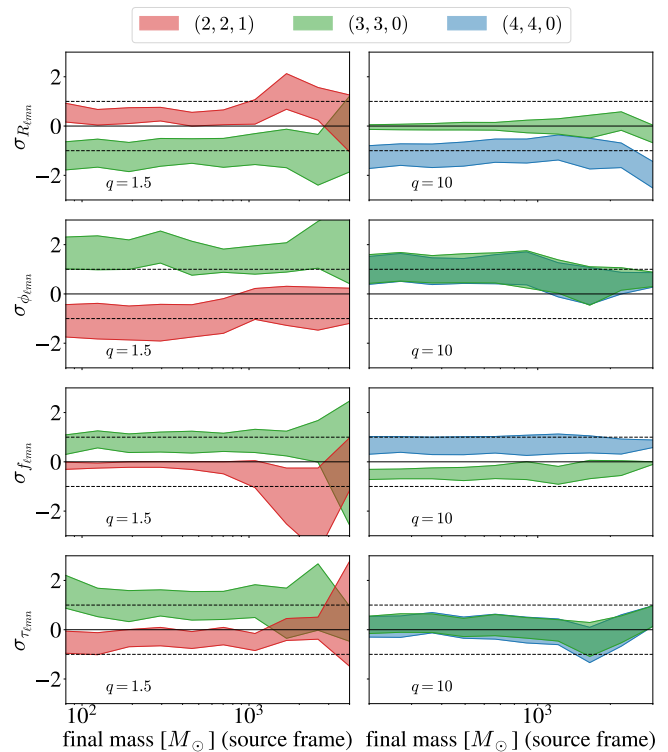


Figure 12. *Identifying the secondary and tertiary modes:* same as Figure 11, but for the secondary and tertiary modes identified for the three-mode model \mathcal{M}_3 : (2, 2, 1) and (3, 3, 0) for $q = 1.5$ and (3, 3, 0) and (4, 4, 0) for $q = 10$.

shows that the injected parameters are recovered.

For the case *inside* the horizon, the Fisher matrix error estimations agree well with the posteriors, but the errors are already noticeably larger for some parameters, such as A , R and τ_{221} . For the case *on* the horizon, the Fisher matrix error for τ_{220} also becomes larger than the Bayes inference estimation. For the case *outside* the horizon all of the errors given by the Fisher Matrix are larger than the Bayes estimations, and for most parameters the black Gaussians appear nearly flat in the ranges shown. However, the posteriors for the (2, 2, 1) parameters outside the horizon simply recover their priors, as the signal is not strong enough to be informative.

As expected, for large distances (low SNR) the Fisher matrix estimations are incorrect. For the case inside the Bayes factor horizon with $z = 0.03$ the Fisher estimations are good, but this redshift is outside the Rayleigh criterion horizon $z_{221}^{\text{spec,R}}$ for $M = 156.3$ (see Figure 4). This indicates that the poor error predictions of the Fisher matrix are not solely responsible for restricting the horizons, and the Rayleigh criterion itself seems to be too restrictive to determine the evidence of a secondary mode.

The very high SNR needed to satisfy both Rayleigh conditions was addressed by [49], where they proposed a generalized likelihood ratio test (GLRT) to identify the presence of a secondary mode in the signal. The GLRT is similar to the Bayes factor criterion, and in both criteria

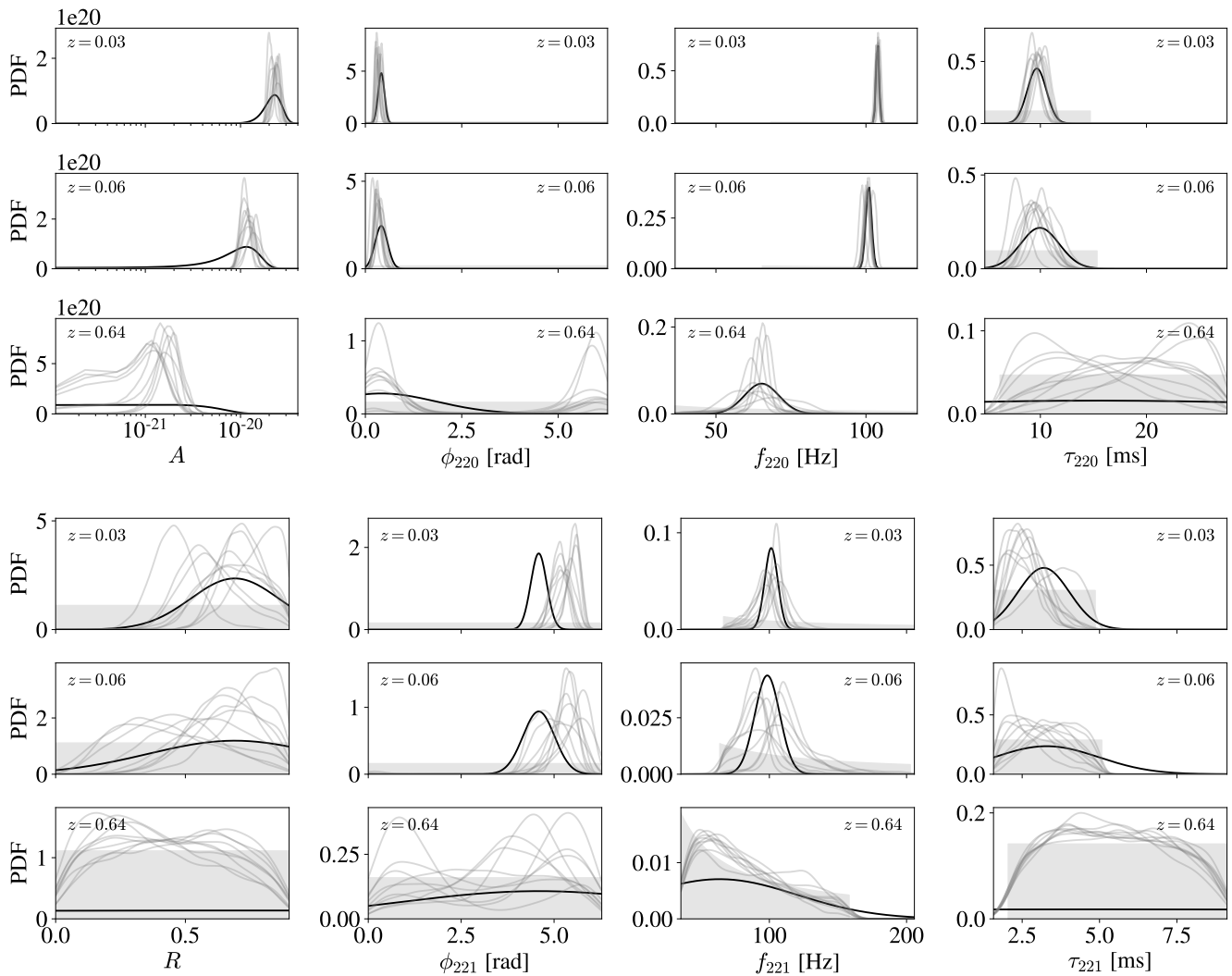


Figure 13. Posterior distributions for the eight parameters of a ringdown signal with the modes $(2, 2, 0)$ (top half of the figure) and $(2, 2, 1)$ (bottom half of the figure). We fixed the final BH mass $M = 156.3$ (similar to GW190521 [10]) and $q = 1.5$, and chose three distances: a distance inside the horizon $z_{221}^{\text{spec,B}}$ (1st and 4th rows), the distance of the horizon $z_{221}^{\text{spec,B}}$ (2nd and 5th rows) and a distance compatible with the event GW190521, well outside the horizon $z_{221}^{\text{spec,B}}$ (3rd and 6th rows). The gray filled regions are the prior probability distributions used in the Bayes inference, the gray curves show 10 examples of the posterior probability distributions (obtained with independent noise realizations) for each case and the black curves represent a Gaussian distribution constructed with errors estimated by the Fisher matrix. The injected parameters correspond to the peak of the black curves.

the confidence level can be increased in order to make them equivalent to the Rayleigh criterion, which results in a more restrictive detectability condition.

We computed the Bayes factor \mathcal{B}_1^2 for the $(2, 2, 1)$ mode at the corresponding Rayleigh horizon for the LIGO detector and $q = 1.5$, considering 500 noise realizations for 50 log-distributed masses. We found very high values, $\ln \mathcal{B} \gtrsim 60$ (or $\mathcal{B} \gtrsim 10^{26}$). Moreover, the Bayes factor value is not constant at the Rayleigh horizon, it has a minimum near $M = 45$ and monotonically increases for smaller and larger masses. For higher harmonics, the Rayleigh criterion is more restrictive due to the damping

time condition, resulting in even larger Bayes factors.

An alternative could be to require the resolvability of the quality factor $Q_{\ell mn} = \pi f_{\ell mn} \tau_{\ell mn}$ instead of the damping time $\tau_{\ell mn}$, or to require the resolvability of just one of the Rayleigh conditions, as done by [37, 39]. In appendix A we explore these and other variations of the Rayleigh criterion and compare them with the result of our Bayesian analysis for LIGO. A modification in the Rayleigh criterion that is compatible with the Bayes inference analysis is desirable for the extension of this kind of analysis for a larger set of systems, as the Bayesian analysis requires extensive computational resources.

V. CONCLUSIONS

The possibility of testing the no-hair theorem and GR with black hole spectroscopy is at hand. There is already some evidence for the detection of a secondary mode [13, 27, 36], and the detection of another event like GW150914 with current detector sensitivity could provide significantly improved constraints. We propose here a conservative criterion for detectability: a Bayes factor threshold $\ln \mathcal{B}_1^2 > 8$, which simultaneously guarantees that the secondary mode will be detectable in the signal and distinguishable from the fundamental mode $(2, 2, 0)$.

With this criterion we calculate the BH spectroscopy horizons, that is the maximum distance (averaged over sky location and binary inclination) up to which two or more modes can be detected in the ringdown of a binary black hole merger. Our analysis is restricted to non-spinning circular binaries with mass ratios equal to 1.5 and 10; the first mass ratio is compatible with most LVC detections [7, 8] and the second has an asymmetry that is important for the excitation of the higher harmonics in the ringdown. Importantly, we start analyzing all ringdown signals at $t = 10(M_1 + M_2)$ after the peak amplitude.

The best prospects for detection of a secondary mode arrive for nearly equal mass binaries and higher masses. In such cases, the overtone mode $(2, 2, 1)$ will be the easiest subdominant mode to detect, followed by the $(3, 3, 0)$ mode, as we determine in an agnostic multimode analysis. For very asymmetric binaries with a 10:1 mass ratio, the secondary and tertiary modes are the harmonics $(3, 3, 0)$ and $(4, 4, 0)$. At the high mass end of our analysis ($M \gtrsim 10^3 M_\odot$), the secondary and tertiary modes are switched, as the modes with higher frequency stand out more against the low-frequency noise of the detector.

For LIGO, the expected rate for an event similar to GW190521 within the corresponding BH spectroscopy horizon redshift $z_{221}^{\text{spec,B}} \sim 0.06$ is $0.03 - 0.10 \text{ yr}^{-1}$. For CE, this rate is dramatically improved: $(0.6 - 2.4) \times 10^3 \text{ yr}^{-1}$. Intermediate mass black holes with masses between approximately 10^2 and 10^3 solar masses have the largest horizons for ground-based detectors, but their rates are still unknown. Binary BH merger rates for LISA sources are also still very uncertain. However, LISA

should have no problem resolving more than one extra mode in the ringdown.

Other proposals for appraising the resolvability of a secondary mode have been suggested in the literature. We present a detailed analysis of the Rayleigh criterion, and contrast its predictions with our Bayesian analysis. We find the Rayleigh criterion to be excessively restrictive, but useful for quicker estimates and establishing general trends. A correction of the Rayleigh criterion that could reproduce the Bayesian results would have many practical advantages. Different combinations and variations of the Rayleigh conditions provide constraints at varying levels of required significance (evidence) for the secondary mode.

Finally, a generalization of this work could include a network of detectors in the analysis, which would increase the BH spectroscopy horizons by a factor $\lesssim 2$. Coherent mode stacking [74] has been proposed to use multiple measurements to increase the SNR of the subdominant quasinormal modes. The original proposal is not easily applied to overtones, but data from individual signals can be combined in a Bayesian framework [75] to determine how many ringdown events, with the current merger rate estimates, would be needed to detect the higher harmonics with an statistical evidence of $\ln \mathcal{B} > 8$.

ACKNOWLEDGMENTS

We are thankful to Quentin Baghi for providing the LISA noise spectral density curve, Tito Dal Canton for answering our questions about LIGO noise and Lucas Sanches for help with Slurm/HPC. We also thank Cole Miller and the Flatiron CCA GW Group for useful discussions, and Emanuele Berti, Vitor Cardoso, Will Farr, Max Isi and Xisco Jimenez Forteza for comments on our manuscript. IO was supported by grant 2018/21286-3 of the São Paulo Research Foundation (FAPESP). CC acknowledges support by NASA under award number 80GSFC17M0002. This work was completed at the Aspen Center for Physics, which is supported by National Science Foundation grant PHY-1607611. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at Goddard Space Flight Center.

-
- [1] N. Yunes and X. Siemens, Living Reviews in Relativity **16**, 9 (2013), arXiv:1304.3473 [gr-qc].
 - [2] R. Abbott *et al.*, ApJ **896**, L44 (2020), arXiv:2006.12611 [astro-ph.HE].
 - [3] K. Mehgagan, J. Thomas, R. Saglia, X. Mazzalay, P. Erwin, R. Bender, M. Kluge, and M. Fabricius, ApJ **887**, 195 (2019), arXiv:1907.10608 [astro-ph.GA].
 - [4] K. D. Kokkotas and B. G. Schmidt, Living Rev. Rel. **2**, 2 (1999), arXiv:gr-qc/9909058.
 - [5] H.-P. Nollert, Classical and Quantum Gravity **16**, R159 (1999).
 - [6] E. Berti, V. Cardoso, and A. O. Starinets, Class. Quant. Grav. **26**, 163001 (2009), arXiv:0905.2975 [gr-qc].
 - [7] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. X **9**, 031040 (2019), arXiv:1811.12907 [astro-ph.HE].
 - [8] R. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. X **11**, 021053 (2021), arXiv:2010.14527 [gr-qc].
 - [9] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev.

- Lett. **116**, 061102 (2016), arXiv:1602.03837 [gr-qc].
- [10] R. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. Lett. **125**, 101102 (2020), arXiv:2009.01075 [gr-qc].
- [11] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. Lett. **116**, 221101 (2016), [Erratum: Phys.Rev.Lett. 121, 129902 (2018)], arXiv:1602.03841 [gr-qc].
- [12] R. Abbott *et al.* (LIGO Scientific, Virgo), Astrophys. J. Lett. **900**, L13 (2020), arXiv:2009.01190 [astro-ph.HE].
- [13] R. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. D **103**, 122002 (2021), arXiv:2010.14529 [gr-qc].
- [14] O. Dreyer, B. J. Kelly, B. Krishnan, L. S. Finn, D. Garrison, and R. Lopez-Aleman, Class. Quant. Grav. **21**, 787 (2004), arXiv:gr-qc/0309007.
- [15] E. Berti, V. Cardoso, and C. M. Will, Phys. Rev. D **73**, 064030 (2006), arXiv:gr-qc/0512160.
- [16] V. Cardoso and L. Gualtieri, Class. Quant. Grav. **33**, 174001 (2016), arXiv:1607.03133 [gr-qc].
- [17] E. Berti, K. Yagi, H. Yang, and N. Yunes, Gen. Rel. Grav. **50**, 49 (2018), arXiv:1801.03587 [gr-qc].
- [18] R. Brito, A. Buonanno, and V. Raymond, Phys. Rev. D **98**, 084038 (2018), arXiv:1805.00293 [gr-qc].
- [19] E. Berti, V. Cardoso, J. A. Gonzalez, U. Sperhake, M. Hannam, S. Husa, and B. Bruegmann, Phys. Rev. D **76**, 064034 (2007), arXiv:gr-qc/0703053.
- [20] I. Kamaretsos, M. Hannam, S. Husa, and B. S. Sathyaprakash, Phys. Rev. D **85**, 024018 (2012), arXiv:1107.0854 [gr-qc].
- [21] E. Thrane, P. D. Lasky, and Y. Levin, Phys. Rev. D **96**, 102004 (2017), arXiv:1706.05152 [gr-qc].
- [22] R. Cotesta, A. Buonanno, A. Bohé, A. Taracchini, I. Hinder, and S. Ossokine, Phys. Rev. D **98**, 084028 (2018), arXiv:1803.10701 [gr-qc].
- [23] V. Baibhav and E. Berti, Phys. Rev. D **99**, 024005 (2019), arXiv:1809.03500 [gr-qc].
- [24] G. Carullo, W. Del Pozzo, and J. Veitch, Phys. Rev. D **99**, 123029 (2019), [Erratum: Phys.Rev.D 100, 089903 (2019)], arXiv:1902.07527 [gr-qc].
- [25] A. Maselli, P. Pani, L. Gualtieri, and E. Berti, Phys. Rev. D **101**, 024043 (2020), arXiv:1910.12893 [gr-qc].
- [26] N. Uchikata, T. Narikawa, K. Sakai, H. Takahashi, and H. Nakano, Phys. Rev. D **102**, 024007 (2020), arXiv:2003.06791 [gr-qc].
- [27] C. D. Capano and A. H. Nitz, Phys. Rev. D **102**, 124070 (2020), arXiv:2008.02248 [gr-qc].
- [28] F. H. Shaik, J. Lange, S. E. Field, R. O’Shaughnessy, V. Varma, L. E. Kidder, H. P. Pfeiffer, and D. Wysocki, Phys. Rev. D **101**, 124054 (2020), arXiv:1911.02693 [gr-qc].
- [29] C. D. Capano, M. Cabero, J. Westerweck, J. Abedi, S. Kastha, A. H. Nitz, A. B. Nielsen, and B. Krishnan, (2021), arXiv:2105.05238 [gr-qc].
- [30] E. W. Leaver, Phys. Rev. D **34**, 384 (1986).
- [31] R. F. Stark and T. Piran, Phys. Rev. Lett. **55**, 891 (1985).
- [32] A. Buonanno, G. B. Cook, and F. Pretorius, Phys. Rev. D **75**, 124018 (2007), arXiv:gr-qc/0610122.
- [33] L. London, D. Shoemaker, and J. Healy, Phys. Rev. D **90**, 124032 (2014), [Erratum: Phys.Rev.D 94, 069902 (2016)], arXiv:1404.3197 [gr-qc].
- [34] V. Baibhav, E. Berti, V. Cardoso, and G. Khanna, Phys. Rev. D **97**, 044048 (2018), arXiv:1710.02156 [gr-qc].
- [35] M. Giesler, M. Isi, M. A. Scheel, and S. A. Teukolsky, Physical Review X **9** (2019), 10.1103/physrevx.9.041060, arXiv:1903.08284 [gr-qc].
- [36] M. Isi, M. Giesler, W. M. Farr, M. A. Scheel, and S. A. Teukolsky, Phys. Rev. Lett. **123**, 111102 (2019), arXiv:1905.00869 [gr-qc].
- [37] S. Bhagwat, X. J. Forteza, P. Pani, and V. Ferrari, Phys. Rev. D **101**, 044033 (2020), arXiv:1910.08708 [gr-qc].
- [38] I. Ota and C. Chirenti, Phys. Rev. D **101**, 104005 (2020), arXiv:1911.00440 [gr-qc].
- [39] X. Jiménez Forteza, S. Bhagwat, P. Pani, and V. Ferrari, Phys. Rev. D **102**, 044053 (2020), arXiv:2005.03260 [gr-qc].
- [40] P. Mourier, X. Jiménez Forteza, D. Pook-Kolb, B. Krishnan, and E. Schnetter, Phys. Rev. D **103**, 044054 (2021), arXiv:2010.15186 [gr-qc].
- [41] M. Okounkova, (2020), arXiv:2004.00671 [gr-qc].
- [42] J. C. Bustillo, P. D. Lasky, and E. Thrane, Phys. Rev. D **103**, 024041 (2021), arXiv:2010.01857 [gr-qc].
- [43] E. Finch and C. J. Moore, Phys. Rev. D **103**, 084048 (2021), arXiv:2102.07794 [gr-qc].
- [44] M. Isi, W. M. Farr, M. Giesler, M. A. Scheel, and S. A. Teukolsky, Phys. Rev. Lett. **127**, 011103 (2021), arXiv:2012.04486 [gr-qc].
- [45] M. Isi and W. M. Farr, (2021), arXiv:2107.05609 [gr-qc].
- [46] X. J. Forteza and P. Mourier, (2021), arXiv:2107.11829 [gr-qc].
- [47] A. Dhani, Phys. Rev. D **103**, 104048 (2021), arXiv:2010.08602 [gr-qc].
- [48] A. Dhani and B. S. Sathyaprakash, (2021), arXiv:2107.14195 [gr-qc].
- [49] E. Berti, J. Cardoso, V. Cardoso, and M. Cavaglia, Phys. Rev. D **76**, 104044 (2007), arXiv:0707.1202 [gr-qc].
- [50] L. S. Finn, Phys. Rev. D **46**, 5236 (1992), arXiv:gr-qc/9209010.
- [51] <https://dcc.ligo.org/LIGO-T1800044/public>.
- [52] P. Amaro-Seoane *et al.*, “Laser interferometer space antenna,” (2017), arXiv:1702.00786 [astro-ph.IM].
- [53] M. Punturo *et al.*, Classical and Quantum Gravity **27**, 194002 (2010).
- [54] B. P. Abbott *et al.*, Classical and Quantum Gravity **34**, 044001 (2017), arXiv:1607.08697 [astro-ph.IM].
- [55] Planck Collaboration, N. Aghanim, *et al.*, A&A **641**, A6 (2020), arXiv:1807.06209 [astro-ph.CO].
- [56] C. J. Moore, R. H. Cole, and C. P. L. Berry, Class. Quant. Grav. **32**, 015014 (2015), arXiv:1408.0740 [gr-qc].
- [57] T. Regimbau *et al.*, Phys. Rev. D **86**, 122001 (2012), arXiv:1201.3563 [gr-qc].
- [58] T. Robson, N. J. Cornish, and C. Liu, Class. Quant. Grav. **36**, 105011 (2019), arXiv:1803.01944 [astro-ph.HE].
- [59] <https://dcc.cosmicexplorer.org/cgi-bin/DocDB/ShowDocument?docid=T2000017>.
- [60] <http://www.et-gw.eu/index.php/etsensitivities>.
- [61] <https://pages.jh.edu/~eberti2/ringdown/>.
- [62] S. Bhagwat, D. A. Brown, and S. W. Ballmer, Phys. Rev. D **94**, 084024 (2016), [Erratum: Phys.Rev.D 95, 069906 (2017)], arXiv:1607.07845 [gr-qc].
- [63] E. E. Flanagan and S. A. Hughes, Phys. Rev. D **57**, 4535 (1998), arXiv:gr-qc/9701039.
- [64] M. Boyle *et al.*, Class. Quant. Grav. **36**, 195006 (2019), arXiv:1904.04831 [gr-qc].
- [65] <https://data.black-holes.org/waveforms/index.html>.
- [66] S. Talts, M. Betancourt, D. Simpson, A. Vehtari, and A. Gelman, arXiv e-prints, arXiv:1804.06788 (2018), arXiv:1804.06788 [stat.ME].
- [67] I. M. Romero-Shaw *et al.*, MNRAS **499**, 3295 (2020),

arXiv:2006.00714 [astro-ph.IM].

- [68] R. E. Kass and A. E. R. Raftery, *Journal of the American Statistical Association* **90**, 773 (1995).
- [69] F. Feroz, M. P. Hobson, and M. Bridges, *Mon. Not. Roy. Astron. Soc.* **398**, 1601 (2009), arXiv:0809.3437 [astro-ph].
- [70] F. Feroz, M. P. Hobson, E. Cameron, and A. N. Pettitt, *Open J. Astrophys.* **2**, 10 (2019), arXiv:1306.2144 [astro-ph.IM].
- [71] J. Buchner, A. Georgakakis, K. Nandra, L. Hsu, C. Rangel, M. Brightman, A. Merloni, M. Salvato, J. Donley, and D. Kocevski, *Astron. Astrophys.* **564**, A125 (2014), arXiv:1402.0004 [astro-ph.HE].
- [72] R. Abbott *et al.*, *ApJ* **913**, L7 (2021), arXiv:2010.14533 [astro-ph.HE].
- [73] M. Vallisneri, *Phys. Rev. D* **77**, 042001 (2008), arXiv:gr-qc/0703086.
- [74] H. Yang, K. Yagi, J. Blackman, L. Lehner, V. Paschalidis, F. Pretorius, and N. Yunes, *Phys. Rev. Lett.* **118**, 161101 (2017), arXiv:1701.05808 [gr-qc].
- [75] A. Zimmerman, C.-J. Haster, and K. Chatziioannou, *Phys. Rev. D* **99**, 124044 (2019), arXiv:1903.11008 [astro-ph.IM].

Appendix A: Variations of the resolvability criteria

Here we explore different variations of the Rayleigh criterion proposed in the literature and present the resulting BH spectroscopy horizons for each case.

The formulation of the Rayleigh conditions (6) was first introduced in [15], where it was used to compute the “critical” SNR ρ_{crit} to resolve *either* frequency or damping time and the larger SNR ρ_{both} needed to resolve *both* conditions. It was found in [15] that ρ_{crit} is one or two orders of magnitude smaller than ρ_{both} .

We require that both conditions be satisfied in Section II. Some recent works have confirmed that requiring a single condition to be satisfied is not very restrictive [39, 62]. Accordingly, we present in Figure 14 the LIGO spectroscopy horizons by requiring the resolvability of the frequency *or* the damping time. As expected, the horizons are much larger than in the case when both conditions are required, shown in Figure 4. We can see that requiring just one condition favors the modes with $\ell \neq 2$, which have easily resolvable frequencies. The (2, 2, 1) horizon is further reduced by the very large errors in the damping times estimated by the Fisher matrix analysis at the horizon distances (see discussion in Section IV).

More recently, [45] reinforced the idea that requiring both conditions should not be necessary to distinguish between two ringdown modes and suggested a single condition for the distinguishability of the 2-dimensional posteriors of frequencies and damping times,

$$\frac{(f_{\ell mn} - f_{\ell' m' n'})^2}{\sigma_{f_{\ell mn}}^2 + \sigma_{f_{\ell' m' n'}}^2} + \frac{(\tau_{\ell mn} - \tau_{\ell' m' n'})^2}{\sigma_{\tau_{\ell mn}}^2 + \sigma_{\tau_{\ell' m' n'}}^2} \gtrsim 1. \quad (\text{A1})$$

In Figure 15 we show the LIGO BH spectroscopy horizons obtained with equation (A1). We can see that these

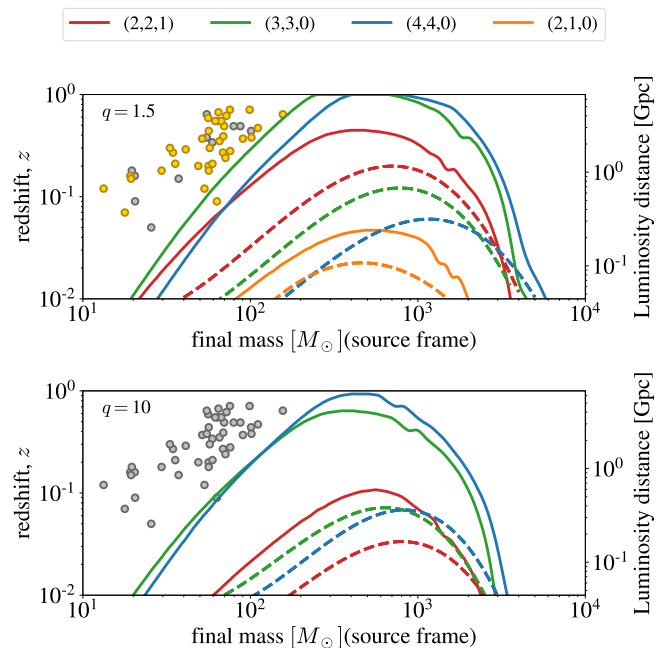


Figure 14. LIGO BH spectroscopy horizons obtained by requiring only one of Rayleigh conditions (6a) or (6b) to be satisfied (solid lines). The errors are estimated with a Fisher matrix analysis. The dashed curves show the LIGO BH spectroscopy horizons obtained with a Bayes factor threshold $\ln \mathcal{B} > 8$ (same as in Figure 8). Circles show detections from GWTC-1 and GWTC-2; yellow circles are compatible with the mass ratio in each case, whereas gray circles are not. The one-condition Rayleigh horizons are much less restrictive and favor subdominant modes with $\ell \neq 2$.

horizons are very close to the single-condition Rayleigh horizons in Figure 14 and the same considerations apply here. If the Fisher matrix error estimates are replaced with values derived from the posterior distributions, the (2, 2, 1) horizon reaches farther and the constraints become compatible with a somewhat lower Bayes factor threshold than the conservative value we used. Obviously, in each *individual* event detected, properties such as the mass ratio, sky location and binary inclination will also impact the detectability of a secondary mode.

In [39] the use of the quality factor $Q_{\ell mn} = \pi f_{\ell mn} \tau_{\ell mn}$ instead of the damping time was proposed, replacing equation (6b) by

$$|Q_{\ell mn} - Q_{\ell' m' n'}| > \max(\sigma_{Q_{\ell mn}}, \sigma_{Q_{\ell' m' n'}}). \quad (\text{A2})$$

The damping times are also replaced by the quality factors for the Fisher matrix parameters, which results in different error estimation for the frequencies.

In Figure 16 we show the LIGO spectroscopy horizons by requiring the resolvability of *both* frequencies and quality factors. The horizons for the harmonics (3, 3, 0) and (4, 4, 0) are very close to the Bayes factor horizons, but the horizons of the overtone and the (2, 1, 0) mode are still very restrictive. This is expected, as the quality factor “relaxes” the second Rayleigh condition on the

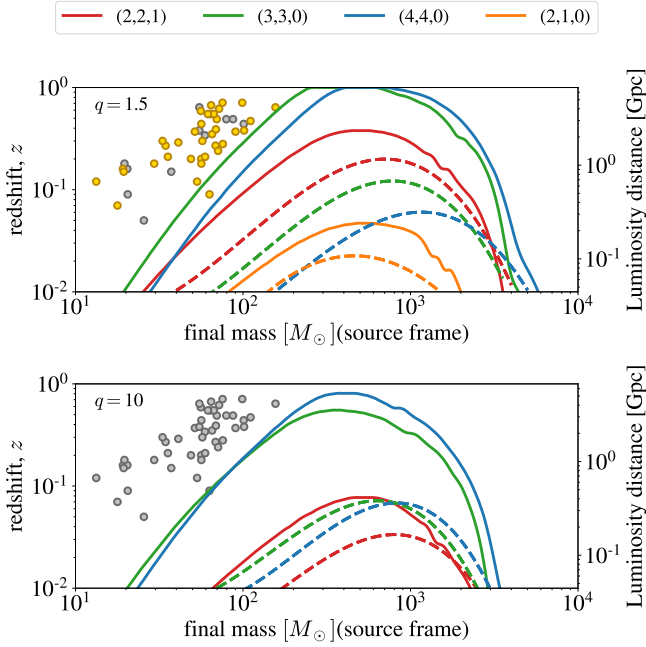


Figure 15. Same as Figure 14, but here the solid lines show the LIGO BH spectroscopy horizons obtained with condition (A1). These horizons are very similar to the single-condition Rayleigh horizons of Figure 14.

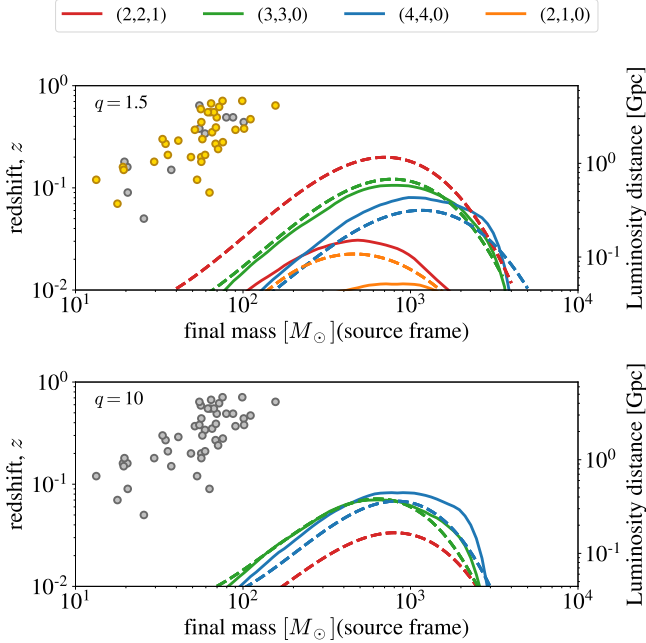


Figure 16. Same as Figure 14, but using the quality factors $Q_{\ell mn} = \pi f_{\ell mn} \tau_{\ell mn}$ instead of the damping times in the Rayleigh conditions (6) (solid lines). The horizons obtained for modes with $\ell \neq 2$ are very close to the horizons obtained with the Bayes factor threshold (dashed lines), but the horizons of the modes with $\ell = 2$ are more restrictive.

fundamental higher harmonic modes, which have damp-

ing times very similar to the dominant mode (2, 2, 0).

In Figures 14, 15 and 16 we impose different prescriptions for the resolvability criteria of two modes in the ringdown, in the spirit of the original Rayleigh criterion. However, we do not impose the detectability requirement that each mode (ℓ, m, n) should have a SNR $\rho_{\ell mn} > 8$, which we also impose in our Rayleigh horizons in Section II. (In Section II, however, this condition is superfluous: requiring both Rayleigh conditions to be satisfied is more restrictive than the SNR requirement.)

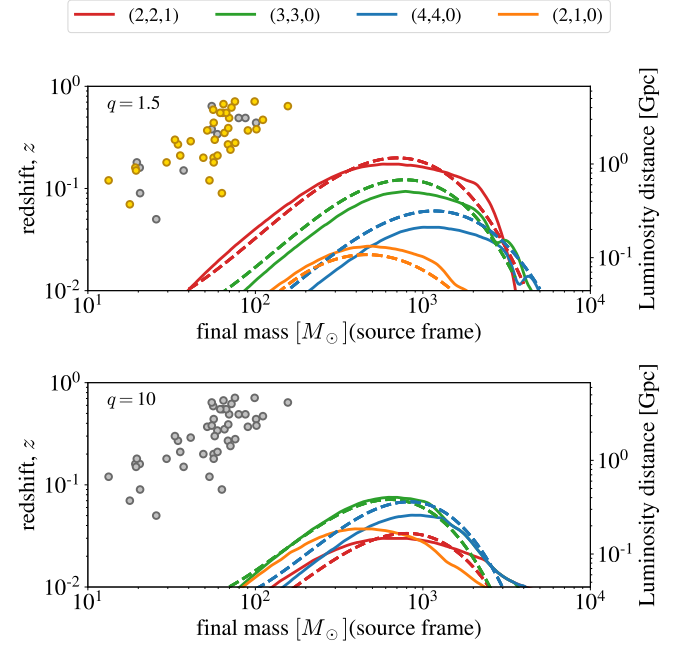


Figure 17. Same as Figure 14, but this time we show with solid lines the LIGO SNR horizons obtained by requiring $\rho_{\ell mn} > 8$ for each subdominant mode (ℓ, m, n) . The SNR horizons are close to the Bayes factor threshold, with the exception of the (2, 1, 0) mode for $q = 10$.

In Figure 17 we show the SNR = 8 horizons, which are much more restricted than the horizons in Figures 14 and 15. The SNR criterion estimates horizons very close to the Bayes factor horizon, indicating that a SNR threshold for the subdominant modes is necessary for a high Bayes factor, but it is not sufficient, as we can see that the (2, 1, 0) SNR horizon for $q = 10$ is very large. We note that our SNR horizons are somewhat smaller than the corresponding ringdown horizons reported for fundamental modes by [23] due to the different mode amplitudes assumed in our analysis, as we take the start time to be $t = t_{\text{peak}} + 10M$ (see also [34]).

Finally, there are additional resolvability criteria proposed in the literature that we do not explore here. In [45], it is proposed that a subdominant mode is detected if its amplitude posterior excludes zero in the 90% credible interval, and they find that this condition is satisfied for the first overtone of a GW150914-like event; this was also confirmed by other works [36, 42, 44]. In [39],

the authors require the similar condition $\sigma_R < R$, for the mode amplitude ratio R to exclude zero at the 1σ level, and they also propose a “measurability” criterion, which requires a relative accuracy for the estimated QNMs parameters.