Testing the instability of overtone models.

Yi Qiu[1], Xisco Jiménez Forteza[2,3], Pierre Mourier[2,3]
[1] *School of Astronomy and Space Science, Nanjing University, Nanjing 210023, P. R. China*
[2] *Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Callinstraße 38, 30167 Hannover, Germany and*
[3] *Leibniz Universität Hannover, 30167 Hannover, Germany*

The ringdown phase gravitational waves are extremely useful for the testings of general relativity (GR). While the distinguishable ringdown signal are required to constrain modified theories of gravity, the particular considerations on the instability inside overtone models are often absent. We investigate the uncertainties of overtone models in both the QNM values' deviation to the Kerr spectrum and alternative forms of damped sinusoids. Using our grid methods to replace part of the fitting, the results can present in an unprecedented way while providing extra information on the recovery consistency of final mass and spin of the remnant black hole. Throughout this work, a GW150914-like waveform data BBH:0305 from the SXS catalogue is mostly used. We found very high instability in the highest overtone of models while the fundamental tone can be better constraint with more overtones included. Our results of the 0, 1 tone perturbed model with the fractional deviation factors $\alpha_0 = -0.04$, $\beta_0 = 0.08$, $\alpha_1 = -0.24$, $\beta_1 = 0.48$ also give a new reference for the realistic gravitational waves testing of GR by providing the intrinsic uncertainty ranges of the overtone models in 2 tone black hole spectroscopic analysis.

I. INTRODUCTION

The number of gravitational wave (GW) observations is increasing along with the coming of future GW detectors. Up to date, the LIGO-Virgo collaboration has reported a total of 48 binary black hole merger candidates [1, 2]. Those observations are providing unprecedented constraints on general relativity in its strong regime, with the post-merger phase in particular providing a promising channel for such studies.

Each binary black hole coalescence event can be generally decomposed into three phases of evolution: inspiral, merger and ringdown (RD). Usually, the dynamics of these phases are theoretically described as the following: inspiral regime is depicted by post-Newtonian and effective-one-body theories; the merger phase, at when time the two objects get closer to each other and merge, can only be described by the strong-field numerical relativity; while the final stage is predicted by linear perturbation theory applying on the remnant black hole.

Particularly, black hole RD provides one of the most excellent observational prospects on the nature of gravity as it encodes the information about the structure of black hole.

In the RD regime, Teukolsky equation [3] masters the Kerr space-time, and thus, also tells us about how gravitational waves behave. The strain $h(t, \theta, \phi)$ of the RD waveform are composed of damped sinusoids:

$$h(t,\theta,\phi) = \sum_{l,m,n} \mathcal{A}_{lmn} e^{-\iota\omega_{lmn}(t-t_0)}\, {}_{-2}\mathcal{Y}_{lm}(\theta,\phi)\,, t \geq t_0\,. \tag{1}$$

Here, $l = 2, 3, \ldots$ and $m = -l, -l+1, \ldots, l-1, l$ denote the two angular indices of the spheroidal decomposition, while $n = 0, 1, 2, \ldots$ labels the tone index; ${}_{-2}\mathcal{Y}_{lm}(\theta, \phi)$ are the spin-weighted spheroidal harmonics of spin weight $s = -2$, as functions of the polar angle θ and azimuthal angle ϕ; $\mathcal{A}_{lmn} = A_{lmn}\, e^{\iota\varphi_{lmn}}$ is the tone complex amplitude; and t_0 is some undefined time beyond which linear perturbation theory is expected to accurately describe the RD regime [4**?** , 5]. In particular, and for non-charged black holes, the $\omega_{lmn} = w_{lmn} - \iota/\tau_{lmn}$ defines an infinite set of complex frequencies solely determined by the final black hole's mass M_f and spin a_f, where the values of ω_{lmn} correspond to poles of the Green function to the inhomogeneous Teukolsky equation [6**?** , 7]. Here $\mathrm{Re}[\omega_{lmn}] = w_{lmn}$ and $-\mathrm{Im}[\omega_{lmn}] = 1/\tau_{lmn}$ take the role of the oscillation frequency and the damping rate (inverse of the damping time) respectively. As a rule of thumb, if one considers fixed the value of the (l, m) indices, the mass M_f and for moderate spins a_f, the values of the damping times τ_{lmn} decrease as the tone index n increases. This sets the $n = 0$ (fundamental) tone as the dominant tone while the $n \geq 1$ (over)tones rank down successively as n goes up.

The black hole no-hair theorem in general relativity (GR) implies that the final state of an uncharged black hole merger, and the associated quasi-normal mode (QNM) spectrum, are fully and uniquely determined by the values of the final mass and spin. This has led to two main avenues to test such theorems [**?**]. The first one consists in performing an inspiral–merger–ringdown (IMR) consistency test, which relies on independently estimating the final black hole mass and spin from both the inspiral–merger and the ringdown phases [**?**]. The second approach is to perform black hole spectroscopy, which typically aims at the independent estimate of the parameters of the fundamental tone of the dominant angular mode, $(l = 2, m = 2, n = 0)$, plus either i) the first corresponding overtone, $(l = 2, m = 2, n = 1)$, or ii) another angular fundamental mode, either the $(l = m = 3, n = 0)$ or the $(l = 2,\ m = 1,\ n = 0)$ mode [8] (in order of importance). So far and for unequal-mass-ratio binaries, the higher angular mode remains the most promising approach to test the implications of the

black hole no-hair theorem [?]. A successful independent evidence of the $(l = m = 2, n = 0)$ and the $(l = m = 3, n = 0)$ modes in the ringdown phase of a GW event (in this case GW190521) has been recently provided in [8]. On the other hand, channel i) becomes a promising possibility when dealing with near equal-mass-ratio nonspinning binaries. For such events, the higher harmonic modes are only weakly excited, while the overtones would still represent a valid channel in the ringdown regime. A first attempt to observe overtones in GW observational data has been performed in [? ?] on GW150914. However, the full spectroscopic analysis performed by [8] on GW190521 could not find evidence of tones other than the fundamental ones.

Current models of the $(22n)$ ringdown modes rely of fits to numerical relativity (NR) waveforms [4? ? ?], which are shown to be consistent with current GW observations. In particular, using NR waveforms [9] has the following advantages: i). the underlying theory which it based is well-known; ii). the mass and the spin of the final black hole are accurately estimated, hence accurately determining the QNM spectrum; and iii). numerical errors in the simulated waveforms are also given, and are typically smaller than current GW detectors noise. In such studies, considering the $(l = m = 2)$ mode of the strain[1], $h_{22}(t)$, from a given time t_0 onwards, one fits for the successive complex amplitudes $\mathcal{A}_{22n}$ of the $(22n)$ QNM tones for a running index $n \in \{0, \ldots, N_{\rm max}\}$, with various choices for the total number $N_{\rm max}$ of overtones to be included in the model. The $N_{\rm max} = 7$ model has been shown to provide the best estimates of the true final parameters (mass and spin) [4?], and the models beyond $N_{\rm max} = 7$ have also been studied recently [10].

Particularly, current black hole spectroscopic analysis, whether done with real data or numerical data, seem to pay their attentions only to GR basis. However, the instability of the overtone model itself which is theory-dependent might be neglected. The inadequate consideration on it and the corresponding careful treatments to the data analysis techniques can result in tone missing due to biased measurements and potentially erroneous conclusions. For example, in principle, the numerical relativity waveform should naturally have the best fit with GR, but if the instability on the overtone models are high enough, the best candidates might turn out to be some theories of modified gravity. Thus, we must understand to what extent can we constraint the tones, discern between different gravity theories, and trust our results in spectroscopy. In previous work, the instability of $N_{\rm max} = 1$ tone model have been test by the simulation of injections. In this work, since we are focusing on the numerical relativity data, we can not only extend the analysis to higher overtones to see a clearer trend, but also be capable to examine the mass/spin consistency aspect of the models (because the final mass/spin values are also given in very high accuracy).

We discuss the basis of the RD overtone models and three of the most widely used quantities associated with the RD data analysis in Sec. III below and show how they measure the model instabilities. In Sec. IV we introduce the definitions of two classes of modified overtone model candidates in which the instability could be test, including the tone perturbed models in Sec. IV A and other modified models in Sec. IV B. In Sec. V, we provide a comprehensive discussion on the most important computation method used in this work, i.e., the grid method, and some possible issues related to it. In Sec. VI, we turn to the results presenting, in which the deviation to Kerr spectrum in the highest tones or a 0,1 tones perturbed model are given in Sec. VI B and Sec. VI D. Finally, in Sec. VII we offer an summary of the main conclusions in the paper as well as alternative interpretations, and discuss the prospects of black hole spectroscopy.

II. THE RINGDOWN ANALYSIS

A. overtone models

The QNM values in overtone models with $N_{\rm max}$ up to 7 are computed through the open source qnm package [11].

III. MEASURING THE INSTABILITY

There are three quantities that we believe could be illustrative to the measurement of instability in overtone models: χ^2, $\mathcal{M}$ and ϵ. In the following subsections, we will give a detailed discussion on the way in which they represent instability differently and how we could consider them jointly.

A. The χ^2 analysis

In gravitational wave data analysis, a parameter space will be initially built on one specific template, then comes the matched-filtering process which works on all the combinations of test values (also known as the injections). Notice that once the values of the mass and spin are fixed, the RD ansatz (1) is linear in the complex amplitudes $\mathcal{A}_{lmn}$. Therefore, one may use a linear least-squares algorithm to obtain the fit results [4? ?]. That is, for a given value of the (M_f, a_f) pair, the complex amplitudes $\mathcal{A}_{lmn}$ are obtained by minimizing the χ^2:

$$\chi^2 = \sum_k \left| \bar{h}_{22}\left(\vec{\lambda}\right)(t_k) - h_{22}(t_k) \right|^2 , \qquad (2)$$

[1] It is worth mentioning here that in NR codes the strain $h(t, \theta, \phi)$ is decomposed in terms of the spin-weighted spherical harmonics basis instead of the spheroidal harmonics $\mathcal{Y}_{lm}$ used to define QNMs, since it is a better adapted basis to the inspiral-merger regime. This adds mode-mixing artifacts principally at modes other than the (22) mode [? ?].

where the subscript k labels the values of the time axis of the NR waveform, $t_k \in [t_0, t_f]$ for a certain fit starting time t_0, and $\bar{h}_{22}(\vec{\lambda})$ denotes the model $(2,2)$-mode strain for the set of parameters $\vec{\lambda}$. By default in the following, the starting time is set to $t_0 = 0$, which corresponds to the peak of the (22) mode of the strain $h_{22}(t)$.

B. The mismatch analysis

There is another way that closely related to the above χ^2 analysis which is also recurrently used in gravitational wave astronomy: to assess the fit goodness we might also use the mismatch $\mathcal{M}$ as in [12–15], which is defined as:

$$\mathcal{M} = 1 - \frac{\langle h_{\rm NR}|h_x\rangle}{\sqrt{\langle h_{\rm NR}|h_{\rm NR}\rangle\langle h_x|h_x\rangle}} \tag{3}$$

with:

$$\langle f|g\rangle = \int_{t_0}^{t_f} f(t)g(t)\,{\rm d}t\,. \tag{4}$$

The mismatch $\mathcal{M}$ analysis can usually be done in the fitting procedure of time-domain data. Thus, in numerical relativity gravitational wave analysis, a better (lower) $\mathcal{M}$ represents the more preferable model.

C. The epsilon analysis

The associated value of the minimum $\mathcal{M}$ for each RD $N_{\rm max}$ model is sufficient to assess the fit accuracy but insufficient to determine whether the fitting parameters are physically reliable. A decreasing value of the mismatch $\mathcal{M}$ between different models is particularly sensitive to overfitting, especially if it is applied to nested models such as the RD models we have considered in this work (the RD model with $N_{\rm max}-1$ overtones corresponds to the subclass of the RD model with $N_{\rm max}$ overtones with $\mathcal{A}_{N_{\rm max}}$ set to 0). To overcome this issue we use the mass and spin bias ϵ defined in Eq. (4) of [4],

$$\epsilon = \sqrt{\left(\frac{\delta M_f}{M}\right)^2 + \delta a_f^2}\,, \tag{5}$$

where $\delta M_f = M_f^{\rm fit} - M_f^{\rm true}$ and $\delta a_f = a_f^{\rm fit} - a_f^{\rm true}$. Thus, ϵ measures the combined deviation of the final mass M_f and the final spin a_f with respect to the true parameters $M_f^{\rm true}$ and $a_f^{\rm true}$ of the NR simulation, that are estimated from the mass and spin quasi-local definitions [9], *i.e.*, following Eqs. (**??**) and (**??**).

D. Summary of the analyses

In the above subsections, we demonstrate that the $\mathcal{M}$ measures the fitting goodness in the reconstructions of NR waveform, while the ϵ measures the recovered final mass and spin consistency. Thus, in order to consider them both as key quantities in the instability measurements as well as giving out confidence level of the deviation to the Kerr spectrum, we need to introduce the grid method in our computation as we will discuss in Sec. V.

IV. MODIFIED OVERTONE MODELS

In order to investigate the instability of overtone models, we implement the analysis that we introduced in Sec. III on several modified overtone models.

A. Tone perturbed models

In some modified gravity theories, the QNM values are also able to be obtained [16]. However, they are different from the GR solutions to some extents. Thus, it is natural that we firstly came up with the overtone models, in which we allow for deviations on the complex frequencies from the QNM values. The expression may be replaced by the following:

$$\begin{aligned} {}_sX = \sum_{l\geq s,\,m,\,n} \mathcal{A}_{lmn}\,\exp\left[-\iota\omega_{lmn}(1+\alpha_{lmn})(t-t_r)\right]\times \\ \times\exp\left[-\frac{t-t_r}{\tau_{lmn}(1+\beta_{lmn})}\right]\,{}_sY_{lm}\,, \end{aligned} \tag{6}$$

where α_{lmn} and β_{lmn} are two sets of perturbation parameters. The values of them will measure the deviations to the QNM spectrum (as predicted by perturbation theory within GR), while GR spectrum is recovered for $\alpha_{lmn} = \beta_{lmn} = 0$. To perform black hole spectroscopy in one specific gravity theory, one shall require that the posterior distributions of α_{lmn} and β_{lmn} are consistent with their theoretically predicted values. Or at least, the frequency values can be resolved to a given $n\sigma$ credible value [**?**], i.e. the overtone model should be stable enough.

On the other hand, in a high-overtone model, it is also very difficult to model the deviations on every single tone, simply because the parameter space will become too large and over-fittings very possibly happen. Thus, we will firstly consider the overtone model with just 1 tone perturbed.

B. Other modified overtone models

Apart from the single tone perturbed overtone models, we also introduce five other modified overtone models: four of which replace one or two specific tones (the highest in the model) by new expressions, namely, the power-law model (**PLM**), power-lawcosine mode (**PLCM**), exponential model (**EM**), power-law* model (**PLM***),

and the fifth one undergoes a universal time-coordinate-transformation (**TCTM**). All of these new models are believed to be capable of depicting the trends how the waveform amplitude vary with time as the original model does (i.e. exponentially decaying ringdown wave). Thus, they are used both as comparisons to the former studies and to further investigate the instability in regard to the alternative forms of damped sinusoids in overtone models.

The models read as the followings:

- **PLM** $(N_{\rm max})$: **OM** $(N_{\rm max}-1)+x_p t^{-\gamma}$
- **PLCM** $(N_{\rm max})$: **OM** $(N_{\rm max}-2)+x_p t^{-\gamma}\cos(\omega t+\phi)$
- **EM** $(N_{\rm max})$: **OM** $(N_{\rm max}-1)+x_p \exp\left(-\frac{t}{\tau}\right)$
- **PLM*** $(N_{\rm max})$: **OM** $(N_{\rm max}-1)\times(1+x_p t^{-\gamma})$
- **TCTM** $(N_{\rm max})$: **OM** $(t)\rightarrow$ **OM** $\left(Ae^{-t/\tau}\right)$

Note that the first four modifications present here are constructed with adding one specific non-linear term to the basic overtone model with a less number of overtones (i.e. $N_{\rm max}-1$ or $N_{\rm max}-2$). This setting ensures that all these modifications will depend on $2\,(N_{\rm max}+1)$ real-valued free parameters, which is, the same number of parameters as for the reference overtone models (**OM**). Accordingly, they are the $N_{\rm max}$ or $N_{\rm max}-1$ amplitudes A_{22n} along with the same number of phases ϕ_{22n}, plus the additional amplitude on the highest tone x_p, and either an exponent γ or an extra damping time τ (except for the specific case of **PLCM**, in which presents one another frequency ω and an extra phase ϕ). Particularly, for the **TCTM** cases, the extra term is not adding to the $N_{\rm max}-1$ basic tones, but instead, applying to all the $N_{\rm max}-1$ tones by replacing the ordinary time t present everywhere in the expression by a new coordinate of time with a normalization factor A and a damping factor τ, which represent a transformation systematically. Anyway, **TCTM** also depends on $2\,(N_{\rm max}+1)$ real-valued free parameters as the same to the other ones.

V. THE GRID METHOD

A. Algorithms set up

In the normal fitting procedure, the constructed ansatz will be given guess values of free parameters, and then fit to a series of data points (regarding the variable of function, which is time t in our case) over and over again to ultimately find the best-fit parameters.

Apart from the above, another scenario is to separate the fitting by replacing part of it with the grid method. For example, in our case, the fitting process is iterated over a range of (M_f, a_f) values to find the optimal one. To this aim, we build a two-dimensional adaptive grid on the remnant black hole's final mass M_f and final spin a_f as in [10, 17][2] Every point of the grid is then treated as a linear least-squares minimization problem on the parameters $\vec{\lambda}=\{A_{22n},\varphi_{22n}\}$ as above [17**?**].

Instead of the χ^2, we compute the mismatch for each best-fit RD model[3], Finally, the best-fit mass and spin values $(M_f=M_f^{\rm fit}, a_f=a_f^{\rm fit})$ for the given waveform and the given number of overtones $N_{\rm max}$ of the RD model are selected as the grid point where $\mathcal{M}$ from Eq. (3) is minimal.

The associated value of the minimum $\mathcal{M}$ for each RD $N_{\rm max}$ model is sufficient to assess the fit accuracy but insufficient to determine whether the fitting parameters are physically reliable. A decreasing value of the mismatch $\mathcal{M}$ between different models is particularly sensitive to over-fitting, especially if it is applied to nested models such as the RD models we have considered in this work (the RD model with $N_{\rm max}-1$ overtones corresponds to the subclass of the RD model with $N_{\rm max}$ overtones with $\mathcal{A}_{N_{\rm max}}$ set to 0). To overcome this issue we use the mass and spin bias factor ϵ as defined in Sec. III C to measure the combined deviation of the final mass M_f and the final spin a_f with respect to the true parameters $M_f^{\rm true}$ and $a_f^{\rm true}$ of the NR simulation, that are estimated from the mass and spin quasi-local definitions [9], *i.e.*, following Eqs. (**??**) and (**??**).

To sum up, there are two advantages of conducting the grid method: i). If the key parameters (which we believe having more importance) are well constraint to known ranges, separate them from the parameter space can help us better control the iterations in fitting as well as obtaining clearer distribution properties of them. ii). The over-fitting problem on the final mass M_f and the final spin a_f can be bypassed, otherwise the resulting QNM values on all the tones will be unreliable as they are all fully determined by (M_f, a_f) values.

B. Fit starting time

Usually, having an independent model for the post-merger data is conceptually simpler (since the ringdown analysis is fully agnostic about the pre-merger dynamics). However, truncating the GW signal at a specific time is difficult to handle with the current LIGO-Virgo analysis techniques [50], because it calls for special treatment in the time domain [37, 39], or an equivalent nontrivial procedure in the frequency domain [47]. While we are fully

[2] In [10], the grid method was set with a grid minimum step $3.2\cdot10^{-6}$ in both variables. While in this paper, most of the computation use a bisection method adaptation in grid, which is relatively faster with a factor of 2 5 even with more adaptation step and similar level of grid minimum step $3.05\cdot10^{-6}$.

[3] In this framework, one can easily show that both $\mathcal{M}$ and the χ^2 provide the same qualitative behaviour. In particular, for a model closely fitting the NR waveform, $\chi^2\simeq 2\,\mathcal{M}\,(\sum_k |h_{22}(t_k)|^2)$. Therefore, a minimum on χ^2 directly translates to a minimum in $\mathcal{M}$ and *vice versa*.

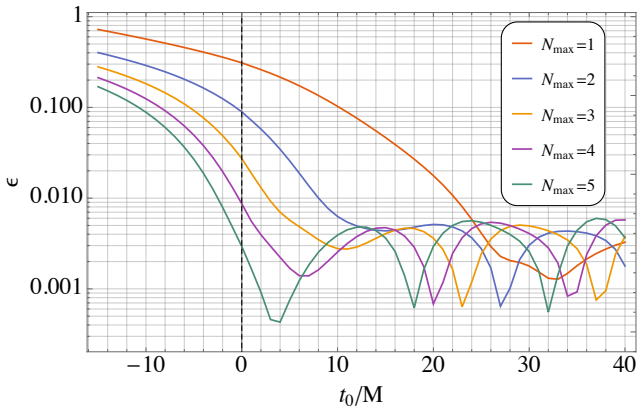

FIG. 1. We show the computed epsilon evolution with the fit starting time in overtone models with $N_{\max} \in [0, 4]$.

rely on the numerical waveform, this choice are much more reasonable to be make.

In Fig. 1, each curve corresponds to an overtone model with different number of tones. The epsilon analysis done with different fit starting time illustrates the oscillation performance[4] in tones respectively. As a result, the choose of a relatively late fit starting time might cause the error emerges within tones. Thus, we choose $t_0 = 0$ to overcome such a problem.

VI. RESULTS

In Sec. IV, we have introduced two types of modified overtone models: tone perturbed models and models with alternative forms of damped sinusoids. During the implementation of mass/spin grid method (which refers to Sec. V), we also use different computation techniques for the searching of optimal free parameters. On the one hand, for the tone perturbed models in Sec. IV A, we not only attempt to recover the optimal mass and spin from the grid computation, but seek the best-fit parameters α_n, β_n also in a 2-dimension grid (or 4-dimension grids for the 0, 1 tone perturbed cases which will be given in Sec.VI D). Therefore, each point in such a grid represents a special deviation to the Kerr spectrum (which might potentially be found the same to QNMs generated in one special modified gravity). After the initial set up of α_n/β_n grid, we insert an entire mass/spin grid into every one of the former grid points then. Which means in the grid points of α_n and β_n where the α_n, β_n values are fixed, the lowest mismatch will be searched and saved in the mass/spin grid computation (and thus the optimal mass and spin saved and fixed, so as QNMs), and then the seeking of global minimum mismatch in all the points span the α_n/β_n grid range will be make and iterate. Note that the order of which grid inside the another is physically meaningful because the reversed one of our case could be interpreted as allowing for the initial existence of mass/spin bias. On the other hand, for the other modified overtone models given in Sec. IV B, the newly added free parameters in completely different forms of ansatz will be searched with a CNMinmizer function created in [17] (an algorithms which attempts to find the global minimum of one expression with complex form, i.e. the real part + the imaginary part) in the mass/spin grid. Our computation results and further discussions are given in the following subsections.

A. Comparison of perturbation on different tones

Firstly, we test the ringdown overtones by examining the toy model which assigns the deviation to Kerr QNMs on only one of the tones in a multi-tone overtone model. In the two panels of Fig. 2, decreasing trends of both the lowest mismatch and epsilon values in tones as $N_{\max}$ increases can be observed in all the models. While all the lowest mismatches calculated in models with the nth tone perturbed are below the corresponding unperturbed overtone models, the resulting epsilons are on the opposite. Besides, in terms of the tone comparison in perturbed models: on the one hand, we can see that the highest tone always has the lowest optimal mismatch values compared to other tones in one specific $N_{\max}$ model. In other words, if we assume a single tone perturbed model (which has only one tone's QNMs deviating from Kerr), the highest tone seem to be the most preferred case statistically. Therefore, we then particularly apply most of the investigations into the highest tone perturbed models and present the results as well as further discussion in Sec. VI B. Being that said, more considerations on this aspect need to be taken as for the cases of high overtone number models. Because according to the epsilon evaluation, the trend on optimal epsilon variation with the nth perturbed tone seem to be more fluctuating in higher $N_{\max}$ models, which makes the certain feature observed in mismatch evaluation harder to tell. On the other hand, the deviation on the fundamental tone also gives relatively better mismatches when comparing to the first overtone's, which make it also worth paying attention to. Not to mention the importance of 0, 1 tone models in current ringdown analysis (because of signal loudness) and very natural thinking about the least loosely constraint characteristics of the fundamental tone as also recognized in [19]. Therefore, the related works focusing on fundamental tones are also done and presented in both Sec.VI D and Sec.VI E below.

[4] There is an argument about the potential different onsets of different tones in a "chromatic" form in [18]. We believe even if it is possible, it won't apply to our cases since it's unstable between different $N_{\max}$ models and somehow makes the models overall unable to compare.

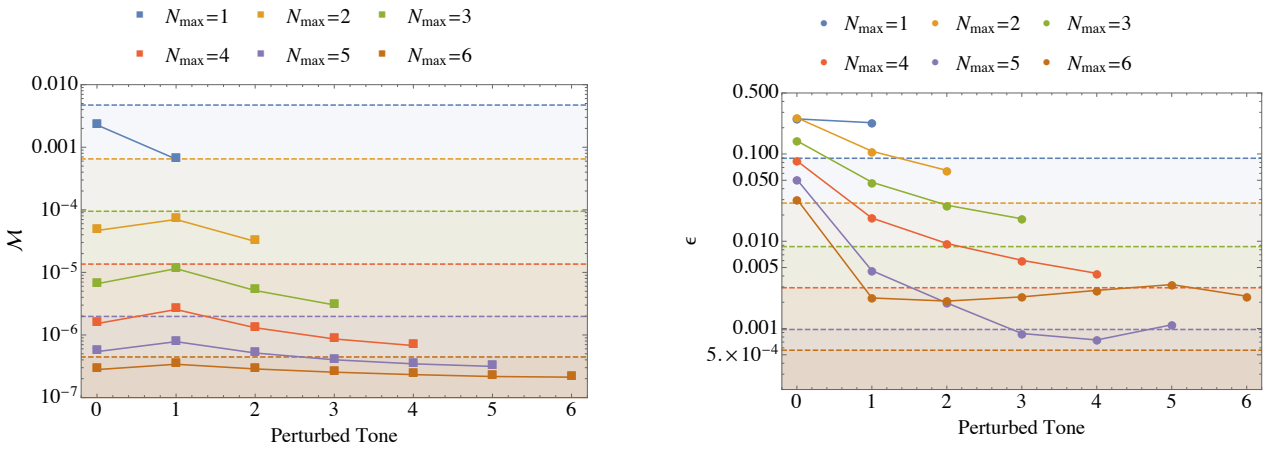

FIG. 2. The left panel shows the lowest mismatch variation in regard to different perturbed tones in overtone models with $N_{\max} \in [0, 7]$. While the right panel shows the lowest epsilon variation in regard to different perturbed tones in overtone models with $N_{\max} \in [0, 7]$. The colored region in both panels stand for the value space of either mismatch or epsilon values that are lower than those of the "unperturbed" overtone models with the corresponding $N_{\max}$.

B. Instability of the highest tones

1. *The mismatch distribution*

In Fig. 3, we can see from all the distributional contour plots that the lowest mismatch points are deviating from the origins, which means: the QNMs in the original unperturbed overtone models may not be the most statistically preferred ones in such analysis based on GR generated waveform. Particularly, if we look at the best-fit $\beta_{N_{\max}}$ values, these deviations become extremely large for $N_{\max} = 1, 2, 3$, where $\beta_3 \approx 100\%$. While we also observe the best-fit fractional deviations on frequencies (which are represented as values of $\alpha_{N_{\max}}$) do exist to some extents for almost all the overtone models we test. Therefore, this fact turns out to be very alarming because it basically just makes it possible that one mistakenly find a modified gravity (which has a different set of QNMs) being preferred over GR using the current fitting techniques. While GR might also mistakenly "pass" the ringdown tests with insufficient considerations of such intrinsic deviations on QNMs even if it shouldn't.

2. *The epsilon distribution*

As mentioned above, another quantity worth examining with the help of grid method computation is the epsilon ϵ, which measures the consistency between the recovered final mass/spin and the true values of them. To this aim, we also plot the distributional contours in regard to epsilon values in 2-dimension planes of $\alpha_{N_{\max}}$ / $\beta_{N_{\max}}$. In this figure, we observe a similar symptom as in the mismatch analysis that the point corresponds to GR in every single plot is not the most preferred one, since the optimal $\alpha_{N_{\max}}$ / $\beta_{N_{\max}}$ points in grids can have significantly lower epsilons comparing to them. Note that, since we are only considering the values included in the grids seeking lowest mismatches, the epsilon values may not be the optimal ones that should be found in grids directly seeking[5] them (which are exactly zero).

In both Fig. 3 and Fig. 9, we observe the optimal α / β values deviating from the origins in all the highest tone perturbed overtone models (from $N_{\max} = 0$ to $N_{\max} = 7$). However, those deviation factors which return the best mass/spin consistency's are different from the former best-fit ones as clearly shown in Fig. 5. Even though for some cases, such differences are minor (for instance, look at the blue curves in Fig. 5, there are less than 10% difference on the $\alpha_{N_{\max}}$ values optimized in epsilon and mismatch grids for the models with $N_{\max} = 1 \to 5$), and we do have to admit that the models with $N_{\max} = 7$ can naturally have quite large uncertainties on the highest tone's QNMs (because of the relatively large numerical errors that initially exist), these inconsistencies between the two analysis are generally leading to the worry that uncertainties on the QNMs of ringdown overtone do exist somewhere. Unfortunately, the error regions are undetermined for the α_N / β_N values computed from both cases, which might make a clearer explanation difficult to obtain. On the other hand, we also observe that the

[5] Here the "directly seeking" means we set a lower epsilon as grid object which it seeks and iterate around. While in our current analysis, we always just simply use the values stored in the grid points originally seeking the lowest mismatch. Thus it is called "not optimal".

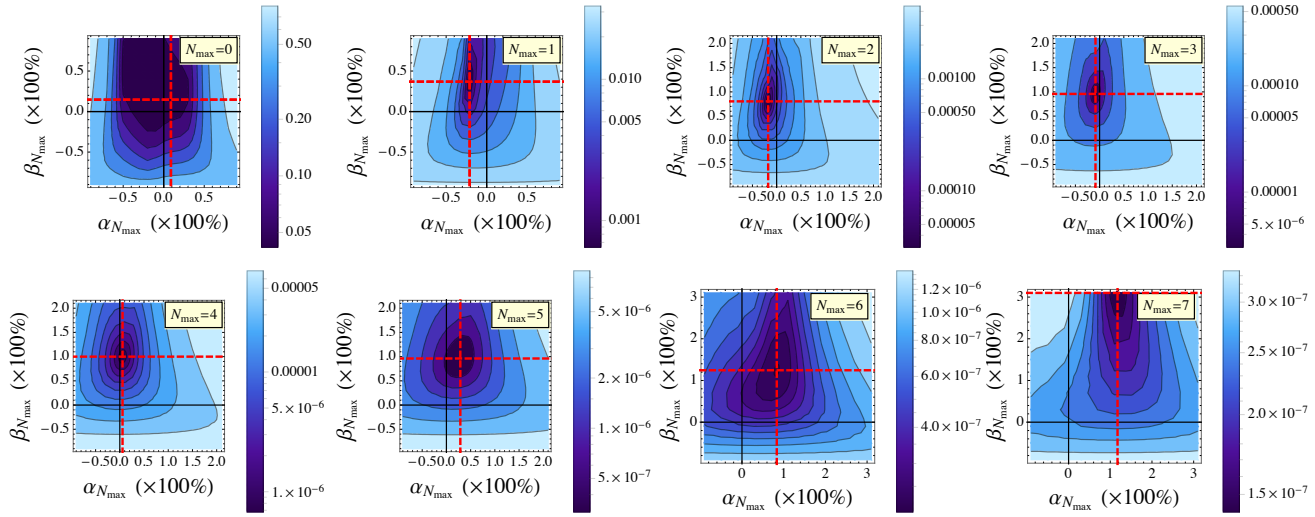

FIG. 3. In this figure we show the distribution of mismatch values in the highest tone perturbed overtone model's α / β grids ranging from $N_{\max} \in [0, 7]$, the fitting was done on the waveform data BBH:0305 from SXS catalog. The horizontal and vertical red lines given in every plot are set as references indicating the optimal values of both $\alpha_{N_{\max}}$ and $\beta_{N_{\max}}$. While the intersections correspond to points calculated from grid iterations seeking the lowest mismatch in the 2D planes.

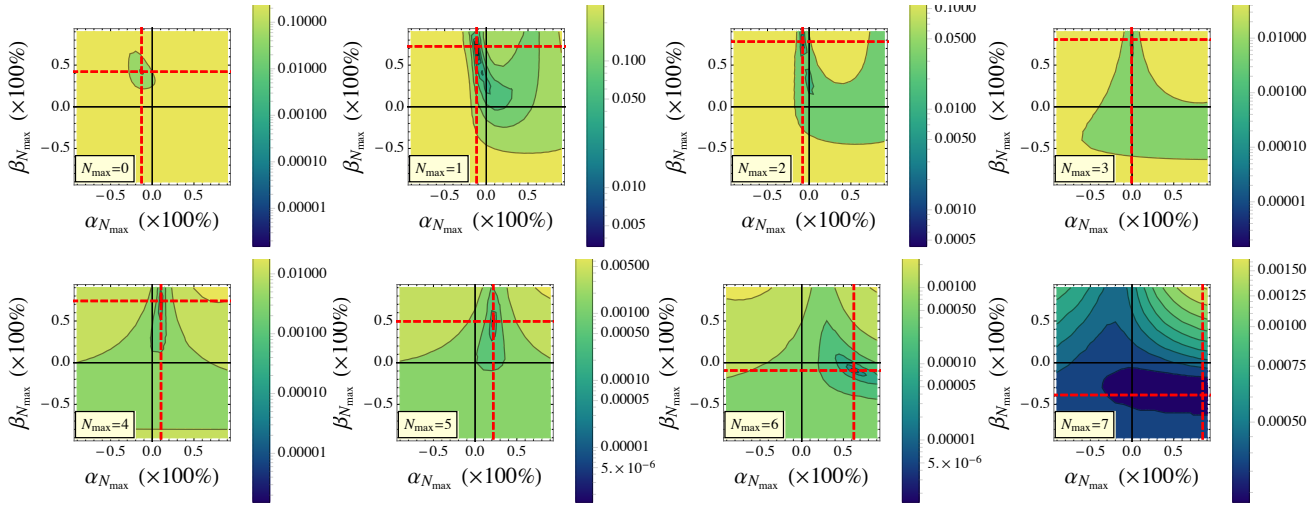

FIG. 4. In this figure we show the distribution of epsilon values in the highest tone perturbed overtone model's α / β grids ranging from $N_{\max} \in [0, 7]$, the fitting was done on the waveform data BBH:0305 from SXS catalog. The horizontal and vertical red lines given in every plot are set as references indicating the optimal values of both $\alpha_{N_{\max}}$ and $\beta_{N_{\max}}$. While the intersections correspond to points calculated from grid iterations seeking the lowest epsilon in the 2D planes.

absolute values of the optimal solutions on the deviating factors of damping time, βs, are usually larger than those of αs. For example, both $\alpha_{N_{\max}}$ with optimal ϵ and $\alpha_{N_{\max}}$ with optimal mismatch are less than 20% for models with $N_{\max} = 0 - 4$, while the $\beta_{N_{\max}}$ with optimal mismatch is around or exceeds 100% for all the cases from $N_{\max} = 3$ to $N_{\max} = 6$. In addition, the optimal βs calculated from the mismatch analysis also seem to be always positive numbers and increase in one direction. This fact suggests that there may exist a global bias factor on the time coordinate of ringdown overtones, which contribute significantly to the fitting procedures.

3. *Mismatch/epsilon consistency*

In order to consider the fitting performance and mass/spin recovery ability jointly as well as seeing the statistical insights of introducing the deviation to Kerr on QNMs, we plot the comparison of the best-fit $\epsilon - \mathcal{M}$ plane for a range of both the original unperturbed overtone models and the highest tone perturbed models in

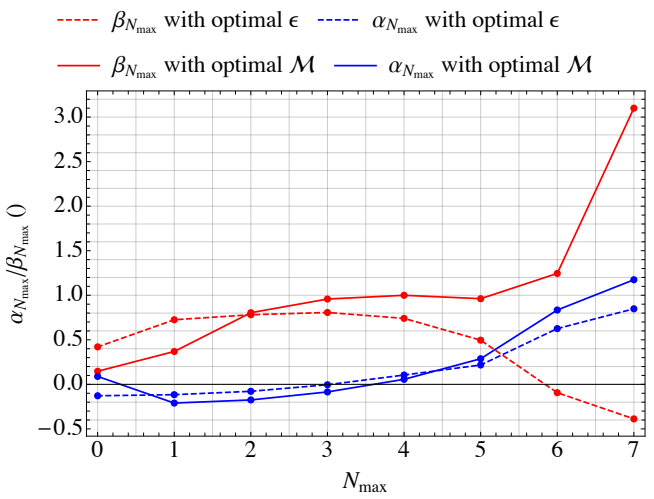

FIG. 5. We show the variation of optimized values of α_N / β_N (which are the points output the lowest mismatch or epsilon values among the 2D grids) respectively. The colors denotes the deviation types, whether on QNM frequencies or damping times, and the solid or dashed lines represent the optimal-mismatch or optimal-epsilon searching results accordingly. While each pair of the $\alpha_{N_{\rm max}}$ and $\beta_{N_{\rm max}}$ belongs to the grid computation done on the highest tone perturbed models with $N_{\rm max} \in [0, 7]$.

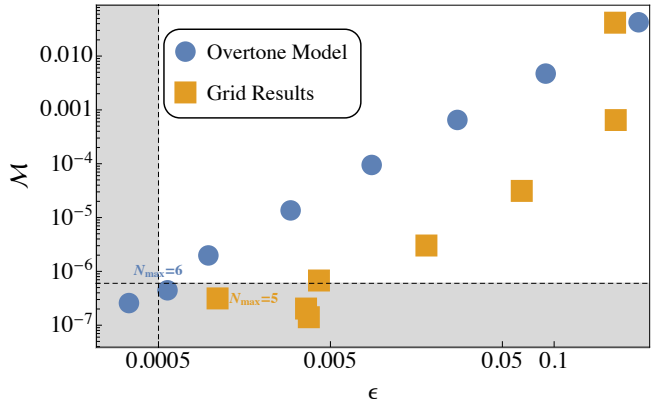

FIG. 6. We show the best-fit $\epsilon-\mathcal{M}$ plane for a range of models with $N_{\rm max} \in \{0, \ldots, 7\}$ through the NR waveform from the SXS catalogue BBH:0305 ($(q = 1.22, \chi_{\rm eff} = -0.0165)$) as for the cases of both unperturbed overtones (blue) and the highest tone perturbed overtones (yellow). The dashed lines and shaded areas at the bottom-left corner of the panel delimit the mismatch and ϵ values that are respectively smaller than the "errors" in terms of the maximum resolution/extrapolation mismatch $\max(\mathcal{M}_{\rm res}, \mathcal{M}_{\rm extr})$, and of the radioactive error $\delta\epsilon_{\rm r}$, of the SXS waveform used. (the exact values of them are cited from the Fig. 5 in [10].)

Fig. 6 with $N_max \in \{0, \ldots, 7\}$. Note that as $N_{\rm max}$ increases, the $\epsilon - \mathcal{M}$ point of one model is progressively shifted to the left bottom corner until $N_{\rm max} = 7$, where the minimum $\mathcal{M}$ is achieved for both types of overtone models. This figure shows that the highest tone perturbed overtone models always have lower mismatches compared to the same tone number $N_{\rm max}$ unperturbed models, while the original overtone models are statistically preferred in the epsilon analysis band. Which is another version of proof of the mismatch/epsilon inconsistency of model assessment apart from Fig. 5. In the mean time, we also observe from the given "error region" in Fig. 6 that the least number of tones one model requires to be interpreted as physically meaningless (because of the relatively too large errors) are $N_{\rm max} = 6$ and $N_{\rm max} = 5$ for the unperturbed (blue) and perturbed (yellow) models respectively. Therefore, the strange patterns that we have observed in high overtone number models of both Fig. 3 and Fig. 9, in which the computed optimal deviations on the highest tone go beyond the ranges we set, are actually, possible results of the numerical errors inside the waveform data.

C. Other modified models

Besides the tone perturbed models, in the test of other modified overtone models (see Sec. IV B), we also consider the cases in which the highest tones are replaced by some new forms. Thus, the clear comparison can be made between the highest tone perturbed model (whose difference with original overtone models only assign in the QNM values deviating from Kerr spectrum) and the modified models with alternative forms of damped sinusoids. Note that the latter ones are only expected to present as references for the robustness test of the former analysis done on the tone perturbed models since the exact physical meaning and clear scientific aims of using them are not identified.

In both panels of Fig. 7, the overtone models with the same numbers of real-valued free parameters are compared in regard to the same $N_{\rm max}$. From the two comparisons, it can be observed that: all the four models with alternative forms of damped sinusoids adding on the highest tones are not compatible with the original overtone models, since according to the mismatch and epsilon quantities present in the plots, they neither give better fitting to the waveform data, nor provide any more constraint on the recovery accuracy of the final mass and spin. However, the last modified model we introduce, **TCTM**, shows an exceptional advantage over all the other models on the mismatch aspect in the comparison of models with $N_{\rm max} = 2 - 4$, and also slightly lower epsilons compared to the unperturbed overtones in models with $N_{\rm max} = 2 - 3$. This result is particularly interesting because it might help us to further confirm the point that the globally re-scaling of time may significantly assist to the fitting (see previous discussions on Fig. 5). But since we are still not very clear about the insight of **TCTM**, we shall stop the further discussion now and leave it to the future works. Note that here in the comparison, the

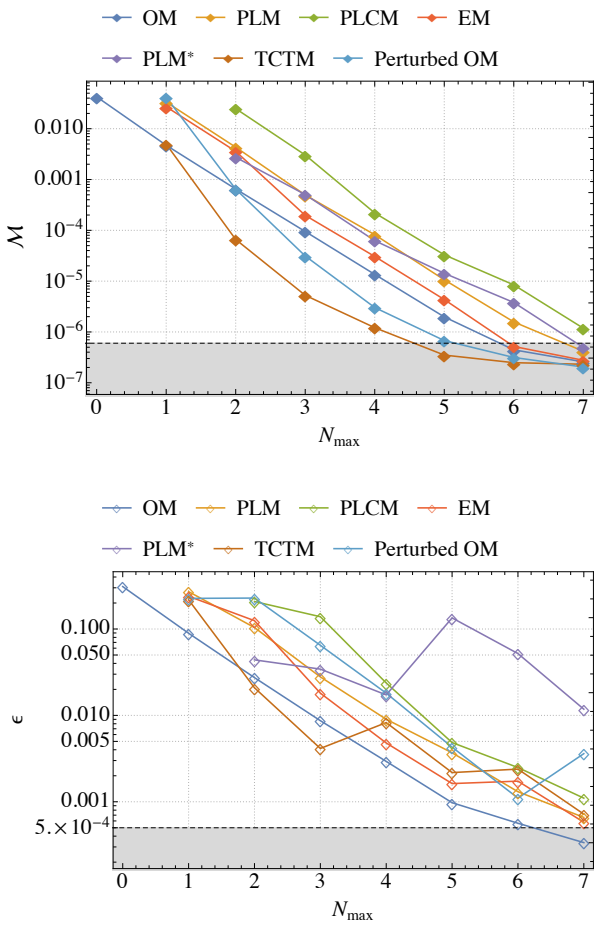

FIG. 7. In the upper panel, we show the comparison of mismatch evolution with $N_{\max} \in [0, 7]$ in original overtone models, highest tone perturbed model (Perturbed OM), and other modified models as introduced in Sec. IV B, the computation was done on the BBH:0305 waveform data from SXS catalogue. In the lower panel, the comparison of epsilon evolution with $N_{\max} \in [0, 7]$ in these models are presenting as well. The dashed lines and shaded areas at the bottom of both panels delimit the mismatch (upper) or ϵ (lower) values that are respectively smaller than the "errors" in terms of the maximum resolution/extrapolation mismatch max $(\mathcal{M}_{\rm res}, \mathcal{M}_{\rm extr})$, or of the radioactive error $\delta\epsilon_{\rm r}$, of the SXS waveform used.

number of consistent tones[6] are different for all the models, which might cause higher errors on the mass/spin estimations then. Finally, we conclude from such observations that: in future works, we had better look more into the uncertainties on the QNM values than those on the different forms of ansatz when attempting to test the instability of overtones with GR assumptions.

D. 0, 1 tone perturbed models

As mentioned in the Sec. I, one possible channel of doing black hole spectroscopic analysis is to detect the fundamental tone and the first overtone together. The deviation to the Kerr spectrum on such a model and the uncertainties associated have been calculated in several ways recently. While we also try to implement this through our grid method in this paper through two approaches: i). assume only the overtone QNMs are perturbed, ii). assign the fractional deviation parameters on both of the tones. Combining the results of both the two approaches, a more general conclusion on the instability of this 0, 1 tone perturbed overtone models could be made.

In order to facilitate the comparison between ours and other works, we take some extra steps from the above sections' works so that some similar plots can be obtained. Firstly, we use the information saved in our computation of the 2-dimension α_N / β_N grids by marginalizing the mismatch distribution. Then, injections with different SNR values could be tested and inserted to the form with the assuming of Gaussian noise:

$$Distribution = Ne^{-\rho^2\mathcal{M}}. \tag{7}$$

where the N denotes the normalization factor. We then test with $SNR = \{10, 20, 50\}$ injections using the above distribution formula.

E. Stability of the fundamental tone

In the original RD overtone models, the higher overtones are predicted to decay faster. Each time one adds the N-th tone to the $N_{\max} = N$ overtone model, the early time fitting (near the peak) shall be improved. While they always exist a shorter period as N increases, the effects the highest tone could bring to the fitting goodness decreases, so as the QNM uncertainties. Thus, we believe the fundamental tone should be the best constraint one compared to the other tones. In our works, a strange relatively lower mismatch and epsilon values in the fundamental tones are also observed in the tone comparison plot Fig. 2. In this paper, we are particularly interested about the instability of the fundamental tone and how it varies in different $N_{\max}$ models.

We can see from the Fig. 10 which strongly proves the constraint effects higher overtones could bring to the fundamental tone as they added in the overtone models.

[6] Here the consistency within tones are controllable and countable because of the grid method we use. Except for the tone perturbed models, whose QNMs in all the tones are systematically determined by the final mass and spin (the deviation factors are added above them), all the other modified models have different forms of damped sinusoids in the highest tone, and thus don't have physically the same meanings of quasi-normal frequencies and damping times any longer. Therefore, in the 2D grid method computations, there are only $N_{\max} - 1$ or $N_{\max} - 2$ tones' QNMs are determined by and vary with the mass/spin, which then result in the less consistencies within tones theoretically.

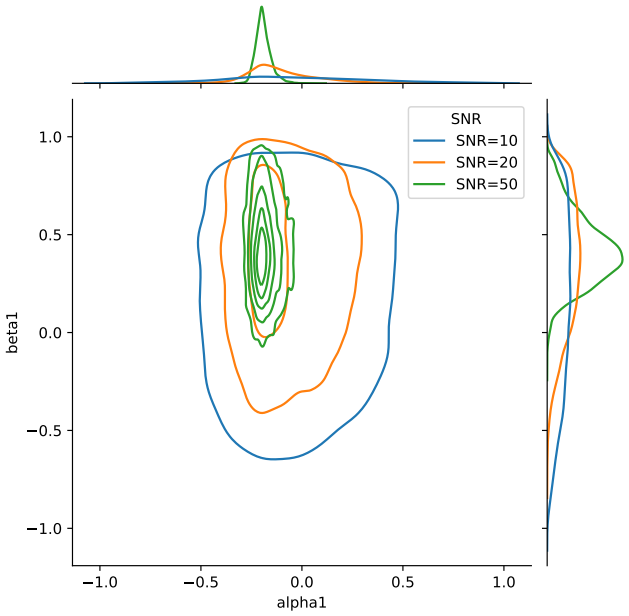

FIG. 8. We show the contour plots of the mismatch in a α_1 / β_1 2-dimension plane with injection SNR=10, 20, 50, the computation was done on the assuming of only the first overtone perturbed 2 tone models.

VII. CONCLUSIONS

The aim of this work is to test the instability of overtone models. To this end, we firstly give an overview of the key quantities $\mathcal{M}$ and ϵ and techniques used in RD analysis. Secondly, we briefly talk about the modified overtone models we consider and how the instabilities can possibly present and be measured in both the QNMs fractional deviation to Kerr spectrum and the modified forms of damped sinusoids as in the original models. Then we introduce our grid method, with which we believe can tackle the challenges in considering recovered mass/spin consistency and fitting performance jointly, while presenting the results in an very clear and unprecedented way. Finally, all the results including the instability came from QNMs deviation or alternative damped sinusoids are presented in figures and given with detailed explanations and discussions. The main conclusions of this paper can be split into four points:

- The higher tones have higher QNM uncertainties.
- The instability are more on the QNM values than the altenative forms of damped sinusoids.
- The fundamental tone could be better constraint with higher overtones involved.
- The intrinsic uncertainties of the 0, 1 tone overtone models maybe larger than expected.

The overall results are somehow unexpected since the numerical relativity data implemented in our analysis are fully generated with GR. Thus, they can only be interpreted as the intrinsic uncertainties of the original overtone models' QNMs.

In the future RD analysis, adequate tones should be included if possible, otherwise one should give particularly careful considerations to the priori uncertainties inside the overtone models before concluding anything on the gravity theory testing aspect.

ACKNOWLEDGMENTS

This paper is a product of the summer internship carried out virtually under the Max Planck Institute for Gravitational Physics. We acknowledge the Max Planck Gesellschaft for support and we are grateful to the Atlas cluster computing team at AEI Hannover for their help. The authors are also thankful to all the members including supervisors and students in this intern: Alex Nitz, Amanda Garcia, Collin Capano, Sarah Sweeney, Simran Dave, Soham, Sumit Kumar, and Yifan Wang for useful discussions.

[1] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. X **9**, 031040 (2019), arXiv:1811.12907 [astro-ph.HE].
[2] R. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. X **11**, 021053 (2021), arXiv:2010.14527 [gr-qc].
[3] S. A. Teukolsky, Phys. Rev. Lett. **29**, 1114 (1972).
[4] M. Giesler, M. Isi, M. A. Scheel, and S. A. Teukolsky, Phys. Rev. X **9**, 041060 (2019), arXiv:1903.08284 [gr-qc].
[5] L. London, D. Shoemaker, and J. Healy, Phys. Rev. **D90**, 124032 (2014), [Erratum: Phys. Rev.D94,no.6,069902(2016)], arXiv:1404.3197 [gr-qc].
[6] E. Leaver, Proc. Roy. Soc. Lond. A **A402**, 285 (1985).
[7] S. L. Detweiler, Astrophys. J. **239**, 292 (1980).
[8] 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].
[9] The SXS Collaboration, "SXS Gravitational Waveform Database," (2016).
[10] X. J. Forteza and P. Mourier, arXiv:2107.11829 [gr-qc] (2021), arXiv:2107.11829 [gr-qc].
[11] L. C. Stein, J. Open Source Softw. **4**, 1683 (2019), arXiv:1908.10377 [gr-qc].
[12] M. Giesler, M. Isi, M. Scheel, and S. Teukolsky, Phys. Rev. X **9**, 041060 (2019), arXiv:1903.08284.
[13] S. Bhagwat, X. J. Forteza, P. Pani, and V. Ferrari, Phys. Rev. D **101**, 044033 (2020).
[14] X. J. Forteza, S. Bhagwat, P. Pani, and V. Ferrari, Phys. Rev. D **102**, 044053 (2020), arXiv:2005.03260.
[15] P. Mourier, X. Jimenez-Forteza, D. Pook-Kolb, B. Krish-

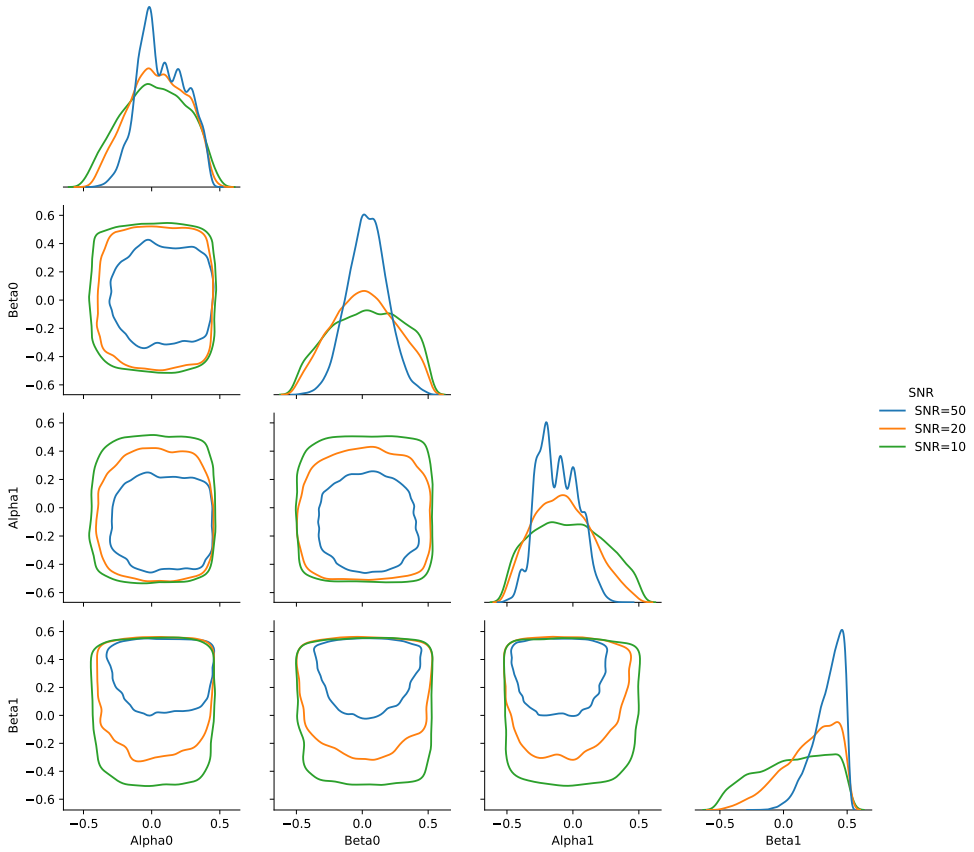

FIG. 9. We show the pair plot of the mismatch distribution in a α_0 / β_0, α_1 / β_1 4-dimension plane with injection SNR=10, 20, 50, the computation was done on the assuming of all the overtones perturbed 2 tone models.

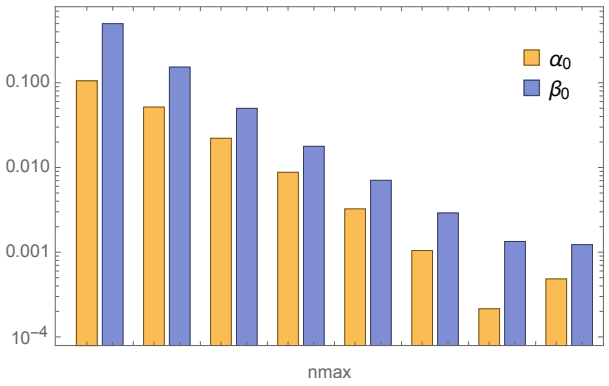

FIG. 10. We show the absolute best-fit α_0 / β_0 values computed in the fundamental tone perturbed models with $N_{\max} \in [0, 7]$, the results are taken averaged within 5 SXS waveform: BBH:0259, BBH:0295, BBH:0297, BBH:0300, BBH:0305.

nan, and E. Schnetter, Phys. Rev. D **103**, 044054 (2021), arXiv:2010.15186.
[16] L. Manfredi, J. Mureika, and J. Moffat, Physics Letters B **779**, 492 (2018).
[17] X. Jimenez Forteza, `https://github.com/frcojimenez/GW_Rdown`.
[18] H.-T. Wang, Y.-M. Hu, and Y.-Z. Fan, arXiv:2109.03113 [astro-ph, physics:gr-qc] (2021), arXiv:2109.03113 [astro-ph, physics:gr-qc].
[19] M. Isi and W. M. Farr, arXiv:2107.05609 [astro-ph, physics:gr-qc] (2021), arXiv:2107.05609 [astro-ph, physics:gr-qc].