

# On the spectroscopy of binary black hole ringdown using overtones and angular modes

Xisco Jiménez Forteza<sup>1</sup>, Swetha Bhagwat<sup>2</sup>, Paolo Pani<sup>2</sup>, Valeria Ferrari<sup>2</sup>

<sup>1</sup> *Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Callinstraße 38, 30167 Hannover, Germany and*

<sup>2</sup> *Dipartimento di Fisica, “Sapienza” Università di Roma & Sezione INFN Roma1, Piazzale Aldo Moro 5, 00185, Roma, Italy*

The black hole uniqueness and the no-hair theorems imply that the quasinormal spectrum of any astrophysical black hole is determined solely by its mass and spin. The countably infinite number of quasinormal modes of a Kerr black hole are thus related to each other and any deviations from these relations provide a strong hint for physics beyond the general theory of relativity. To test the no-hair theorem using ringdown signals, it is necessary to detect at least two quasinormal modes. In particular, one can detect the fundamental mode along with a subdominant overtone or with another angular mode, depending on the mass ratio and the spins of the progenitor binary. Also in the light of the recent discovery of GW190412, studying how the mass ratio affects the prospect of black hole spectroscopy using overtones or angular modes is pertinent, and this is the major focus of our study. First, we provide ready-to-use fits for the amplitudes and phases of both the angular modes and overtones as a function of mass ratio  $q \in [0, 10]$ . Using these fits we estimate the minimum signal-to-noise ratio for detectability, resolvability, and measurability of subdominant modes/tones. We find that performing black-hole spectroscopy with angular modes is preferable when the binary mass ratio is larger than  $q \approx 1.2$  (provided that the source is not located at a particularly disfavoured inclination angle). For nonspinning, equal-mass binary black holes, the overtones seem to be the only viable option to perform a spectroscopy test of the no-hair theorem. However this would require a large ringdown signal-to-noise ratio ( $\approx 100$  for a 5% accuracy test with two overtones) and the inclusion of more than one overtone to reduce modelling errors, making black-hole spectroscopy with overtones impractical in the near future.

## I. INTRODUCTION

In the general theory of relativity (GR), the post-merger remnant of a binary black hole (BBH) coalescence settles to a Kerr black hole (BH) at a sufficiently late time after the merger. To reach this stationary state, the perturbed BH remnant emits gravitational waves (GWs) during a process known as the ringdown (RD). The GW corresponding to the RD phase is described by the linear perturbation theory of a Kerr BH [1–8]. The strain takes the following form,

$$h(t) = \sum_{lmn} \mathcal{A}_{lmn} e^{-i\omega_{lmn}t} e^{-(t-t_0)/\tau_{lmn}} {}_s\mathcal{Y}_{lm}. \quad (1)$$

Here  $\omega_{lmn}$  and  $\tau_{lmn}$  are the quasinormal-mode (QNM) frequencies and damping times, respectively, the  $(l, m)$  indices describe the angular decomposition of the modes,  ${}_s\mathcal{Y}_{lm}$  are the spin-weighted  $s = 2$  spheroidal harmonics,  $n$  accounts for the  $n$ -tone excitations of a given  $(l, m)$  mode, with  $n = 0$  being the fundamental “tone”<sup>1</sup>, and  $t_0$  is the starting time of the RD, i.e., a suitable time where the linear perturbation theory is expected to describe the dynamics accurately [9, 10]. The amplitude  $\mathcal{A}_{lmn} = A_{lmn} e^{i\phi_{lmn}}$  is a complex number that depends on the perturbation conditions set up during the inspiral-plunge-merger phase of the BBH evolution and is determined by the mass ratio and the spins of the BBH system.

As a consequence of the BH uniqueness and the no-hair theorems [11–17], the QNM frequencies  $\omega_{lmn}$  and the damping times  $\tau_{lmn}$  define a countably infinite set of modes uniquely related to each other by the final BH mass ( $M_f$ ) and final spin ( $a_f$ ) only.

Customarily, there are at least two approaches to test the BH no-hair theorem with RD signals (see [18, 19] for reviews). The first consists of performing an *inspiral-merger-RD consistency test* [20] that checks whether the final mass and spin of the remnant BH estimated using the inspiral signal only is consistent with the final mass and spin measured from its RD phase only, or from the merger-RD phase. This approach is a null-hypothesis test and assumes that GR is the correct theory of gravity: an inconsistency between the different measurements of  $\{M_f, a_f\}$  (within statistical errors) would provide evidence for new physics beyond GR describing the BBH coalescence. The second and more direct approach is performing *BH spectroscopy*, where one aims at extracting several QNM parameters from the RD signal. These are then used both to measure  $\{M_f, a_f\}$  and to check whether the QNM spectrum is consistent with the Kerr hypothesis [8, 21]. This is a more stringent test but requires a higher signal to noise ratio (SNR) in the RD, as it relies on *independent measurements* of at least two QNM modes. While measuring only the fundamental  $l = m = 2$  QNM frequency and damping time provides an estimate of  $M_f$  and  $a_f$ , subdominant mode parameters (frequencies and damping times of other modes/tones) are needed to perform a consistency test of the QNM spectrum [22]. The premise of this work is to investigate the prospects of performing a BH spectroscopy with BBH RD.

<sup>1</sup> In this paper, we call the overtones simply as the “tones” and the angular modes as “modes”. We shall often refer to a *2-tone* or to a *2-mode* RD model, when considering Eq. (1) with two different tones ( $n = 0, 1$ ) or two different angular modes, respectively.

As a rule of thumb, for a given value of  $\{M_f, a_f\}$ , the frequency of a mode scales approximately as  $\omega_{lmn} \approx \frac{1}{2}\omega_{22n}$ . Thus, the relative difference between the frequencies of two angular modes ( $l \neq l'$ ), is approximately  $\gtrsim 50\%$ . However, the damping times are typically comparable for similar values of  $l$  [7, 8]. Contrarily, the frequencies difference between two overtones (with same angular indices  $(l, m)$  but  $n \neq n'$ ) is typically less than a few percent. For instance, for a GW150914-like BBH event [23]  $1 - \omega_{221}/\omega_{220} \sim 2\%$ . Several earlier studies of BH RD spectroscopy focused solely on the angular modes [22, 24–29] and typically neglected the overtones (with some notable exceptions [27, 30–32]). However, more recently, RD tests with overtones were explored in more details [10, 33, 34] and an attempt to perform such tests with GW150914 data was demonstrated in Ref. [35].

Loosely speaking, the accuracy to which a particular set of modes or tones in a BBH RD allows for a BH spectroscopy depends on the power contained in them. The relative amplitudes to which different RD modes/tones are excited depend on the initial perturbation conditions that are set up during the plunge-merger phase. This, in turn, depends on the mass ratio and the spins of the progenitor BBH. Asymmetric RD modes (such as  $l = m = 3$  and  $l = 2, m = 1$ ) are excited when the merging BHs are either of unequal masses or have misaligned spins. However, the BBH GW events detected by the LIGO-Virgo observations show a high concentration of comparable mass binaries with low spins [36] that prevents extracting higher angular QNMs from their RD.

The recent discovery of a BBH system with mass ratio significantly different from unity, GW190412 [37], provides further motivation to detect asymmetrical RD angular modes from stellar mass BBHs. This state of affairs might become common in the third and future observational runs, where several events with unequal masses and possibly nonvanishing spins are expected.

Thus, it is timely to study the role of overtones and angular modes in the RD signal as a function of the BBH mass ratio. A recent analysis in this direction was performed in Ref. [10] and, more recently, Ref. [34] quantified the amplitude ratio of different tones/modes relative to the dominant  $l = 2, m = 2, n = 0$  mode as a function of the mass ratio for nonspinning binaries.

It is unlikely that a large number of QNMs will be detected in a single event with the current ground-based GW observatories [38], although the third-generation facilities like Einstein Telescope and Cosmic Explorer [39–41] and the space-based mission LISA [42] are expected to detect multiple modes. Therefore, the current and the near-future RD-based GR tests shall rely either on a *2-mode analysis* or on a *2-tone analysis*. This motivates the question we try to address in this work, namely what is the optimal set of QNMs to perform 2-mode/2-tone BH spectroscopy. In particular, we quantify the prospects of observing the fundamental  $l = 2, m = 2, n = 0$  mode along with either:

a) its first overtone, i.e., the  $l = 2, m = 2, n = 1$

mode, or

b) another angular mode, either the ( $l = m = 3, n = 0$ ) or the  $l = 2, m = 1, n = 0$  mode .

The optimal set of QNMs to analyse a BBH RD signal depends on the mass ratio and initial spins of the progenitor BBH. We quantify the notion of an optimal set of QNMs by estimating the minimum SNR required to perform BH spectroscopy with a 2-mode/2-tone analysis.

In order to address the above question, we introduce three specific criteria: the *detectability*, the *resolvability*, and the *measurability* of the QNMs. The detectability criterion demands that the amplitude ratio between the subdominant mode/2-tone and the dominant mode should be nonzero (within measurement errors). Furthermore, the resolvability criterion requires that the subdominant mode/2-tone QNM parameters should be distinguishable from the corresponding parameters of the fundamental mode [8]. Finally, the measurability criterion quantifies the measurement uncertainties in the inference of these parameters and requires that their uncertainty must be smaller than a given threshold.

We estimate the minimum SNR for which each of the three criteria (and combinations thereof) are met using a Fisher information matrix framework. Thus, our estimate of the minimum SNR provides an optimistic lower bound to what is required to perform a Bayesian analysis using GW data. Compared to Ref. [34], our analysis has three notable differences: (i) we assess the role of the starting time of the RD differently; (ii) we use the combination of the three criteria mentioned above – detectability, resolvability, and measurability – as the discriminator between an overtone-based and an angular-mode-based BH spectroscopy, as compared to the criteria used in Ref. [34]; (iii) in our RD model we use the quality factor  $Q_{lmn} \equiv \pi f_{lmn} \tau_{lmn}$  in place of the damping time  $\tau_{lmn}$ . As discussed below, this choice leads to some differences for what concerns the resolvability criterion, but it does not affect our final conclusions.

For this study, we fit for the amplitude  $A_{lmn}$  and phase  $\phi_{lmn}$  in Eq. (1) to a set of numerical relativity (NR) simulations corresponding to nonspinning binaries with mass ratio  $q \in [1, 10]$  from the Simulating eXtreme Space-times (SXS) catalogue [43]. In Secs. II and II C we provide ready-to-use fits for the amplitude ratio and the phase difference of the different modes/tones for various choices of the RD starting time. Then, in Sec. III A we address the issue of choosing a set of subdominant mode/2-tone ( $l, m, n$ ) to perform BH spectroscopy. Finally in Sec. IV we discuss the qualitative aspects of our results.

## II. FITTING NR-RD WAVEFORMS

In this section we discuss our procedure to fit the NR waveforms listed in Table I with the RD model given in Eq. (1) for different modes/tones and discuss the choice of the starting time  $t_0$ .

Sim No.	SXS ID	$q$	$a_f$	$M_f$
1	SXS:0066	1.00	0.686	0.952
2	SXS:0070	1.00	0.686	0.952
3	SXS:1143	1.25	0.680	0.953
4	SXS:0007	1.50	0.664	0.955
5	SXS:1354	1.83	0.638	0.959
6	SXS:0169	2.00	0.623	0.961
7	SXS:0201	2.32	0.596	0.965
8	SXS:0259	2.50	0.581	0.967
9	SXS:0191	2.51	0.580	0.967
10	SXS:0030	3.00	0.541	0.971
11	SXS:1221	3.00	0.541	0.971
12	SXS:0294	3.50	0.504	0.975
13	SXS:1220	4.00	0.472	0.978
14	SXS:0182	4.00	0.472	0.978
15	SXS:0107	5.00	0.417	0.982
16	SXS:0296	5.50	0.393	0.984
17	SXS:0181	6.00	0.372	0.985
18	SXS:0166	6.00	0.372	0.985
19	SXS:0298	7.00	0.337	0.988
20	SXS:0186	8.27	0.300	0.990
21	SXS:0301	9.00	0.282	0.991
22	SXS:1107	10.00	0.261	0.992

TABLE I. Set of NR simulations used in this work obtained from the SXS public catalogue [43, 44]. Here,  $q = m_1/m_2 \geq 1$  is the mass ratio while  $a_f$  and  $M_f$  are the final dimensionless spin and the final mass (in units of the binary total mass), respectively.

### A. Setup

We fit the set of simulations listed in Table I following the prescription described in [10, 33]. For each of the SXS-BBH simulations, we select the NR-RD waveform that corresponds to that with the best resolution (see Appendix A). For each simulation the SXS catalogue provides a mode decomposition,  $h_{lm}$ , in a basis of *spherical* harmonics, which are related to the spheroidal harmonics in Eq. (1) as discussed in Appendix B. The reference time  $t = 0$  is defined as the time at which the amplitude of  $l = m = 2$  mode peaks.

The data set consists of BBH systems with nearly zero initial dimensionless spins  $a_{1,2}$  and mass ratio  $q = m_1/m_2 \in [1, 10]$ . The motivation for neglecting the binary component spins is twofold. First, all binaries detected in the first two observational runs of LIGO/Virgo are compatible with small or negligible spins<sup>2</sup>, except possibly for one event [36]. Second, neglecting the spins reduces the parameter space significantly, since all dimensionless quantities (e.g., mode amplitude ratios and phases) depend solely on  $q$ .

In Table I, we also list the final mass  $M_f$  and spin  $a_f$  of the BH remnant as provided in the metadata of the

SXS catalogue; the details of their computation can be found in Ref. [44]. Loosely speaking, for a nonspinning BBH system, the final BH spin monotonically decreases as a function of  $q$  as a consequence of the conservation of angular momentum [45–49]. Our study spans a range of final spin with  $a_f \approx 0.69$  for  $q = 1$  to  $a_f \approx 0.26$  when  $q \approx 10$ .

The fits are performed by fixing the frequencies  $\omega_{lmn}$  and the damping times  $\tau_{lmn}$  to their corresponding values predicted by GR [8], choosing a starting time  $t_0$ , and then fitting for the complex mode/tone amplitudes  $\mathcal{A}_{lmn}$ . A complex least-square fit is used and the best-fit parameters correspond to the one that minimises the  $\chi^2$  value,

$$\chi^2 = \sum_i |\bar{h}(\vec{\lambda})_i - h_i|^2, \quad (2)$$

where  $\vec{\lambda} = \{A_{lmn}, \phi_{lmn}\}$  are the fit parameters for a given  $t_0$ . To quantify the deviations of the fits with respect to the NR waveform, we compute the mismatch  $\mathcal{M}$  which is defined as

$$\mathcal{M} = 1 - \frac{\langle h_{\text{NR}} | h_x \rangle}{\sqrt{\langle h_{\text{NR}} | h_{\text{NR}} \rangle \langle h_x | h_x \rangle}}, \quad (3)$$

where

$$\langle f | g \rangle = \int_{t_i}^{t_f} f(t)g(t)dt. \quad (4)$$

Here  $h_{\text{NR}}$  is the NR-RD waveform with the highest resolution [43, 44] and  $h_x$  stands respectively for the fit model ( $h_x = h_M$ ) or for the NR waveform with the next-to-the-highest resolution (when  $h_x = h_{\text{LNR}}$ ) when computing the NR error estimates. The integration domain ranges from  $t_i = t_0$  and  $t_f = 60M$  for  $h_M$ , and from  $t_i = \delta t_{lm}^p$  (where  $\delta t_{lm}^p$  is the peak time of the strain  $h_{lm}$ , see Sec. II B 2 below) to  $t_f = 60M$  for  $h_{\text{LNR}}$ .

### B. Role of the starting time

#### 1. Overtones

In Fig. 1, we show the mismatch  $\mathcal{M}$  between the  $l = m = 2$  mode of the NR-RD waveforms from Table I and a *2-tone* RD model – i.e., Eq. (1) with the fundamental mode  $l = m = 2, n = 0$  and its first overtone – as a function of  $t_0$ . The thin horizontal lines in the plot correspond to the NR error estimates for each of the NR-RD simulations (see Appendix A) and are marked using the same colour scheme as their corresponding mismatch curves. These horizontal lines correspond to the mismatch between the two highest resolutions of the NR waveforms available in the SXS catalogue.

The mismatch has a similar behaviour for all mass ratios considered in this study: it decreases monotonically up to a minimum at  $t_0 \in [15M, 20M]$ , and then gradually

<sup>2</sup> Nonetheless, in the light of the very recent BBH detection GW190412 during the third observational run [37], including the binary spins is a natural and urgent extension of our analysis.

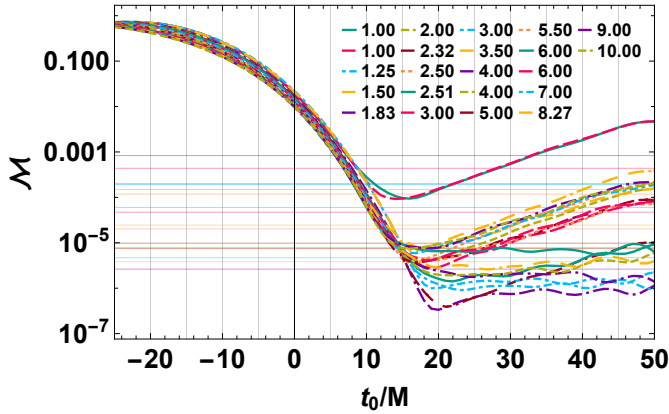


FIG. 1. Mismatch computed between a 2-tone model fit and its respective NR-RD waveform as a function of the starting time  $t_0$  for  $l = m = 2$ . Each curve corresponds to an entry in Table I, while the horizontal lines show the NR error estimates using the same colour scheme (see Fig. 5 below). Note that, for all cases, the error lines cross the solid curves at  $t_0/M \approx 5 - 15$ . The reference time  $t = 0$  is defined by the peak of the  $l = m = 2$  strain mode. The plot starts at  $t_0 = -25M$  (i.e. before the peak of the strain) in order to explore the pre-merger behaviour.

rises up as the NR simulation saturates with numerical noise. One expects that the accuracy of the RD model given by Eq. (1) increases at late times in the BBH post-merger and, therefore,  $\mathcal{M}$  should monotonically decrease. However the mismatch curves reach a minima due to the numerical noise floor, which is exacerbated by the exponential damping of the RD waveform.

Furthermore, we note that when  $t_0/M \approx 5 - 15$  (depending on the value of  $q$ ) or higher, the mismatch between the highest-resolution NR simulation and the analytical RD model is smaller than the corresponding NR error estimate. This happens because at late times the waveform is accurately described by a superposition of the QNM excitations. This suggests that  $t_0/M \in [5, 15]$  would provide a reasonable range for the choice for the starting time for the 2-tone RD model. Ideally, to ensure that the system is in a quasi-linear regime, one would pick the largest possible value of  $t_0$ . However, the fast exponential decay of the  $n = 1$  overtone implies that its amplitude is very small for large values of  $t_0$ . Therefore, we opt for an agnostic strategy<sup>3</sup> and in the following sections we present the results for  $t_0/M = \{0, 5, 10, 15\}$ . We

<sup>3</sup> This is different from Ref. [34], in which  $t_0$  has been identified with the value that minimizes  $\mathcal{M}$  using either the strain or the time-derivative of the phase of the waveform. In such case both methods identify the range  $t_0/M \in [5, 18]$  for  $q \in [1, 10]$ . We did not follow this prescription since the minimum of  $\mathcal{M}$  is a quantity solely determined by the accuracy of the NR waveforms, and not by the faithfulness of the RD model, as explained in the main text. Nonetheless, our prescription identifies a range compatible with that of Ref. [34], namely  $t_0/M \in [10, 15]$ .

emphasise that a 2-tone RD model is insufficient to describe a BBH RD accurately from  $t = 0$ ; higher overtones should be included to ensure accuracy [10, 33, 35]. However, including a large number of overtones is impractical from the point of view of parameter estimation in the context of BH spectroscopy with current detectors [38].

## 2. Angular modes

The choice of the starting time is much less of a problem for angular modes [9] as the damping time of subdominant angular modes (with  $n = 0$ ) is comparable to that of the fundamental  $l = m = 2$  mode. For the case of a 2-mode RD model (with either  $l = m = 3$  or  $l = 2$ ,  $m = 1$  along with the fundamental  $l = m = 2$  mode), we choose two values of the starting time:  $t_0 = 10M$  and  $t_0 = 15M$ . Unlike for the case of a 2-tones model, the choice of the starting time does not affect the amplitude-ratio and phase-difference fits significantly, as we will discuss later in detail.

Furthermore, the amplitudes of different modes do not peak at the same time. We denote the peak amplitude of the  $h_{lm}$  mode as  $\delta t_{lm}^p$  (with  $\delta t_{22}^p = 0$  conventionally). In Fig. 2 we show the time shifts  $\delta t_{lm}^p$  as a function of  $q$  for our set of waveforms. Note that both  $l = 2$ ,  $m = 1$  and  $l = m = 3$  modes peak at a later time compared to the  $l = m = 2$  mode (i.e.  $\delta t_{lm}^p > 0$  for these cases), for all values of  $q$ . The time shifts  $\delta t_{21}^p$  corresponding to the  $l = 2$ ,  $m = 1$  mode (green circles) approximately increases linearly with  $q$  up to  $\delta t_{21}^p \approx 9M$  at  $q = 10$ , while  $\delta t_{33}^p$  (red squares) tends to be approximately constant ( $\delta t_{33}^p \approx 4M$ ) for all  $q$ 's. Thus, our choices  $t_0 = 10M$  and  $t_0 = 15M$  ensure that  $t_0 > \delta t_{lm}^p$ , i.e., that we start the RD analysis after all the modes peak.

## C. Amplitude and phase fits

Using the procedure described in the previous section, we can compute the fits for the amplitude and phases of each mode/tone. Since  $\mathcal{A}_{lmn} = A_{lmn}e^{i\phi_{lmn}}$  in Eq. (1), the total phase for the  $(lmn)$  mode is

$$\Phi_{lmn}(t) = \omega_{lmn}t - \phi_{lmn}. \quad (5)$$

The phase  $\phi_{lmn}$  depends on the initial orbital phase of the BBH system, which is generically different for different NR simulations. Thus, to make a meaningful comparison among different simulations, we align the phases by adding an extra constant phase to the NR waveform for each mode, such that  $\Phi_{lmn}(10M) = 0$ . The choice of the reference time  $t = 10M$  is arbitrary and does not affect our results.

We choose to fit the amplitude ratio,

$$A_{R,lmn} = \frac{A_{lmn}}{A_{220}}, \quad (6)$$

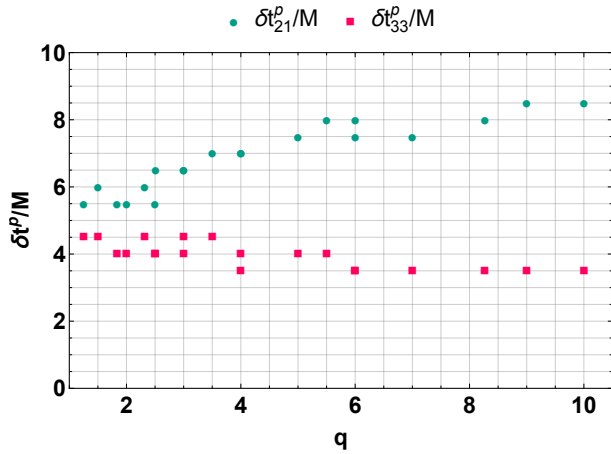


FIG. 2. Difference between the time at which the (22) mode strain peaks (conventionally chosen at  $t = 0$ ) and the peak time of the higher harmonics,  $\delta t_{lm}^p$ . The (21) peak time  $\delta t_{21}^p$  raises up as  $q$  increases and eventually reaches  $\delta t_{21}^p \approx 9M$ . On the other hand,  $\delta t_{33}^p$  remains approximately constant,  $\delta t_{33}^p \approx 4M$ . In both cases  $\delta t_{lm}^p$  is smaller than the starting time chosen for the  $(lm)$  modes fit,  $t_0 = 10M$  or  $t_0 = 15M$ .

and the phase difference

$$\Delta\phi_{lmn} = \phi_{220} - \phi_{lmn}, \quad (7)$$

relative to the  $l = m = 2, n = 0$  mode.

We find that these quantities can be conveniently expressed by the following closed-form approximations

$$A_{R,lmn} = a_{0,lmn} + \frac{a_{1,lmn}}{q} + \frac{a_{2,lmn}}{q^2} + \frac{a_{3,lmn}}{q^3}, \quad (8)$$

$$\Delta\phi_{lmn} = b_{0,lmn} - \frac{b_{1,lmn}}{b_{2,lmn} + q^2}, \quad (9)$$

where the fit parameters  $a_{i,lmn}$  and  $b_{i,lmn}$  are obtained for different choices of  $t_0$ . Their values are listed in Tables II, III, and IV, for the 2-tone model ( $l = m = 2, n = 1$ ) and for both the 2-mode models ( $l = m = 3$  or  $l = 2, m = 1$ , both with  $n = 0$ ), respectively. For the subdominant angular modes, the parameter  $a_{0,lmn}$  is fixed in terms of the others by requiring that  $A_{R,210} = A_{R,330} = 0$  at equal mass ratio,  $q = 1$ , as imposed by symmetry arguments.

These fits are discussed in the following sections.

### 1. Assessing the accuracy of the RD model by varying the starting time

The RD waveform model (1) has the following symmetry

$$\begin{cases} t_0 \rightarrow t_0 + \Delta t \\ A_{lmn} \rightarrow A_{lmn} e^{-\frac{\Delta t}{\tau_{lmn}}} \\ \phi_{lmn} \rightarrow \phi_{lmn} \end{cases}. \quad (10)$$

$t_0$	$0M$	$5M$	$10M$	$15M$
$a_{0,221}$	0.901327	0.641972	0.373974	0.170129
$a_{1,221}$	0.705107	0.297579	0.074412	0.10284
$a_{2,221}$	-0.386356	0.218103	0.416288	0.0918048
$a_{3,221}$	0.045801	-0.25445	-0.322963	-0.0765307
$b_{0,221}$	0.426307	0.396928	0.381152	0.415696
$b_{1,221}$	5.59052	4.84138	6.3857	6.70373
$b_{2,221}$	56.7092	18.6836	14.9772	13.2363

TABLE II. Fitting coefficients of Eqs. (8) and (9) for the 2-tone model (i.e.  $l = m = 2$  fundamental mode plus its first overtone) for different values of the starting time  $t_0$ .

$t_0$	$10M$	$15M$
$a_{0,210}$	0.473846	0.479966
$a_{1,210}$	-1.22756	-1.23848
$a_{2,210}$	1.61047	1.61757
$a_{3,210}$	-0.85676	-0.859064
$b_{0,210}$	1.8082	1.82032
$b_{1,210}$	9.9702	8.79577
$b_{2,210}$	10.3096	9.30836

TABLE III. Same as Table II for for the 2-mode RD model with  $l = m = 2$ , and  $l = 2, m = 1$  fundamental modes.

$t_0$	$10M$	$15M$
$a_{0,330}$	0.439698	0.437926
$a_{1,330}$	-0.611581	-0.651738
$a_{2,330}$	0.199865	0.301015
$a_{3,330}$	-0.0279826	-0.0872038
$b_{0,330}$	2.66306	2.68764
$b_{1,330}$	6.81421	6.39255
$b_{2,330}$	6.65011	6.03077

TABLE IV. Same as Table II for for the 2-mode RD model with  $l = m = 2$ , and  $l = m = 3$  fundamental modes.

Thus, if Eq. (1) is a faithful model for the real signal, shifting the starting time and rescaling the amplitudes as presented in the above formula should not affect our best-fit parameter values. Conversely, the dependence of the fit parameters on the starting time after the above rescaling would imply that the Eq. (1) (for a certain number of modes/tones) does not reproduce the NR RD accurately.

We expect this symmetry to be broken for small values of  $t_0$  for the following reasons: (i) we included only one overtone ( $n = 1$ ), whereas higher overtones with  $n > 1$  should be included, especially for small values of  $t_0$  [35]; and (ii) there could be possible nonlinearities in the source frame that cannot be accounted for by a superposition of the QNMs in the asymptotic frame, especially near the peak at  $t \approx 0$  [10, 30, 32, 50–52]. These inaccuracies in the 2-tone RD model lead to systematic errors when performing BH spectroscopy. In Ref. [10], we study these effects and their impact on the QNM spectrum and find that the inferred QNM frequencies can be significantly biased if  $t_0$  is chosen very close to the time of peak amplitude. In principle, one can alleviate this prob-

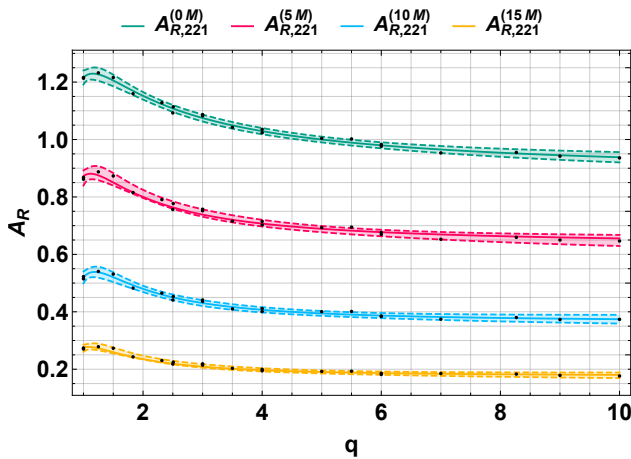


FIG. 3. Amplitude ratio fits  $A_{R,221}$  for the  $n = 1$  overtone computed for the 2-tone RD model using Eq. (8) and the fitting coefficients in Table II for four different starting times,  $t_0/M = \{0, 5, 10, 15\}$ . Solid curves denote the best fit for each case while shaded regions denote the 90% credible range. The agreement between the data (dots) and the best fit is good for all mass ratios and choices of  $t_0$ . The amplitude among different choices of  $t_0$  is not rescaled as in Eq. (10), see Fig. 4 for the corresponding rescaled quantity.

lem by choosing a sufficiently large value of  $t_0$ , but this also reduces the power contained in the highly-damped  $n = 1$  overtone.

Finally, note that the symmetry (10) can be broken when either the amplitude is not rescaled appropriately or when the phase  $\phi_{lmn}$  does not remain constant upon shifting  $t_0$ .

### 2. Fits for overtones

In Fig. 3, we show the amplitude ratio  $A_{R,221}$  (see Eq. (6) and Table II) for the 2-tone RD models corresponding to four different starting times  $t_0/M = \{0, 5, 10, 15\}$ . The fits reproduce the NR amplitude ratio and phase difference for all the simulations considered in this study with a 90% credibility. For a given choice of  $t_0$ , the amplitude ratio for the overtones decreases as the mass ratio increases until the slope flattens out and  $A_{R,221} \rightarrow a_{0,221}$  as  $q \gg 1$ .

The ratio between the test-particle limit  $a_{0,221} = A_{R,221}^{q \rightarrow \infty}$  and the equal mass-ratio case  $A_{R,221}^{q=1}$  is  $\{0.71, 0.71, 0.69, 0.59\}$  for the four values of  $t_0$  considered in the plot. This shows that – independent of the choice of  $t_0$  – the relative amplitude does not vary significantly with the mass ratio and can be used to improve the coherent mode stacking algorithm such as the one outlined in Refs. [34, 53]. This is especially important for third-generation GW interferometers, where one expects to detect  $\mathcal{O}(100 - 1000)$  BBH RDs. Improvements to current stacking algorithms using this empirical observation will be explored in a forthcoming work.

Note that the amplitude ratios shown in Fig. 3 is not rescaled by Eq. (10) for different choices of  $t_0$ . The rescaled amplitude ratios and phase differences are presented in the left and right panels of Fig. 4, respectively. To rescale the ratio in Eq. (6), we use Eq. (10) for both the numerator and the denominator, i.e.,

$$A_{R,lmn} = \frac{A_{lmn}}{A_{220}} \rightarrow A_{R,lmn} e^{-\Delta t(1/\tau_{lmn} - 1/\tau_{220})}, \quad (11)$$

and we rescale the amplitude ratios with respect to that at the reference starting time  $t_0 = 10M$ .

As shown in the left panel of Fig. 4, the rescaled ratios at  $t_0 = 0$  and  $t_0 = 5M$  (green and red curves) are not compatible with the 90% confidence intervals for all mass ratios and they break the symmetry (10) by a large amount. The same is true for  $t_0 = 10M$  and  $t_0 = 15M$  (blue and yellow curves) although, in this case, the differences are smaller for all  $q$ 's considered.

The situation is similar for the phase difference  $\Delta\phi_{221}$ . From Eq. (10), we should expect that  $\Delta\phi_{221}$  be the same independently of  $t_0$ . Instead, we observe that, for fixed  $q$ ,  $\Delta\phi_{221}$  depends on  $t_0$  especially when  $t_0 = 0M, 5M$  (green and red curves) and mostly for low mass ratios. The differences between the  $t_0 = 10M$  and the  $t_0 = 15M$  curves (yellow) are at most about 10% ( $\approx 0.1$  rad absolute error) for low mass ratios, and are within the NR errors on the phase. In Appendix A we estimate the latter to be approximately  $\pm 0.04$  rad.

Using the symmetry criterion in Eq. (10), we argue that one need to wait for at least  $t_0 \geq 10M$  after the peak amplitude to start a reasonable accurate RD analysis with a 2-tone RD model.

In Fig. 5, we show the mismatch between the fits and the NR simulations computed using Eq. (3) for  $t_0/M = \{0, 5, 10, 15\}$  (green to yellow markers). We also show the NR error estimates similar to Fig. 1 (purple triangles). Furthermore, from Fig. 5, we observe that the NR errors can be important at  $t = 15M$  for some of the simulations considered in this study. Loosely speaking, the mismatch decreases as  $t_0$  increases, confirming that a 2-tone model provides a more accurate fit at late times. The mismatch reduces approximately by an order of magnitude for every  $5M$  increase in  $t_0$ .

### 3. Fits for angular modes

Fixing  $t_0 = 10M$  and  $t_0 = 15M$  as a reference starting time, we fit for the amplitude ratio and the phase difference of a 2-mode RD model<sup>4</sup>. The SXS catalogue provides the waveform decomposed into spherical harmonic modes and we use this to perform fits for each angular modes independently. Specifically, we consider a signal

<sup>4</sup> Recall that  $t_0 \geq 10M$  ensures that the peak of the subdominant angular modes is contained in the data.

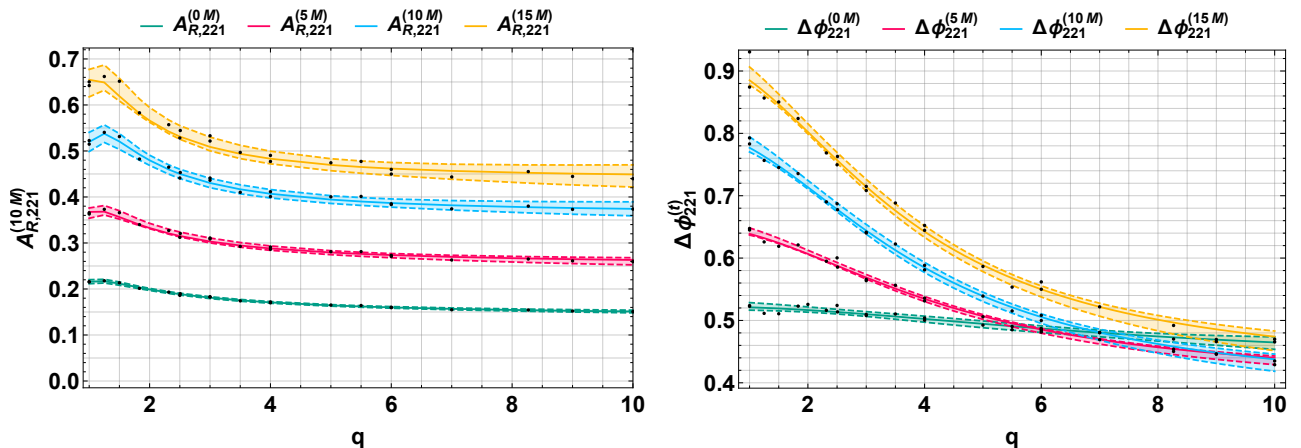


FIG. 4. Fits for the rescaled amplitude ratio  $A_{R,221}$  (left panel) and the phase difference (right panel) as a function of  $q$ . The rescaling (10) is performed using  $t_0 = 10M$  as reference starting time. The bands denote the 90% credible intervals for each fit. The symmetry (10) is broken for  $t_0 = 0M, 5M$  by a large amount and to a minor extent for  $t_0 = 15M$ . Although the variations of the phase difference  $\Delta\phi_{lmn}$  for  $t_0 = 10M, 15M$  are larger than the corresponding 90% credible regions, they are within the phase uncertainty of the NR waveforms, see Appendix A

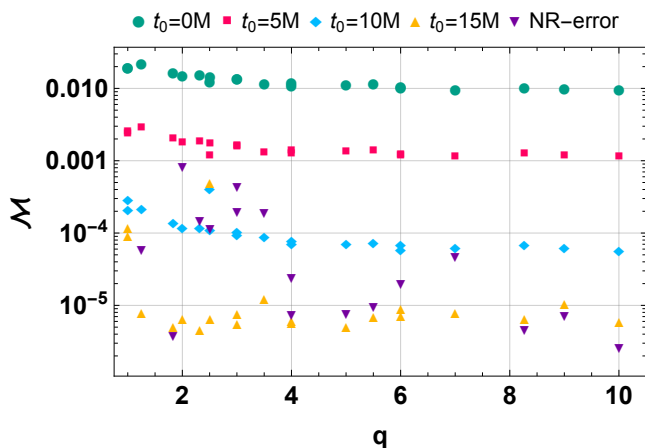


FIG. 5. Measure of the accuracy of the fits for the 2-tone RD model in terms of mismatch between the fits and the NR waveform for  $t_0/M = \{0, 5, 10, 15\}$  (green to yellow markers). The mean values for the mismatch are  $\{0.0134, 0.0016, 0.00012, 0.00004\}$  for  $t_0/M = \{0, 5, 10, 15\}$ , respectively. As a reference, we also show the NR error (purple triangles) as in Fig. 1 (see Appendix A for details).

with a secondary mode with either  $l = m = 3$  or  $l = 2$ ,  $m = 1$ , and  $n = 0$  in both cases.

It is worth noticing that the phase difference and the amplitude ratio presented in this section are inherent to the BBH system and do not include extrinsic factors such as the inclination angle. To perform an analysis on the astrophysical BBH events, one needs to add the effect of the inclination angle as it modifies the amplitude ratio and phase difference between the different angular modes. See Appendix B for details.

In Fig. 6 we show the fits for the rescaled amplitude

ratio (left panel) and the phase difference (right panel) corresponding to  $t_0 = 10M, 15M$ . To illustrate the fit accuracy, we also provide an overlay of the data points that have been fitted. The fits are reasonably accurate (within their own error bars) for both choices of starting time  $t_0$ . A few data points are scattered with respect to the best-fit curve, we attribute this to the inaccuracy of some NR waveforms. Similarly, we notice that  $A_{R,210}^{(15M)}$  happens to lose accuracy at low mass ratio regime, due to errors of the corresponding NR simulations (see Fig. 5). Note also that for subdominant angular modes the deviation of the rescaling relative to the symmetry (10) is much smaller than the one observed in Fig. 4 for overtones. This happens because the damping factors  $\tau_{lmn}$  for different angular modes ( $l, m$ ) of equal tone index  $n$  are comparable. We also note that for  $t_0 = \{10M, 15M\}$ , the values for  $A_{R,210}$  and  $A_{R,330}$  increase monotonically with  $q$  until they asymptote to  $A_{R,210} = a_{0,210}$  and  $A_{R,330} = a_{0,330}$  in the high- $q$  limit. For these modes, the amplitude ratio is constrained approximately in the range  $A_{R,lmn} \in [0, 0.5]$ . For the phase difference  $\Delta\phi_{lmn}$  we observe two opposite trends for the modes:  $\Delta\phi_{330}$  monotonically decreases with the mass-ratio  $q$  whereas the opposite is true for  $\Delta\phi_{210}$ . The phase  $\Delta\phi_{lm0}$  is constrained within  $[0.5b_{210}, b_{210}]$  and  $[b_{330}, 1.3b_{330}]$  for the two modes considered here. This result is independent of the arbitrary phase alignment at  $t_0 = 10M$  that is used in this study.

#### 4. Test-particle limit

In the test-particle limit ( $q \rightarrow \infty$ ), our fitting formulas (8) and (9) reduce to  $A_{R,lmn} \rightarrow a_{0,lmn}$  and  $\Delta\phi_{lmn} \rightarrow b_{0,lmn}$ . For  $t_0 = 10M$ , we obtain

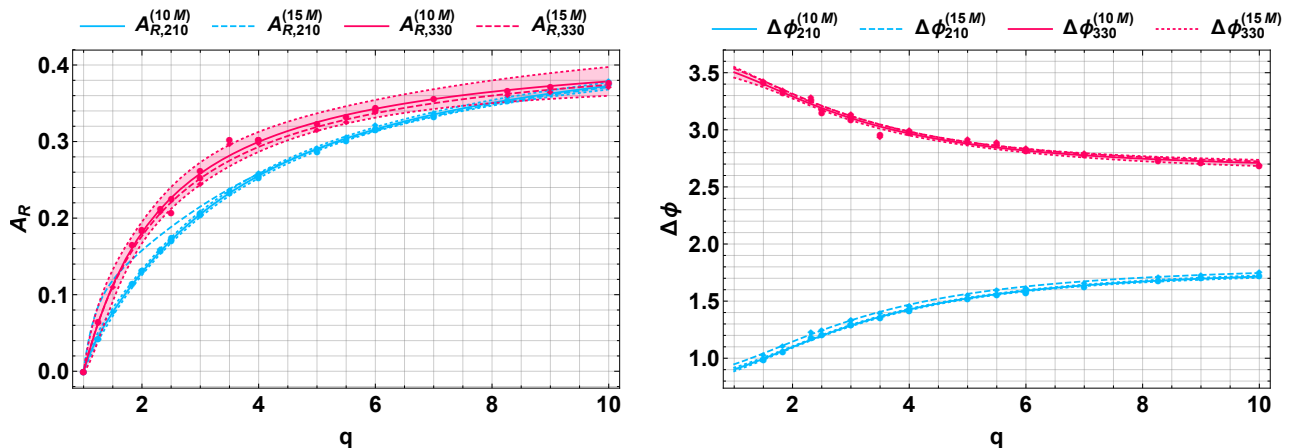


FIG. 6. Fits for the rescaled amplitude ratio and phase difference as a function of  $q$  using the fundamental  $l = m = 2$  mode and different (fundamental) angular modes (either  $l = m = 3$  or  $l = 2, m = 1$ ). Unlike for the overtone case shown in Fig. 4, in this case the symmetry (10) is preserved between  $t_0 = 10M$  (solid) and  $t_0 = 15M$  (dashed). The bands show the 90% credible intervals of the  $t_0 = 10M$  fit. The NR data are shown with solid circles and diamonds at  $t_0 = 10M$  and  $t_0 = 15M$  respectively.

$a_{0,lmn} = \{0.374 \pm 0.07, 0.474 \pm 0.1, 0.440 \pm 0.03\}$  for the ( $l = m = 2, n = 1$ ), ( $l = 2, m = 1, n = 0$ ) and ( $l = m = 3, n = 0$ ) modes respectively, where the uncertainties mark the 90% credible intervals. Our fits are consistent with previous work on the test-particle limit of QNM excitation factors. For instance, we obtain results similar to Ref. [54], wherein the authors estimated the test-particle limit amplitude ratio  $\approx 0.43$  for the  $l = m = 3$  mode by fitting a set of nonspinning NR simulations. A different approach was taken in Ref. [55], where the authors provide an independent estimates of the test-particle limit for amplitudes of the  $l = m = 3$  and  $l = 2, m = 1$  modes computed at the peak of the strain,  $t_0 = 0$ , by solving numerically the Teukolsky equation. Their results at  $t_0 = 10M$  translate<sup>5</sup> to  $\{0.36, 0.40\}$  for the amplitude ratio of the  $l = m = 3$  and  $l = 2, m = 1$  modes, respectively. Although these values are compatible with our fit errors, the small difference could arise from the numerical errors of the NR waveforms analysed in this study, especially in the high mass-ratio regime. Furthermore, in Ref. [34] the amplitude ratio in the test-particle limit is estimated to be approximately 0.3 – 0.4 for ( $l = 3, m = 3, n = 0$ ), ( $l = 2, m = 1, n = 0$ ), and ( $l = 2, m = 2, n = 1$ ), also in agreement with our fits.

For what concerns the phase difference, to the best of our knowledge no similar estimates have been made for in the test particle limit that can be used to benchmark our results. For  $t_0 = 10M$ , we obtain the following phase difference in the test-particle limit:  $b_{0,lmn} = \{0.381 \pm 0.018, 1.808 \pm 0.080, 2.663 \pm 0.103\}$  for the ( $l = m = 2, n = 1$ ), ( $l = 2, m = 1, n = 0$ ), ( $l = m = 3, n = 0$ ) modes, respectively. The trend descends mono-

tonically for  $\Delta\phi_{221}$  and  $\Delta\phi_{330}$ , whereas it increases for  $\Delta\phi_{210}$ . Since these values depend on the phase alignment (see Sec. II C), they may be freely shifted by a constant phase term. However, the quantity physically relevant  $\Delta\phi_{lmn}(q = 1) - b_{0,lmn}$  remains constant independently of the phase alignment. Thus, at  $t_0 = 10M$ , we obtain  $\Delta\phi_{lmn}(q = 1) - b_{0,lmn} = \{0.400, -0.882, 0.891\}$  for the ( $l = m = 2, n = 1$ ), ( $l = 2, m = 1, n = 0$ ), ( $l = m = 3, n = 0$ ) modes, respectively.

### III. QUANTIFYING THE PROSPECTS FOR BH SPECTROSCOPY WITH OVERTONES OR ANGULAR MODES

RD signals from BBH mergers serve as a powerful probe to test the nature of the remnant compact object as well as the behaviour of gravity around it. Experimental validation of certain predictions of GR can be performed to some extent even when explicit spectroscopy of BBH RD is not feasible. For example: i) an inspiral-merger-RD consistency test is based on the comparison of  $\{M_f, a_f\}$  of the final BH, estimated from the inspiral phase, to those estimated directly from the merger-RD phase; ii) even when the signal is not strong enough to measure the subdominant QNMs confidently, an inference can be made that disfavors large deviations from GR by using methods such as the Bayes factor, or by checking whether there is strong support of the posterior distribution of the subdominant QNM parameters in unexpected frequency ranges. These tests are, however, not as stringent as performing an explicit BH spectroscopy and confirming that a *measurement* of the QNM mode frequencies and damping times is consistent with the prediction of GR. BH spectroscopy can then be used to validate the predictions of GR such as the no-hair

<sup>5</sup> To translate the results we need to apply Eq. (11) with  $\Delta t = 10M$ .



theorem, the BH uniqueness theorem and, indirectly, the area-increase theorem.

We note that in this section all the mode amplitudes are computed starting the RD analysis at  $t = t_0$  (see Tables II, III, and IV) without performing the rescaling discussed above.

### A. Detectability, resolvability, and measurability criteria for BH spectroscopy

The possibility to perform BH spectroscopy relies on some necessary criteria that a given RD signal should satisfy. To quantify this issue, for a 2-mode or 2-tone RD model we define the following criteria:

#### 1. Detectability criterion:

$$\sigma_{A_R} < A_R \quad (12)$$

where  $A_R$  is the amplitude ratio between the dominant and the sub-dominant mode, and  $\sigma_{A_R}$  is the uncertainty in the recovery of the parameter  $A_R$ . We name this as the detectability criterion as it is a necessary (albeit not sufficient) condition to claim the presence of the subdominant mode in the RD signal at a  $1\sigma$  level<sup>6</sup>.

#### 2. Resolvability criterion:

$$\max[\sigma_{f_{220}}, \sigma_{f_{\text{sub}}}] < |f_{220} - f_{\text{sub}}|, \quad (13a)$$

$$\max[\sigma_{Q_{220}}, \sigma_{Q_{\text{sub}}}] < |Q_{220} - Q_{\text{sub}}|, \quad (13b)$$

where  $\sigma_X$  is the uncertainty in recovering a quantity  $X$ ,  $f_{220}$  and  $f_{\text{sub}}$  (resp.,  $Q_{220}$  and  $Q_{\text{sub}}$ ) are the QNM frequencies (resp., quality factors, with  $Q_{lmn} = \pi f_{lmn} \tau_{lmn}$ ) of the dominant mode and subdominant mode, respectively. Henceforth, we call this the resolvability criterion as it ensures that a putative measurement of the subdominant mode can be resolved from that of the dominant mode at a  $1\sigma$  level. Note that this is the traditional Rayleigh resolvability criterion that was introduced in the context of BBH RD in Refs. [54, 56]. The resolvability criterion requires that the RD must satisfy either Eq. (13a) or Eq. (13b), depending on which quantity is measured (see below). We expect that for angular modes it is much easier to satisfy Eq. (13a), whereas for overtones it is much easier to satisfy Eq. (13b), since their frequency is close to that of the fundamental mode.

#### 3. Measurability criterion:

$$\left\{ \frac{\sigma_{f_{220}}}{f_{220}}, \frac{\sigma_{Q_{220}}}{Q_{220}}, \frac{\sigma_{f_{\text{sub}}}}{f_{\text{sub}}} \right\} \leq T, \quad (14a)$$

$$\left\{ \frac{\sigma_{f_{220}}}{f_{220}}, \frac{\sigma_{Q_{220}}}{Q_{220}}, \frac{\sigma_{Q_{\text{sub}}}}{Q_{\text{sub}}} \right\} \leq T, \quad (14b)$$

where  $T$  is a given threshold. In other words, we require that at least three QNM parameters are measured within a relative accuracy  $T$ . For concreteness, below we shall consider  $T = \{1, 5, 10\}\%$ . The combined requirement of measurability and resolvability imposes Eqs. (14a) and (13a) or Eqs. (14b) and (13b).

We define the minimum SNR that allows for a spectroscopic analysis of the BH as the one for which *all of the above three conditions* (i.e., either Eqs. (12), (13a) and (14a) or Eqs. (12), (13b) and (14b)) are satisfied.

We estimate the errors on the RD waveform parameter using a Fisher matrix framework, which is valid in the high SNR limit and when the statistical properties of the noise can be assumed to be Gaussian. Therefore, our estimates of the minimum SNR required for performing BH spectroscopy are *optimistic* lower bounds. The details of the Fisher-matrix framework used in this study are similar to that described in Sec. III of Ref. [10], to which we refer for technical aspects. In particular, we note that the analysis does not depend significantly of the noise curve of the detector, since the latter is approximately flat in the frequency range of interest [62]. Since the amplitude ratio for different angular modes depends on the inclination angle of the source, we have averaged out the GW strain on the location angles  $\theta$  and  $\phi$ , as discussed in Appendix B.

### B. Progenitor mass ratio and its effect on BH spectroscopy

The amplitudes to which different RD modes are excited depend on the properties of the progenitor system, in particular the mass ratio and spins of the two coalescing BHs. As a rule of thumb, more the asymmetry in the BBH system i.e., higher the mass ratio and spins, larger is the excitation of angular modes. However, the dependence of overtone excitation on the progenitor mass ratio is less intuitive. We note from Fig. 4 that the amplitude ratio for the  $n = 1$  overtone slightly decreases as the mass ratio of the BBH increases, while the opposite is true for angular modes (see Fig. 6).

The  $l = m = 2$ ,  $n = 0$  mode is always the dominant one when the progenitor BBH system undergoes inspiral, plunge and merger; these are the main signals of interest for GW RD tests using LIGO/Virgo. As in the previous section, we consider the following three subdominant modes: a)  $l = m = 2$ ,  $n = 1$ ; b)  $l = m = 3$ ,  $n = 0$ ; c)  $l = 2$ ,  $m = 1$ ,  $n = 0$ .

<sup>6</sup> See also Ref. [54], where a different criterion for amplitude resolvability of angular modes was adopted.

As discussed in our earlier work [10], picking the optimal value of starting time depends on the interplay between the systematic error of RD modelling and statistical uncertainty due to SNR. For overtones, here we adopt a more agnostic approach and perform the analysis using four different choices of the starting time,  $t_0/M = \{0, 5, 10, 15\}$ . As discussed in the previous section, the fact that the symmetry (10) is broken for  $t_0 = 0, 5M$ , and to a minor degree at  $t_0 = 15M$ , shows that a simple 2-tone RD model is not accurate at early times, and one should include higher-order overtones to accurately capture features close to the peak amplitude of the waveform [33, 35]. We include  $t_0 = 0, 5M$  for the sake of completeness and because using two tones is the minimum requirement to perform BH spectroscopy (measuring *three* modes/tones will require even higher SNR). Nonetheless, as discussed below, our qualitative results do not significantly depend on the choice of  $t_0$ .

On the other hand, from Fig. 6 we note that the amplitude ratio between the angular modes does not change significantly with the starting time of RD, and therefore, for these, we assume that the RD starts  $15M$  after the peak amplitude of the waveform (note that any other choices in the range  $t_0 \in [10, 15]M$  would not change our analysis significantly).

Although the total mass of the progenitor BBH system is a simple rescaling factor for all dimensionful quantities, for concreteness we set  $m_1 + m_2 = 70M_\odot$  and rescale all quantities accordingly.

We can now proceed to study the minimum SNR required to satisfy the detectability, resolvability and measurability criteria for different values of  $q$ .

First, we investigate the minimum SNR required to satisfy the three criteria listed in Sec. III A individually and present the results in Figs. 7, 8 and 9. Then, in Fig. 10 (which is one of the main results of this work), we combine the three criteria and provide the minimum SNR required in the RD to allow for BH spectroscopy using angular modes and overtones. We note that the results of this section are based on the amplitude ratios and phase difference obtained for the list of NR simulations in Table I. In Appendix D we provide contour plots for the minimum SNR needed to satisfy the three criteria above as a function of  $A_R$  and  $\Delta\phi$ , i.e., for generic initial configurations.

In Fig. 7 we present the minimum SNR required in the RD to satisfy the detectability criterion as a function of the mass ratio of the BBH system. We find that the SNR required for detectability of the subdominant angular mode rapidly decreases with increasing mass ratio while that for overtones slowly increases. For a nearly equal-mass system, an extremely high SNR is required for detecting the subdominant angular mode (asymptoting to infinity for an equal mass system) while an SNR between 8 to 15 is sufficient for the case of overtone, depending on the starting time of RD assumed in the analysis. From this plot we see that subdominant angular modes become a more promising candidate for detection

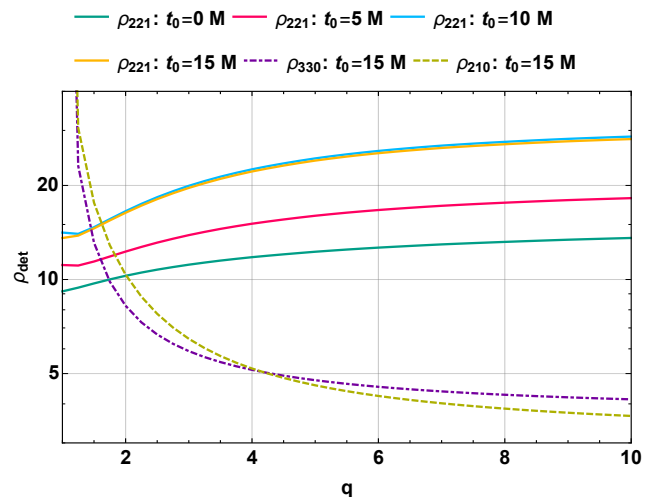


FIG. 7. The minimum SNR required to detect the subdominant mode [see Eq. (12)]. The dotted purple curve corresponds to the angular mode  $l = m = 3, n = 0$  and the dashed green curve corresponds to  $l = 2, m = 1, n = 0$  mode, with  $t_0 = 15M$  as starting time. The continuous teal, pink, blue and orange curves show the minimum SNR for detectability of the overtone for different starting times,  $t_0/M = \{0, 5, 10, 15\}$ .

when  $q \gtrsim 1.5$ .

Next, in the left panel of Fig. 8 we present the minimum SNR required in the RD to satisfy the resolvability criterion<sup>7</sup> for the frequency, Eq. (13a). We find that, except for nearly equal-mass BBHs, it is always easier to resolve the subdominant angular mode frequency from the dominant mode frequency. The angular mode  $l = m = 3$  generally has a QNM frequency that is separated from the dominant mode frequency by tens of Hz (for total mass  $m_1 + m_2 = 70M_\odot$ ) and can be resolved even at a low SNR,  $\rho \sim 3$ . For angular modes, therefore, frequency resolvability is not a limiting factor for BH spectroscopy unlike for the case of overtones.

Furthermore, in the right panel of Fig. 8 we present minimum SNR required to resolve the quality factor of the sub-dominant mode  $Q_{\text{sub}}$  from that of the dominant mode, Eq. (13b). We find that  $Q_{\text{sub}}$  can be resolved for

<sup>7</sup> Our results for the resolvability of angular modes agree (both numerically and analytically) with those of Ref. [56] (the latter differ by a factor of  $2\pi$  with the results of Ref. [54]). However, note that, differently from Refs. [54, 56] we also include  $\Delta\phi$  as one of the parameters of the Fisher matrix. Compared to the resolvability criterion used in Ref. [34], our results for the overtones and angular modes are in agreement, after taking into account two notable differences: (i) we use  $Q_{\text{sub}}$  instead of  $\tau_{\text{sub}}$  in the RD model. Since  $dQ_{\text{sub}}/Q_{\text{sub}} \sim df_{\text{sub}}/f_{\text{sub}} + d\tau_{\text{sub}}/\tau_{\text{sub}}$ , it is easier to resolve  $Q_{\text{sub}}$  rather than  $\tau_{\text{sub}}$  (we thank Cecilia Chirenti and Lara Ota for this comment). Assessing whether using  $Q_{lmn}$  or  $\tau_{lmn}$  in a proper RD analysis is more convenient is an interesting question that we postpone for future work; (ii) we require resolvability of *either* frequencies *or* quality factors, whereas Ref. [34] required resolvability of both frequencies and damping times.

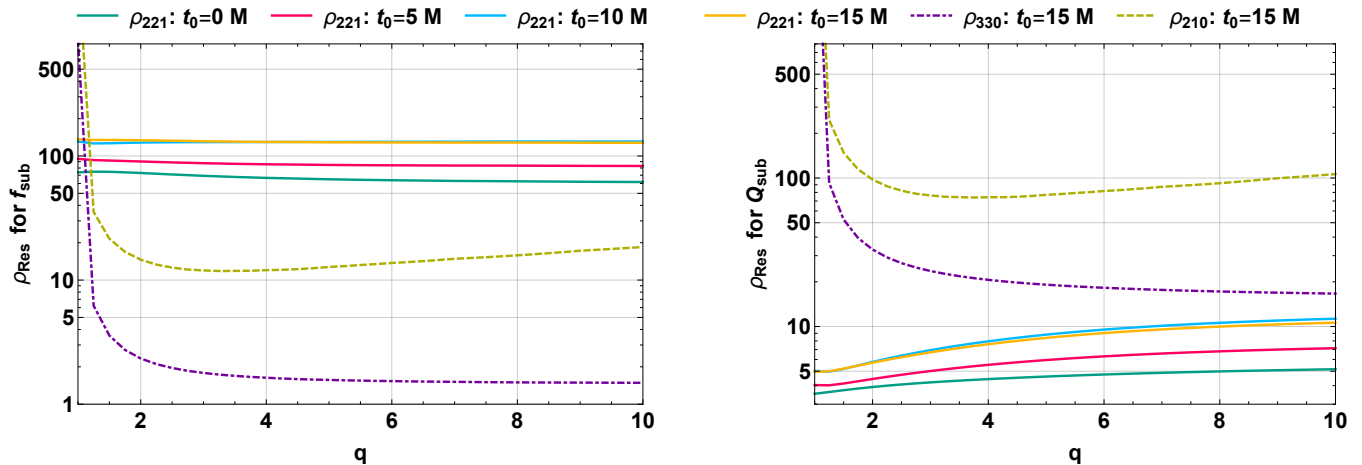


FIG. 8. The resolvability criterion [see Eq.(13)]. We show the minimum SNR required to resolve the subdominant mode frequency  $f_{\text{sub}}$  (left panel) and quality factor  $Q_{\text{sub}}$  (right panel) from the dominant mode. The colour scheme is identical to that used in Fig. 7. Notice that  $Q_{\text{sub}}$  for the overtones and  $f_{\text{sub}}$  for the angular modes are easily resolved.

$\rho \approx 10$  or smaller for the case of overtones for all mass ratios considered in this study. Conversely, for nearly equal mass-ratio systems, resolving  $Q_{\text{sub}}$  for angular modes requires very high SNR:  $\rho \sim 20 - 100$  for  $q \gtrsim 2$ .

In Fig. 9 we present the minimum SNR required to measure the subdominant mode frequency (left panel) and quality factor (right panel) to a precision of 10%, 5%, and 1% (top to bottom). The uncertainty in the measurements of QNMs using overtones slowly increases at higher mass ratios, whereas the uncertainty in the measurements of QNMs using angular mode decreases rapidly as the mass ratio of the progenitor BHs of the BBH system increases. We find that, for a mass ratio greater than 1.5, the  $l = m = 3, n = 0$  subdominant angular mode allows for the most precise measurement among the subdominant modes considered in this study. For a near equal mass system, since angular modes can be measured poorly, the overtones provide a much better measurement precision.

Furthermore, comparing the SNR values in the left panel with that in the right panel, we infer that, for a given SNR, the frequency of a given mode can be estimated with higher accuracy than its quality factor, as expected.

Finally, in Fig. 10, we combine the three above criteria (Sec. III A) to establish the minimum SNR required to perform BH spectroscopy. Solid (dashed) lines provide an estimate of the minimum SNR required to perform BH spectroscopy to a level of 10% (5%) accuracy. To compare the performance of a set of modes/tones, one needs to compare the solid (dashed) curves to solid (dashed) curves in this figure. Among the three criteria discussed in Sec. III A, which is the most stringent one depends on the value of  $q$ . The transition from one criteria to another corresponds to the derivative discontinuity seen in Fig. 10. For example, for the  $l = 2, m = 1, n = 0$  mode

the most stringent criterion at  $q \approx 1$  is detectability, whereas for  $q \gtrsim 3.7$  is resolvability, and for intermediate values of  $q$  the most stringent criterion is measurability.

To summarise, we find that for all values of  $q \gtrsim 1.2$ , the  $l = m = 3$  mode allows for the best spectroscopic analysis of BBH RD, with a minimum required SNR  $\approx 10$  or less when  $q \gtrsim 1.75$  for a 10% accuracy test using the combination  $\{f_{220}, Q_{220}, f_{\text{sub}}\}$ , and slightly larger for a 5% accuracy test. However, the minimum SNR required for BH spectroscopy with angular modes quickly deteriorates as  $q \rightarrow 1$  and becomes larger than that required with overtones in this limit. For nearly equal mass systems ( $q \lesssim 1.2$ ) overtones allow us to perform a spectroscopic analysis of the RD at a 10% accuracy level, but only for high SNR RD signals ( $\rho_{\text{RD}} \gtrsim 100$ ).

#### IV. DISCUSSION

In this work we have investigated the prospects for a spectroscopic analysis of the BBH RD using the angular modes as well as the overtones. Specifically, we have investigated the minimum SNR required to detect, resolve, and measure the QNM frequencies and quality factors of a single subdominant mode/tone for 3 specific choices: the  $l = m = 3$  and  $l = 2, m = 1$  subdominant angular modes (both with  $n = 0$ ), and the first overtone ( $l = m = 2, n = 1$ ) of the dominant angular mode.

We presented the fits for the amplitude ratio  $A_{R,lmn}$  and phase difference  $\Delta\phi_{lmn}$  of the subdominant mode using the NR RD waveforms corresponding to nonspinning BBHs with different mass ratios and for different choices of the RD starting time  $t_0$ . We found that choosing  $t_0 \gtrsim 10M$  is appropriate to reduce the impact of the systematic errors in the RD modelling (either due to a limited number of overtones or to possible nonlineari-

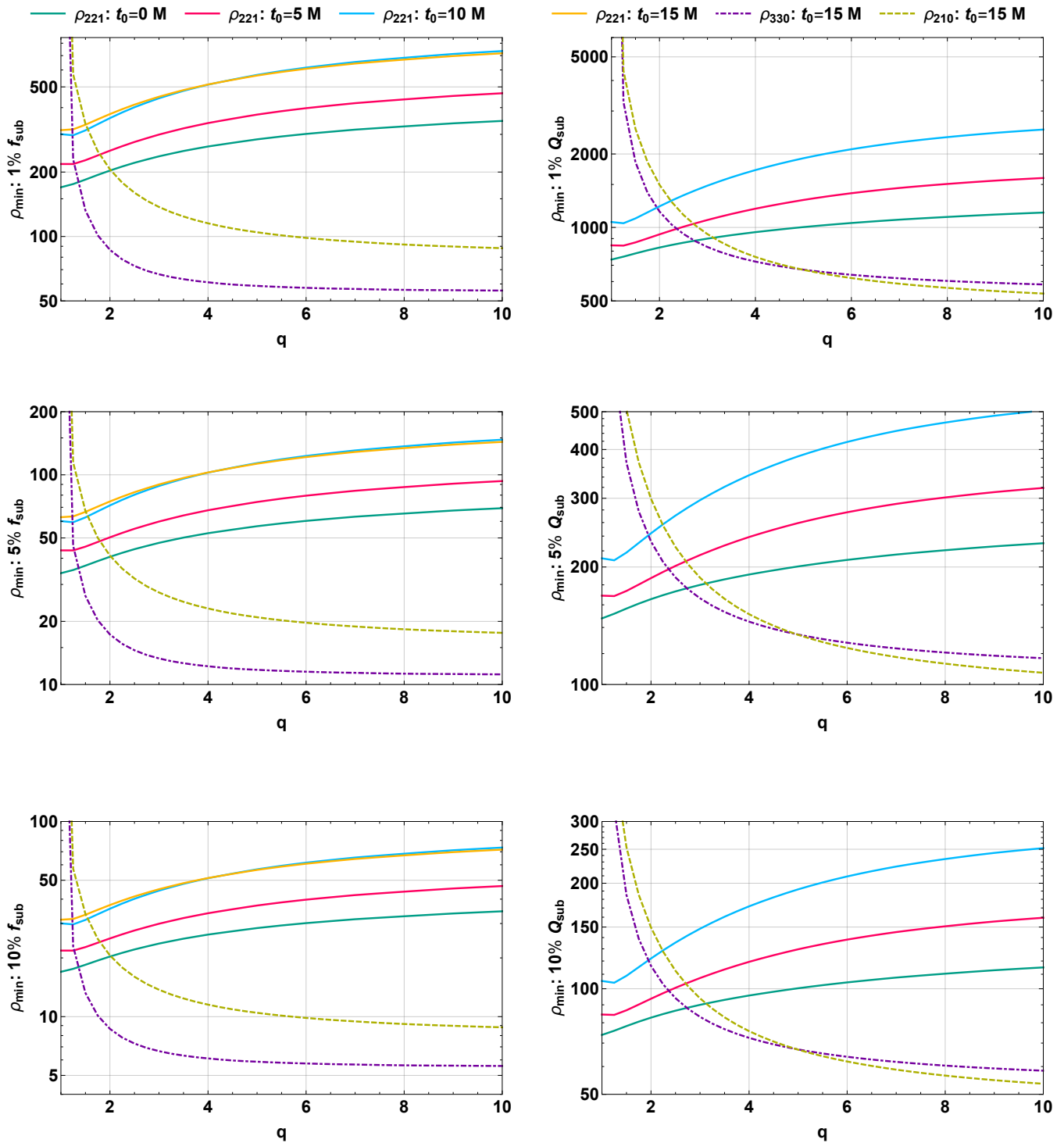


FIG. 9. Measurability of the subdominant mode. Top to bottom panels show the minimum RD SNR required to measure the QNM parameters with 10%, 5% and 1% precision, respectively. We use the same colour scheme as in Fig. 8. The left and right panels correspond to the recovery of the frequency and of the quality factor, respectively. Notice that for a given RD SNR, the subdominant mode frequency can be measured with much less uncertainty compared to the quality factor. The orange curve corresponding to  $l = m = 2, n = 1$  for  $t_0 = 15M$  does not appear in the right panels since it is above the scale of the plots.

ties). However, this choice of the starting time is imprac-

tical for a 2-tone RD model, as the overtones are quickly

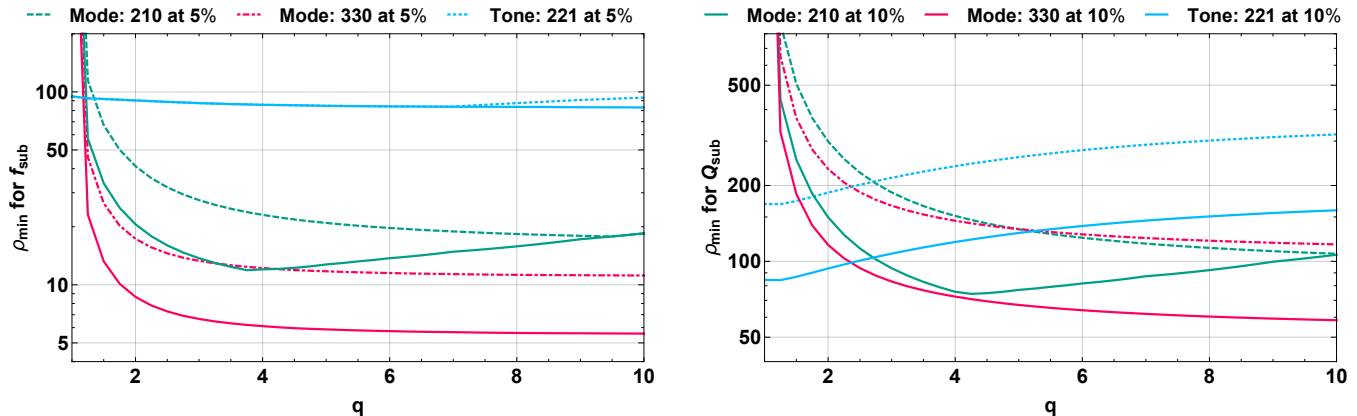


FIG. 10. Minimum SNR for BH spectroscopy with a secondary mode, combining the detectability, resolvability, and measurability criteria [Sec. III A]. Left panel refers to measuring and resolving the triad  $\{f_{220}, Q_{220}, f_{\text{sub}}\}$ , whereas the right panel refers to  $\{f_{220}, Q_{220}, Q_{\text{sub}}\}$ . The solid teal, pink and blue curves correspond to measurement with 10% precision of  $l = 2, m = 1, n = 0$ ,  $l = m = 3, n = 0$  and  $l = m = 2, n = 1$  subdominant mode frequency, respectively. The dot-dashed curves correspond to 5% precision. In this figure, for the overtones, we choose  $t_0 = 5M$ . For unequal mass-ratio BBHs,  $l = m = 3$  is the most promising subdominant mode to perform BH spectroscopy and for near-equal mass ratio systems, overtone performs better. This is a conservative estimate, since the choice of small values of  $t_0$  leads to higher SNR in the overtones.

damped. Therefore, for the cases of a subdominant overtone the analysis has been repeated for a starting time of  $\{0M, 5M, 10M, 15M\}$ , an appropriate value needs to be chosen using techniques such as that outlined in Ref. [10] (see also Appendix C). Our amplitude fits are consistent with the results in Ref. [34].

The amplitude ratio  $A_{R,lmn}$  has been fit to a fourth-degree polynomial while the phase difference  $\Delta\phi_{lmn}$  has been fit to a Lorentzian function. In the range  $q \in [1, 10]$ , the amplitude ratio  $A_{R,lmn}$  is approximately bounded within  $[a_{0,221}, 2.5 a_{0,221}]$ ,  $[0, a_{0,330}]$ ,  $[0, a_{0,210}]$ , where  $a_{0,lmn}$  (Tables II, III, and IV) is the asymptotic value at  $q \rightarrow \infty$ . Similarly, for the phase difference we obtain that  $|\Delta\phi_{lmn}(q=1) - b_{0,lmn}| \lesssim 0.9$  for the all three modes considered. The fact the amplitude ratio and phase difference do not change significantly across the entire range of  $q$  is an interesting empirical feature that could be used to tailor a more efficient mode stacking algorithm for RD tests of GR [34, 53]. We plan to explore this issue in a future work.

One of the main results of this work is the estimate of the minimum SNR required for BH spectroscopy with 2-modes/tones as a function of the mass ratio  $q$  of the progenitor BBH. For this, we introduced a *detectability* criterion – i.e., whether the error in the secondary mode amplitude ratio is smaller than 100% –, a *resolvability* criterion – i.e., whether the secondary QNM can be resolved from the dominant  $l = m = 2, n = 0$  mode –, and a *measurability* criterion – i.e., whether we can constrain at least three out of the four QNM parameters of a 2-mode/tone model within a threshold accuracy. By combining these three criteria we found that BH spectroscopy with angular modes (especially  $l = m = 3$ ) is the most promising channel for nonspinning BBHs with

$q \gtrsim 1.2$ . Our conclusions differ from those of [34] for unequal-mass binaries due to the different set of criteria we require to perform BH spectroscopy (see Sec. III A and footnote 7), but our results are otherwise in good agreement.

In particular, BH spectroscopy could be successfully performed with the  $l = m = 3$  mode and at 10% accuracy level if a signal with  $\rho_{\text{RD}} \approx 8$  and  $q \sim 2$  is observed. Analogously, the minimum RD SNR required for these tests is  $\rho_{\text{RD}} \approx 20$  if the  $l = 2, m = 1$  mode is used. On the other hand, the minimum required SNR for BH spectroscopy with angular modes diverges in the  $q \rightarrow 1$  limit since in this case  $l = m = 3$  and  $l = 2, m = 1$  modes are not excited. For nonspinning binaries with  $q \lesssim 1.2$ , BH spectroscopy with overtones becomes more convenient, although it requires very loud signals ( $\rho_{\text{RD}} \gtrsim 100$ ). Such loud signals might be detectable with third-generation ground-based detectors [39–41] (such as the Einstein Telescope and Cosmic Explorer) and with the future space mission LISA [42]. However, in these cases (especially for LISA) binaries with a broad distribution of mass ratios are expected so the actual benefit of overtone BH spectroscopy might be limited. We will explore this issue in a future work. Future instruments will also detect RD signals at higher SNR, thus increasing the chances to detect more than 2-modes/tones. Besides allowing for independent BH spectroscopy tests, including more overtones will reduce the model systematics and will allow to start the RD analysis earlier after the merger.

Another possible extension of our analysis is to include spinning binaries (along the lines of Refs. [32, 57]) and quantify the role of the binary component spins in the excitation and detectability of a secondary QNM (see

also Appendix D. This is particularly relevant in the light of GW190412, which has a nonnegligible effective spin parameter [37]. We note that, in our notation,  $q_{\text{GW190412}} \approx 3.8$ . Although the RD SNR for this system was low, our results show that a putative spectroscopic analysis of GW190412 would have been much more accurate if performed with angular modes rather than with overtones.

### ACKNOWLEDGEMENTS

We are indebted to Cecilia Chirenti and Iara Ota for useful suggestions on the draft and for comparing some of their results with ours, and to Emanuele Berti and Vitor Cardoso for useful correspondence on Refs. [54, 56]. We acknowledge financial support provided under the European Union’s H2020 ERC, Starting Grant agreement no. DarkGRA-757480, and under the MIUR PRIN and FARE programmes (GW-NEXT, CUP: B84I20000100001). We also acknowledge networking support by the COST Action CA16104 and support from the Amaldi Research Center funded by the MIUR program “Dipartimento di Eccellenza” (CUP: B81I18001170001).

### Appendix A: NR error estimate

Truncation errors in NR waveforms result from the discreteness of the grid that defines the domain of the simulations [44, 58]. This error gets reduced as one increases the sampling of the numerical domain, that is, by increasing the resolution of the numerical grid. To estimate the NR error of a single NR simulation we need to compare the numerical outputs that result from changing the grid resolution.

The SXS catalogue [43] provides the data for different resolution levels, labeled as *Lev-N*, where the resolution increases with the value of  $N$ . The NR error estimate for GW data is usually given in terms of phase and amplitude deviations that one finds between different resolution levels. This is computed after aligning all the waveforms at different resolution in time (or frequency) and phase for a given mass ratio, since in general there exists a time and phase offset between the levels. We choose to align each resolution level  $N$  such that  $\Phi^{(N)} = 0$  at  $t = \delta t_{lm}^p$ . In Fig. 11 we compute the accumulated (up to  $t = 60M$ ) deviation on the phase,  $\delta\phi$ , between two waveforms with resolution levels  $N$  and  $N - 1$  (*finest*) and between the highest and lowest resolution (*coarsest*) for each SXS NR simulation. We take the mean value as the averaged estimates of the NR error of our phase difference  $\Delta\phi_{lmn}$ . The mean value of  $|\delta\phi|$  is about 0.04 rad for both the coarsest and finest estimates, which can approximately account for the differences of  $\sim 0.1$  rad observed in Fig. 4 at  $t_0 = 10M$  and  $t_0 = 15M$ .

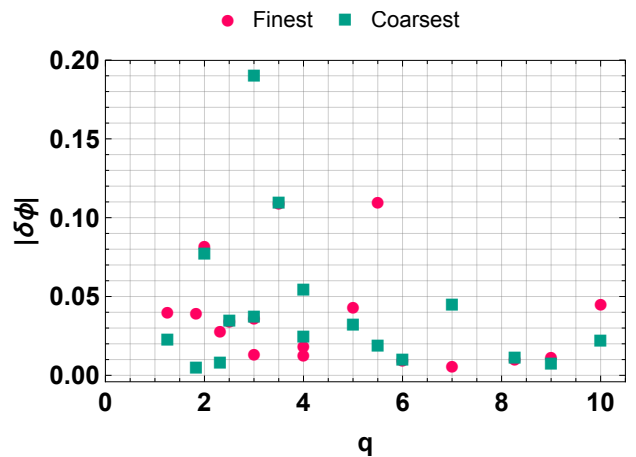


FIG. 11. Deviation accumulated on the phase  $|\delta\phi|$  between two waveforms with the two highest resolutions (red circles) and between two waveforms with the highest and lowest resolution (green circle) for each simulation of the catalogue [43]. The mean values are both approximately  $\langle|\delta\phi|\rangle = 0.04$  rad.

Another possibility is to estimate the NR error by means of the mismatch [Eq. (3)] between two waveforms with two different resolutions. This is the approach followed in this work for the RD signals. For each simulation listed in Table I, we compute Eq. (3) with  $t_i = 0$  and  $t_f = 60M$  between: (i) two waveforms with the highest ( $N$ ) and second highest ( $N - 1$ ) resolution (*finest*); and (ii) two waveforms with the highest ( $N$ ) and the lowest resolution (*coarsest*). This analysis is shown in Fig. 12. As expected, the mismatch is larger for the coarsest resolved waveforms than for finest ones. The median value of the mismatch is  $2 \times 10^{-4}$  and  $2 \times 10^{-5}$  between the coarsest and the finest case, respectively.

The red markers in Fig. 12 correspond to the horizontal grid lines in Fig. 1 and to the purple points of Fig. 5.

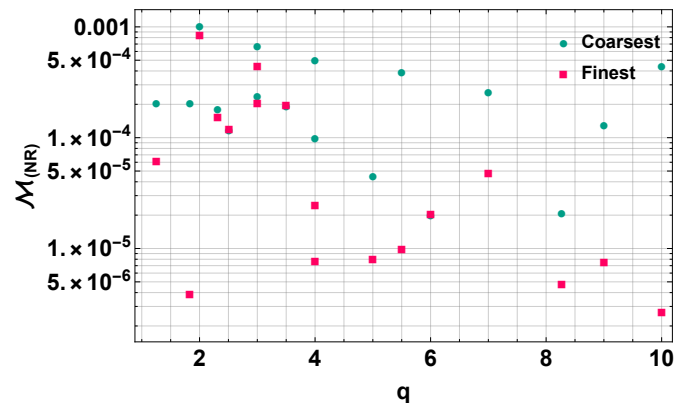


FIG. 12. Mismatch for each SXS NR simulation computed between the two better resolved waveforms (*finest*) and between the best and lowest resolved waveforms (*coarsest*). The median values for both set of points are  $2 \times 10^{-4}$  and  $2 \times 10^{-5}$  for the green and red dots, respectively.

## Appendix B: Spherical versus spheroidal harmonic decomposition and dependence on the inclination angle

The GW signal  $h(t)$  can be decomposed onto a set of basis functions. In the RD regime, a natural decomposition is using spheroidal harmonics. Indeed, each of the  $h_{lmn}(t)$  modes in Eq. (1) are weighted by the spin-two spheroidal harmonics  ${}_2\mathcal{Y}_{lmn}(\theta, \phi)$ . Due to the angular dependence of the spheroidal harmonics, the relative amplitude and phase of different modes depend on the inclination angle  $\theta$  of the source. However, the spacetime during the inspiral-merger phase is better described by a decomposition in spin-two spherical harmonics  ${}_2Y_{lm}(\theta, \phi)$ . Due to the dominant quadrupolar nature of GW emission during inspiral, a BBH merger signal is dominated by the  $l = 2, m = \pm 2$  modes.

The spherical and spheroidal basis functions are related by [59, 60]

$${}_2\mathcal{Y}_{lmn} = {}_2Y_{lm} + \sum_{l' \neq l} \mathfrak{H}_{ll'mn}(a, a^2) {}_2Y_{l'm} + \mathcal{O}(a^3), \quad (\text{B1})$$

where the specific form of  $\mathfrak{H}_{ll'mn}(a, a^2)$  can be found in Appendix A of [61]. In the above equation,  $Y_{lm}$  is the dominant term and  $\mathfrak{H}_{ll'mn} = 0$  for vanishing dimensionless spacetime angular momentum,  $a = 0$ . Thus,  ${}_2\mathcal{Y}_{lmn} = {}_2Y_{lm}$  for Schwarzschild BHs. In the case of a spinning remnant, the spheroidal harmonics differ from the spherical ones as given in Eq. (B1). However, since the overlap between spherical harmonics with different index  $l$  is small, in practice one can neglect the difference at least for moderately spinning remnants [62].

In general, one must account for the dependence on the inclination angle through these angular functions to understand qualitatively the contribution of each  $h_{lmn}$  to the final RD signal  $h(t)$ . The harmonics relevant to us are

$${}_2Y_{22} = \frac{1}{2} \sqrt{\frac{5}{\pi}} \cos^4\left(\frac{\theta}{2}\right) e^{i2\phi}, \quad (\text{B2})$$

$${}_2Y_{33} = -\sqrt{\frac{21}{2\pi}} \sin\left(\frac{\theta}{2}\right) \cos^5\left(\frac{\theta}{2}\right) e^{i3\phi}, \quad (\text{B3})$$

$${}_2Y_{21} = \frac{1}{4} \sqrt{\frac{5}{\pi}} \sin(\theta)(1 + \cos\theta) e^{i\phi}, \quad (\text{B4})$$

where  $\theta$  is the inclination angle and  $\phi$  is the orbital plane (or azimuthal) angle. We can obtain the  $m < 0$  modes by means of the parity transformation  ${}_2Y_{l-m}(\theta, \phi) = (-1)^{l+m} {}_2Y_{lm}^*(\pi - \theta, \pi + \phi)$ .

In Fig. 13 we show that for nearly face-on/face-off systems (i.e., with inclination angle  $\theta \approx 0, \pi$ ) the  $l = 3, m = \pm 3$  and  $l = 2, m = \pm 1$  harmonics are negligible and  $h(t)$  is accurately modelled with only the dominant  $l = 2, m = \pm 2$  mode for a face-on/face-off orientation, respectively. However, for an edge-on orientation ( $\theta = \pi/2$ ) both  ${}_2Y_{21}$  and  ${}_2Y_{33}$  are larger than  ${}_2Y_{22}$ . In general, unless the source is located at a particularly disfavoured

angle ( $\theta \sim 0, \pi$ ), the modes  $h_{21}$  and  $h_{33}$  can in principle (depending on their relative amplitude) carry an important fraction of the amplitude of  $h(t)$  for these inclination.

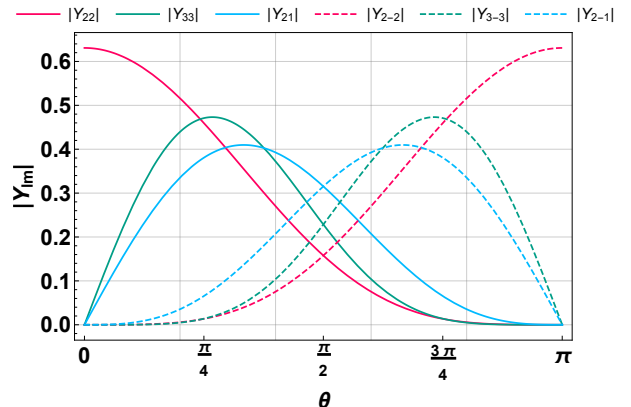


FIG. 13. Absolute value of the  ${}_2Y_{lm}$  harmonics for the  $(2, \pm 2)$ ,  $(2, \pm 1)$ ,  $(2, \pm 3)$  modes. At face-on orientation ( $\theta = 0$ ) the  ${}_2Y_{22}$  harmonic is the only nonzero contribution and dominates until  $\theta \approx \pi/4$ , where the contributions of the  ${}_2Y_{21}$  and  ${}_2Y_{33}$  harmonics become increasingly important. At edge-on orientation ( $\theta = \pi/2$ ) both  ${}_2Y_{21}$  and  ${}_2Y_{33}$  are larger than  ${}_2Y_{22}$ .

The dependence on the inclination angle  $\theta$  implies that the detectability, resolvability, and measurability criteria discussed in the main text will also depend on  $\theta$ . For simplicity, we have averaged out the GW strain on the  $\theta$  and  $\phi$  angles, which is equivalent to replacing  ${}_s\mathcal{Y}_{lm} \rightarrow 1/\sqrt{4\pi}$  in Eq. (1) [56].

## Appendix C: Bias on the subdominant QNM frequencies due to the choice of the RD starting time

To gain an insight into how the choice of starting time  $t_0$  leads to a bias in the estimated frequency for different angular modes, here we perform an analysis similar to the one we presented in Ref. [10] for the 2-tone model (c.f. Ref. [10] for details of the analysis).

In brief, for each mode we assume a modified RD waveform of form

$$h_{lmn}(t) = A_{lmn} e^{-i t \omega_{lmn} (1 + \frac{\alpha_{lmn}}{100})} e^{-(t-t_0)/\tau_{lmn}} \quad (\text{C1})$$

where  $\alpha_{lmn}$  is the relative (percent) deviation of the frequency from the GR-QNM spectrum,  $\omega_{lmn}$ . Since we studied the overtone case in Ref. [10], here we focus on  $n = 0$  only.

Since the overlap between different angular modes is not significant, we fit each angular mode to the corresponding angular decomposition of the NR simulation, and quantify the bias by evaluating the best fit values of  $|\alpha_{lmn}|$ . Ideally, should the above RD model match the NR-RD perfectly, we expect  $\alpha_{lmn} = 0$ , i.e., we expect a zero bias in the recovered QNM frequencies. A nonvanishing value of  $|\alpha_{lmn}|$  quantifies the systematic errors

introduced by the choice in the starting time. Possible nonlinearity and the inclusion of a finite number of tones (here we consider only the fundamental  $n = 0$  tone for each angular mode) may explain this bias. In Fig. 14, we show  $-\alpha_{lm0}$  as a function of  $t_0 - \delta t_{lm}^p$  for various BBH mass ratios. From Fig. 14, we see that the  $l = m = 2$  mode is most sensitive to the choice of the starting time in the sense that the blue curve always corresponds to the largest absolute value among the 3 curves for any given choice of the starting time. In other words, an earlier starting time leads to a larger amount of bias in the dominant mode compared to the sub-dominant mode. Therefore, if one desires that the systematic error due to the choice of starting time be limited to a certain percent, it is sufficient that this criterion be imposed on just  $l = m = 2$  mode.

#### **Appendix D: Minimum SNR for detectability, resolvability, and measurability as a function of amplitudes and phases**

In Fig. 15 we show the minimum SNR required to satisfy each of the criteria discussed in Sec. III A – namely detectability, resolvability, and measurability – as a function of the amplitude ratio  $A_R$  and of the phase difference  $\Delta\phi$  of a secondary mode/tone. This plot can be useful to immediately estimate the minimum SNR once  $A_R$  and  $\Delta\phi$  are known for a given system, also beyond those studied in this work (e.g., when including spinning progenitor binaries).



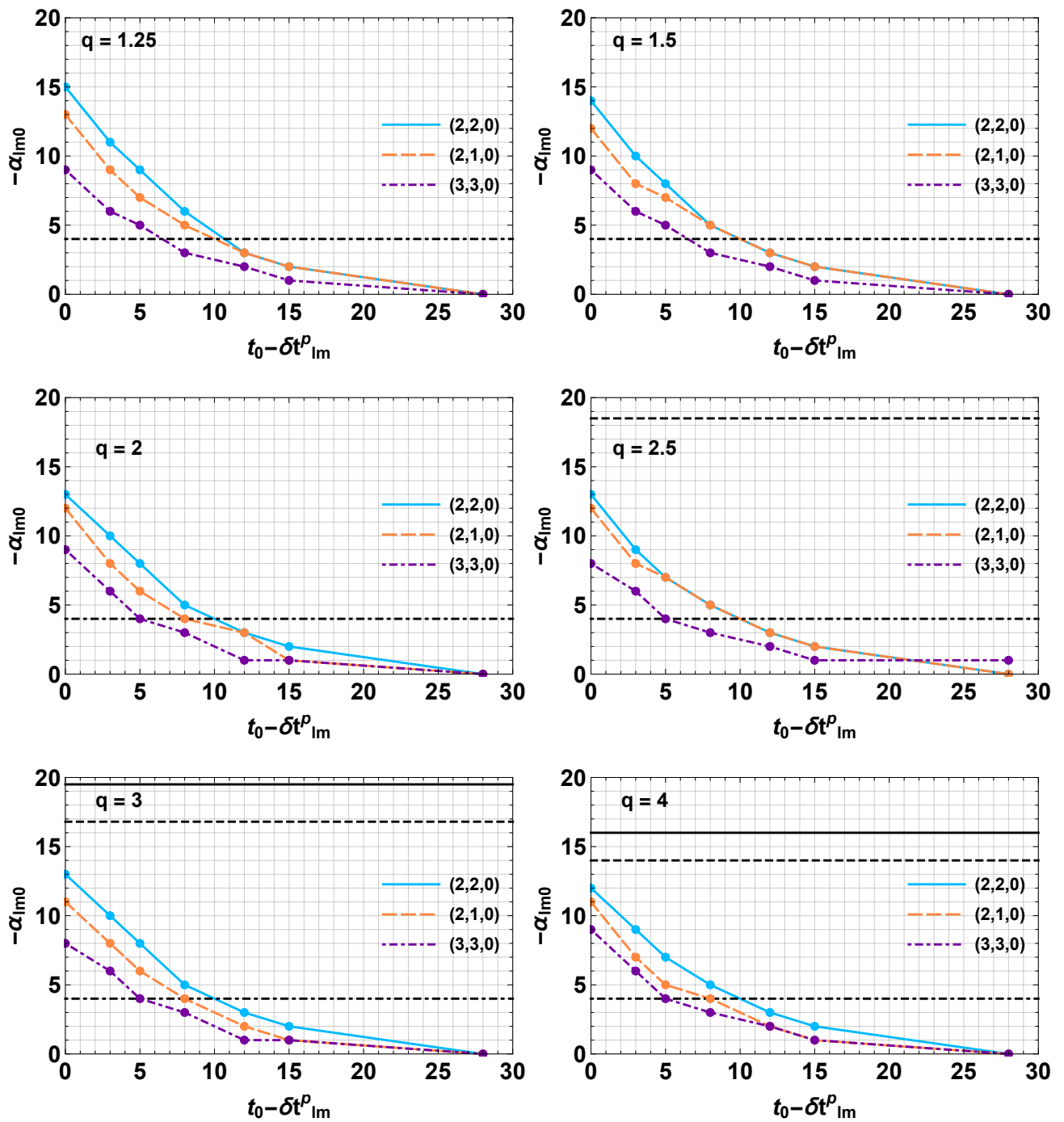


FIG. 14. Bias  $-\alpha_{lm0}$  on the recovery of the QNM frequency of different (fundamental) angular modes. The  $x$ -axis  $t_0 - \delta t_{lm}^p$  denotes the starting time of each mode with respect to the time at which the amplitude of the waveform peaks. The blue, the orange, and the purple curves correspond to the bias in  $l = m = 2$ ,  $l = m = 3$  and  $l = 2, m = 1, n = 0$  angular modes, respectively, neglecting overtones for each angular mode. The dot-dashed, dashed, and solid horizontal lines correspond to the Fisher matrix spread expected in the recover of  $f_{22}$ ,  $f_{33}$ , and  $f_{21m}$  respectively, assuming  $\rho_{RD} = 15$  (the horizontal lines scales as  $1/\rho_{RD}$ ). Notice that the dominant  $l = m = 2$  mode is the one with the largest bias for any mass ratio.

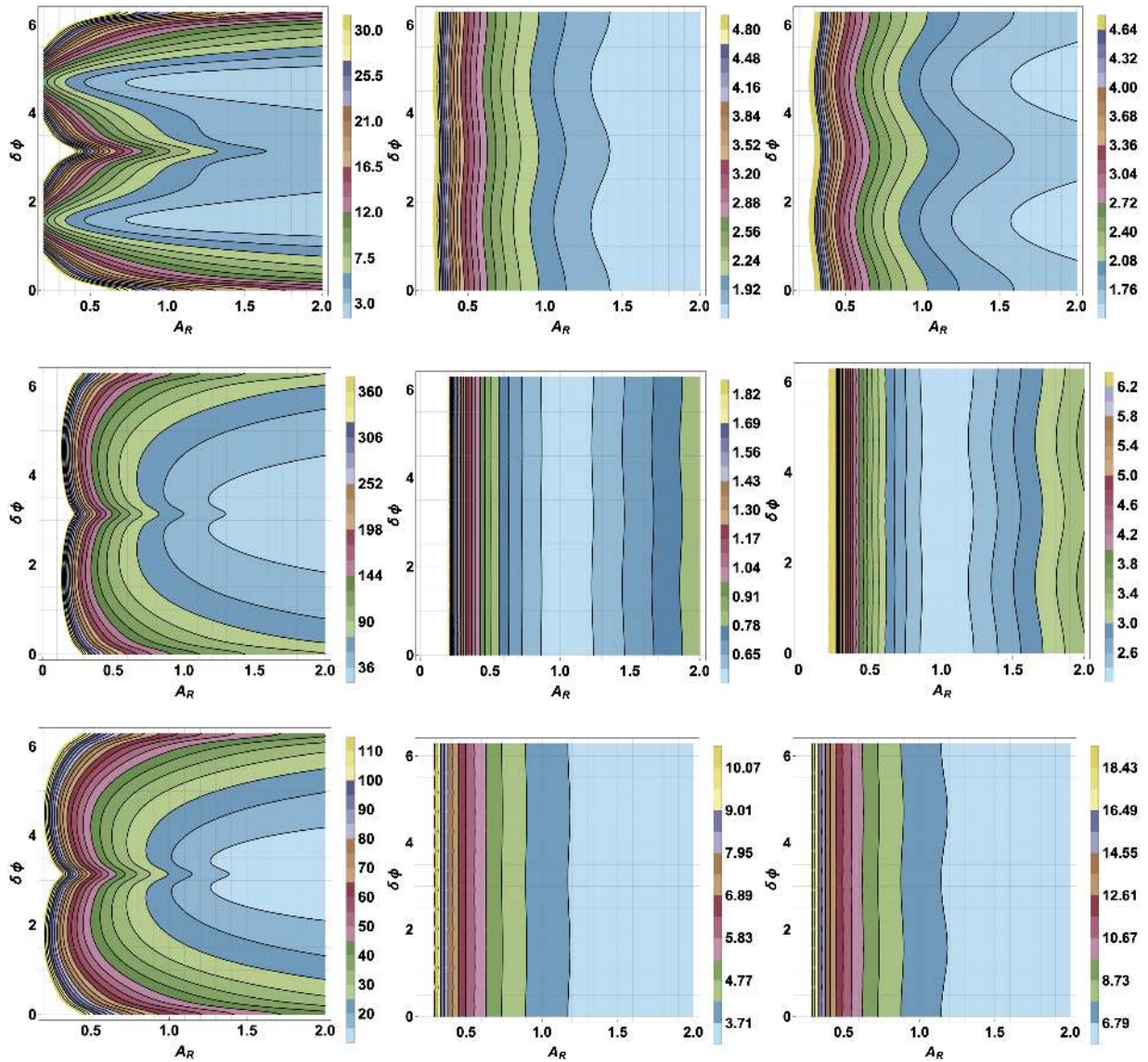


FIG. 15. Contour plot showing the minimum SNR required to satisfy each of the criteria discussed in Sec. III A (detectability, resolvability, and measurability) for a two-mode/one model as a function of the amplitude ratio  $A_R$  and of the phase difference  $\delta\phi$ . The top, middle, and bottom panels correspond to the detectability, the resolvability, and the measurability of the subdominant mode frequency to a 5% precision, respectively. The left, center, and right panels show the case of a secondary QNM with  $l = m = 2$ ,  $n = 1$  (left),  $l = m = 3$ ,  $n = 0$  (center), and  $l = 2$ ,  $m = 1$ ,  $n = 0$  (right), respectively.

- [1] C. V. Vishveshwara, *Nature* **227**, 936 (1970).
- [2] S. A. Teukolsky, *Astrophys. J.* **185**, 635 (1973).
- [3] W. H. Press and S. A. Teukolsky, *Astrophys. J.* **185**, 649 (1973).
- [4] S. A. Teukolsky and W. H. Press, *Astrophys. J.* **193**, 443 (1974).
- [5] S. Chandrasekhar and S. Detweiler, *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* **344**, 441 (1975), <https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1975.0197>.
- [6] E. W. Leaver, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **402**, 285 (1985).
- [7] K. D. Kokkotas and B. G. Schmidt, *Living Rev. Rel.* **2**, 2 (1999), [arXiv:gr-qc/9909058](https://arxiv.org/abs/gr-qc/9909058) [gr-qc].
- [8] E. Berti, V. Cardoso, and A. O. Starinets, *Class. Quant. Grav.* **26**, 163001 (2009), [arXiv:0905.2975](https://arxiv.org/abs/0905.2975) [gr-qc].
- [9] S. Bhagwat, M. Okounkova, S. W. Ballmer, D. A. Brown, M. Giesler, M. A. Scheel, and S. A. Teukolsky, *Phys. Rev. D* **97**, 104065 (2018), [arXiv:1711.00926](https://arxiv.org/abs/1711.00926) [gr-qc].
- [10] S. Bhagwat, X. J. Forteza, P. Pani, and V. Ferrari, *Phys. Rev. D* **101**, 044033 (2020), [arXiv:1910.08708](https://arxiv.org/abs/1910.08708) [gr-qc].
- [11] P. Figueras and J. Lucietti, *Class. Quant. Grav.* **27**, 095001 (2010), [arXiv:0906.5565](https://arxiv.org/abs/0906.5565) [hep-th].
- [12] B. Carter, *Phys. Rev. Lett.* **26**, 331 (1971).
- [13] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, 2011).
- [14] D. Robinson, in *Kerr Fest: Black Holes in Astrophysics, General Relativity and Quantum Gravity Christchurch, New Zealand, August 26-28, 2004* (2004).
- [15] P. O. Mazur, *J. Phys.* **A15**, 3173 (1982).
- [16] G. W. Gibbons, D. Ida, and T. Shiromizu, *Phys. Rev. Lett.* **89**, 041101 (2002), [arXiv:hep-th/0206049](https://arxiv.org/abs/hep-th/0206049) [hep-th].
- [17] W. Israel, *Phys. Rev.* **164**, 1776 (1967).
- [18] L. Barack *et al.*, *Class. Quant. Grav.* **36**, 143001 (2019), [arXiv:1806.05195](https://arxiv.org/abs/1806.05195) [gr-qc].
- [19] E. Berti, K. Yagi, H. Yang, and N. Yunes, *Gen. Rel. Grav.* **50**, 49 (2018), [arXiv:1801.03587](https://arxiv.org/abs/1801.03587) [gr-qc].
- [20] B. P. Abbott *et al.* (Virgo, LIGO Scientific), *Phys. Rev. Lett.* **116**, 221101 (2016), [arXiv:1602.03841](https://arxiv.org/abs/1602.03841) [gr-qc].
- [21] E. Berti, A. Sesana, E. Barausse, V. Cardoso, and K. Belczynski, *Phys. Rev. Lett.* **117**, 101102 (2016), [arXiv:1605.09286](https://arxiv.org/abs/1605.09286) [gr-qc].
- [22] S. Gossan, J. Veitch, and B. Sathyaprakash, *Phys. Rev. D* **85**, 124056 (2012), [arXiv:1111.5819](https://arxiv.org/abs/1111.5819) [gr-qc].
- [23] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Phys. Rev. Lett.* **116**, 131103 (2016), [arXiv:1602.03838](https://arxiv.org/abs/1602.03838) [gr-qc].
- [24] B. P. Abbott *et al.* (LIGO Scientific, Virgo), *Phys. Rev. Lett.* **116**, 221101 (2016), [Erratum: *Phys. Rev. Lett.* **121**, no.12, 129902 (2018)], [arXiv:1602.03841](https://arxiv.org/abs/1602.03841) [gr-qc].
- [25] G. Carullo, L. van der Schaaf, L. London, P. T. H. Pang, K. W. Tsang, O. A. Hannuksela, J. Meidam, M. Agathos, A. Samajdar, A. Ghosh, T. G. F. Li, W. Del Pozzo, and C. Van Den Broeck, *Phys. Rev. D* **98**, 104020 (2018), [arXiv:1805.04760](https://arxiv.org/abs/1805.04760) [gr-qc].
- [26] G. Carullo, W. Del Pozzo, and J. Veitch, *Phys. Rev. D* **99**, 123029 (2019), [arXiv:1902.07527](https://arxiv.org/abs/1902.07527) [gr-qc].
- [27] V. Baibhav and E. Berti, *Phys. Rev. D* **99**, 024005 (2019), [arXiv:1809.03500](https://arxiv.org/abs/1809.03500) [gr-qc].
- [28] S. Bhagwat, D. A. Brown, and S. W. Ballmer, *Phys. Rev. D* **94**, 084024 (2016), [Erratum: *Phys. Rev. D* **95**, no.6, 069906 (2017)], [arXiv:1607.07845](https://arxiv.org/abs/1607.07845) [gr-qc].
- [29] S. Bhagwat, M. Cabero, C. D. Capano, B. Krishnan, and D. A. Brown, *arXiv e-prints*, [arXiv:1910.13203](https://arxiv.org/abs/1910.13203) (2019), [arXiv:1910.13203](https://arxiv.org/abs/1910.13203) [gr-qc].
- [30] L. London, D. Shoemaker, and J. Healy, *Phys. Rev. D* **90**, 124032 (2014), [Erratum: *Phys. Rev. D* **94**, no.6, 069902 (2016)], [arXiv:1404.3197](https://arxiv.org/abs/1404.3197) [gr-qc].
- [31] A. Taracchini *et al.*, *Phys. Rev. D* **89**, 061502 (2014), [arXiv:1311.2544](https://arxiv.org/abs/1311.2544) [gr-qc].
- [32] V. Baibhav, E. Berti, V. Cardoso, and G. Khanna, *Phys. Rev. D* **97**, 044048 (2018), [arXiv:1710.02156](https://arxiv.org/abs/1710.02156) [gr-qc].
- [33] M. Giesler, M. Isi, M. A. Scheel, and S. Teukolsky, *Phys. Rev. X* **9**, 041060 (2019), [arXiv:1903.08284](https://arxiv.org/abs/1903.08284) [gr-qc].
- [34] I. Ota and C. Chirenti, (2019), [arXiv:1911.00440](https://arxiv.org/abs/1911.00440) [gr-qc].
- [35] M. Isi, M. Giesler, W. M. Farr, M. A. Scheel, and S. A. Teukolsky, *Phys. Rev. Lett.* **123**, 111102 (2019), [arXiv:1905.00869](https://arxiv.org/abs/1905.00869) [gr-qc].
- [36] B. P. Abbott *et al.* (LIGO Scientific, Virgo), (2018), [arXiv:1811.12907](https://arxiv.org/abs/1811.12907) [astro-ph.HE].
- [37] (2020), [arXiv:2004.08342](https://arxiv.org/abs/2004.08342) [astro-ph.HE].
- [38] M. Cabero, J. Westerweck, C. D. Capano, S. Kumar, A. B. Nielsen, and B. Krishnan, *Phys. Rev. D* **101**, 064044 (2020), [arXiv:1911.01361](https://arxiv.org/abs/1911.01361) [gr-qc].
- [39] S. Hild *et al.*, *Class. Quant. Grav.* **28**, 094013 (2011), [arXiv:1012.0908](https://arxiv.org/abs/1012.0908) [gr-qc].
- [40] B. P. Abbott *et al.* (LIGO Scientific), *Class. Quant. Grav.* **34**, 044001 (2017), [arXiv:1607.08697](https://arxiv.org/abs/1607.08697) [astro-ph.IM].
- [41] R. Essick, S. Vitale, and M. Evans, *Phys. Rev. D* **96**, 084004 (2017), [arXiv:1708.06843](https://arxiv.org/abs/1708.06843) [gr-qc].
- [42] H. Audley, S. Babak, J. Baker, E. Barausse, P. Bender, E. Berti, P. Binetruy, M. Born, D. Bortoluzzi, J. Camp, C. Caprini, V. Cardoso, M. Colpi, J. Conklin, N. Cornish, C. Cutler, *et al.*, *ArXiv e-prints* (2017), [arXiv:1702.00786](https://arxiv.org/abs/1702.00786) [astro-ph.IM].
- [43] The SXS Collaboration, “SXS Gravitational Waveform Database,” (2016).
- [44] M. Boyle *et al.*, *Class. Quant. Grav.* **36**, 195006 (2019), [arXiv:1904.04831](https://arxiv.org/abs/1904.04831) [gr-qc].
- [45] M. Davis, R. Ruffini, W. H. Press, and R. H. Price, *Phys. Rev. Lett.* **27**, 1466 (1971).
- [46] A. Buonanno, L. E. Kidder, and L. Lehner, *Phys. Rev. D* **77**, 026004 (2008), [arXiv:0709.3839](https://arxiv.org/abs/0709.3839) [astro-ph].
- [47] F. Hofmann, E. Barausse, and L. Rezzolla, *Astrophys. J.* **825**, L19 (2016), [arXiv:1605.01938](https://arxiv.org/abs/1605.01938) [gr-qc].
- [48] J. Healy, C. O. Lousto, and Y. Zlochower, *Phys. Rev. D* **90**, 104004 (2014), [arXiv:1406.7295](https://arxiv.org/abs/1406.7295) [gr-qc].
- [49] S. A. Hughes and R. D. Blandford, *Astrophys. J.* **585**, L101 (2003), [arXiv:astro-ph/0208484](https://arxiv.org/abs/astro-ph/0208484) [astro-ph].
- [50] L. London, S. Khan, E. Fauchon-Jones, C. García, M. Hannam, S. Husa, X. Jiménez-Forteza, C. Kalaghatgi, F. Ohme, and F. Pannarale, *Phys. Rev. Lett.* **120**, 161102 (2018), [arXiv:1708.00404](https://arxiv.org/abs/1708.00404) [gr-qc].
- [51] M. Okounkova, (2020), [arXiv:2004.00671](https://arxiv.org/abs/2004.00671) [gr-qc].
- [52] A. Gupta, B. Krishnan, A. B. Nielsen, and E. Schnetter, *Phys. Rev. D* **97**, 084028 (2018).
- [53] H. Yang, K. Yagi, J. Blackman, L. Lehner, V. Paschalidis, F. Pretorius, and N. Yunes, *Phys. Rev. Lett.* **118**, 161101 (2017), [arXiv:1701.05808](https://arxiv.org/abs/1701.05808) [gr-qc].

- [54] E. Berti, J. Cardoso, V. Cardoso, and M. Cavaglia, *Phys. Rev. D* **76**, 104044 (2007), [arXiv:0707.1202 \[gr-qc\]](#).
- [55] E. Barausse, A. Buonanno, S. A. Hughes, G. Khanna, S. O'Sullivan, and Y. Pan, *Phys. Rev. D* **85**, 024046 (2012), [arXiv:1110.3081 \[gr-qc\]](#).
- [56] E. Berti, V. Cardoso, and C. M. Will, *Phys. Rev. D* **73**, 064030 (2006), [arXiv:gr-qc/0512160 \[gr-qc\]](#).
- [57] V. Baibhav and E. Berti, *Phys. Rev. D* **99**, 024005 (2019), [arXiv:1809.03500 \[gr-qc\]](#).
- [58] I. Hinder *et al.*, *Class. Quant. Grav.* **31**, 025012 (2014), [arXiv:1307.5307 \[gr-qc\]](#).
- [59] W. H. Press and S. A. Teukolsky, *Astrophys. J.* **185**, 649 (1973).
- [60] L. London and E. Fauchon-Jones, *Class. Quant. Grav.* **36**, 235015 (2019), [arXiv:1810.03550 \[gr-qc\]](#).
- [61] E. Berti and A. Klein, *Phys. Rev. D* **90**, 064012 (2014), [arXiv:1408.1860 \[gr-qc\]](#).
- [62] E. Berti, V. Cardoso, and M. Casals, *Phys. Rev. D* **73**, 024013 (2006), [arXiv:gr-qc/0511111 \[gr-qc\]](#).