Searching for a ringdown overtone in GW150914

Eliot Finch and Christopher J. Moore

Searching for a Ringdown Overtone in GW150914

Eliot Finch [ID]* and Christopher J. Moore [ID]†

Institute for Gravitational Wave Astronomy & School of Physics and
Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

(Dated: July 19, 2022)

We reanalyze the GW150914 post-merger data searching for quasinormal modes beyond the fundamental, quadrupolar mode. There is currently an ongoing disagreement in the literature about whether, and to what extent, the data contains evidence for a quasinormal mode overtone. We use a frequency-domain approach to ringdown data analysis that was recently proposed by the authors. Our analysis has several advantages compared to other analyses performed mainly in the time domain; in particular, the source sky position and the ringdown start time are marginalized over (as opposed to simply being fixed) as part of a Bayesian ringdown analysis. We find tentative evidence for an overtone in GW150914, but at a lower significance than reported elsewhere. Our preferred analysis, marginalizing over the uncertainty in the time of peak strain amplitude, gives a posterior on the overtone amplitude peaked away from zero at $\sim 1.8\sigma$.

I. INTRODUCTION

The gravitational-wave (GW) signals from binary coalescences provide a unique opportunity to study gravity in the strong-field and dynamical regimes. Of particular interest here is the signal from the final stages of a binary black-hole (BH) merger, known as the ringdown, which is associated with the remnant BH settling into its final state. The ringdown contains a superposition of exponentially damped oscillations, known as quasinormal modes (QNMs), with a discrete set of (complex) frequencies. Identifying these frequencies allows us to measure properties of the remnant and also provides a particularly clean way to test general relativity (GR) and the Kerr metric; this procedure is known as black hole spectroscopy [1]. QNMs have now been identified in a few tens of binary BH merger signals in the most recent GW catalogs [2, 3].

However, the very first GW event, GW150914 [4], remains probably the best candidate for studying the ringdown. This is a result of several factors, including its large signal-to-noise ratio (SNR) of $\rho \sim 24$ and its total mass of $M \sim 70\,M_\odot$ which places the merger and ringdown in the center of the LIGO [5] sensitive frequency band at $\sim 200\,\mathrm{Hz}$. Additionally, GW150914 is by now the most well-studied GW event and therefore the signal and the properties of the noise in the surrounding data are extremely well understood.

The first tests of GR performed using GW150914 included an investigation of the ringdown [6]. The ringdown signal, after a fixed starting time t_0, was modeled using a single damped sinusoid; the parameters of which were checked for consistency with the predicted least-damped QNM of the remnant BH. This first attempt at a ringdown analysis was performed using the standard Whittle frequency-domain log-likelihood [7], commonly

used in GW data analysis. The ringdown was isolated by choosing a lower limit of $\sim 130\,\mathrm{Hz}$ in the frequency integral, effectively cutting the data mid-signal. This approach suffers from several shortcomings. Firstly the frequency-domain cut at $\sim 130\,\mathrm{Hz}$ only approximately separates the ringdown from the early-time signal due to the breakdown of the stationary phase approximation near merger. Secondly the non-zero amplitude at the start of the signal model breaks the assumption of circularity for the Fourier transform, thereby introducing contamination in the form of spectral leakage. Therefore, this approach does not scale well to higher SNRs where noise will no longer dominate over the systematic errors introduced by the sharp frequency-domain cut. Despite these drawbacks, this approach was successfully used in Ref. [6] to identify the fundamental QNM in the GW150914 signal.

Since this initial attempt, several groups have developed new time-domain frameworks specifically for ringdown analyses [8–10]. The principle motivation for working in the time domain is that it is easy to impose sharp cuts on the data at specific times (without any spectral leakage) and to analyze only data after a chosen start time (see Ref. [11] for details of time-domain analysis methods). These approaches have also enabled going beyond the fundamental mode. Generically, the ringdown can be modeled as a superposition of QNMs with complex frequencies $\omega_{\ell m n} = 2\pi f_{\ell m n} - i/\tau_{\ell m n}$, labeled with angular indices $\ell \geq 2$, $|m| \leq \ell$, and an overtone index $n \geq 0$ [the fundamental mode has $(\ell, |m|, n) = (2, 2, 0)$]. Detecting additional QNMs beyond the fundamental increases the scientific potential of ringdown studies, especially for fundamental tests of the Kerr metric, the no-hair theorems, and the BH area law [1, 8, 9, 12–15].

An early application of the time-domain framework was in Ref. [9], where Isi et al. claimed a detection of the first overtone of the fundamental QNM in the GW150914 signal [that is, the $(2, 2, 1)$ mode]. This was quickly followed by a separate detection claim of the $(3, 3, 0)$ harmonic mode in the signal of the $\sim 150 M_\odot$ binary merger GW190521 [16] by Capano et al. [10] (this was done using

---

an equivalent formulation of the time-domain method, although expressed in the frequency domain). The claimed detection of an overtone was made possible partly because, compared to earlier studies, the authors chose to use an earlier start time for the ringdown; this was motivated by contemporary numerical relativity studies [17] (see also Refs. [18–26]) that demonstrated that when overtones are included the ringdown can be considered to start as early as the time of peak strain amplitude.

However, a recent paper by Cotesta et al. [27] reanalyzed the GW150914 signal using very similar methods and found no significant evidence for an overtone. It was also suggested that the earlier detection claims of Ref. [9] were noise dominated. (This prompted a response from Isi et al. [28] where they restated their claim to have detected an overtone in GW150914.) Ref. [29] also found weaker evidence for an overtone using an analysis method closer to that of Ref. [6]. Similarly, the claim in Ref. [10] that a harmonic had been detected in GW190521 has also been debated and no evidence for a harmonic was found by Ref. [3]. Amid this confusion, it is particularly concerning that the supposedly identical analyses in Refs. [9, 28], and [27] come to such different conclusions concerning which QNMs are in the data. Discrepancies of this sort risk jeopardizing the science that can be done using future ringdown observations.

These discrepancies highlight some of the difficulties inherent in time-domain ringdown analysis, where important choices (that affect the results) for fixed quantities such as the ringdown start time have to be made and care must be taken with the noise covariance estimation. If ringdown studies are to be used to make precision measurements of BH properties or as reliable tests of GR we must first be able to make reliable and reproducible determinations of the QNM content. This is also not a problem that will be removed in future with observations at higher SNR. Even if an event has a higher SNR that is sufficient for a clear detection of the first QNM overtone, the focus will then simply shift to trying to identify the next overtone (or else the next QNM harmonic) in the countably infinite ringdown sum [29].

To complement the time-domain analysis frameworks, the authors recently proposed a new method for ringdown analyses which works in the frequency domain [30]. A flexible sum of sine-Gaussian wavelets, truncated at the ringdown start time, is used to effectively marginalize over the inspiral-merger (i.e. pre-ringdown) part of the signal. The model is completed by attaching this to the usual sum of QNMs which model the ringdown. No continuity is enforced between the two parts of the model in order to keep the ringdown inference independent from the rest of the signal. However, we find the continuity is effectively learned from the data, and any remaining discontinuities disappear entirely when the signal is "whitened" according to the instrumental noise. In a particular limit, this approach can be shown to be formally equivalent to the time-domain analyses described above. However, this frequency-domain approach can

be generalized and offers several advantages over time-domain approaches: well-established GW data analysis methods and pipelines can be used (which are all built in the frequency domain), the inspiral-merger data informs the noise estimation at the start of the ringdown (improving parameter estimation accuracy), and the ringdown start time and the source sky position can be easily treated as free parameters and marginalized over as part of a Bayesian analysis (instead of being fixed). We note, however, that (as discussed in Ref. [30]) a narrow and informative prior on the ringdown start time must be used. Reweighting techniques can be employed to investigate different ringdown start time prior choices computationally efficiently in post processing (see Sec. II C) obviating the need for the large number of analyses performed in [27, 28].

In this paper the new frequency-domain method is applied to reanalyzing the ringdown of GW150914 paying particular attention to the presence (or absence) of an overtone. We perform analyses with and without an overtone and investigate different choices of the ringdown start time. We also perform additional analyses with varying data sampling frequencies and integration limits to verify the stability of our results. Finally, a mock injection study into real detector noise is also performed to further assess the significance of any overtone detection. Sec. II describes the signal model, the data, and the analysis methods used in this paper. Sec. III presents our main results including posteriors on the remnant BH properties and overtone amplitude, and Bayes' factors for the overtone model. The results are discussed further in Sec. IV. Throughout this paper we make use of natural units where $G = c = 1$.

All data products and plotting scripts used to make the figures in this paper are made publicly available at Ref. [31].

II. METHODS

This section briefly describes the frequency-domain method for analyzing BH ringdowns introduced in Ref. [30]: the wavelet-ringdown model is described in Sec. II A; the data, likelihood and priors are described in Sec. II B; and our approach for dealing with changes to the ringdown start time is described in Sec. II C.

A. Wavelet-Ringdown Model

Our model consists of two parts: one for early times before t_0 which is referred to here as the *inspiral-merger*, and another for the *ringdown* after the start time t_0.

First, we describe the ringdown part of the model. After a ringdown start time t_0, which is itself a parameter

in the model, the model takes the form

$$h^{\mathrm{R}}(t) = h^{\mathrm{R}}_+(t) + ih^{\mathrm{R}}_\times(t)$$
$$= \sum_{n=0}^{N} A_n e^{-i[\omega_{22n}(t-t_0)+\phi_n]}, \quad t \geq t_0. \quad (1)$$

Because our focus in this paper is on the presence of an overtone, we fix the angular indices to $\ell = m = 2$ and vary only the number of QNM overtones, N, in the model (N is always taken to be either 0 or 1 in this paper). Note that the form of this equation differs slightly from Eq. 11 in Ref. [30]. This is because the source inclination angle is fixed to be "face-off" (i.e. $\iota = \pi$). In the notation of (for example) Refs. [22, 23, 26], this is equivalent to using the $\ell = -m = 2$ mirror modes. Or, in notation of Ref. [11], using an ellipticity of $\epsilon = -1$. The complex QNM frequencies, $\omega_{\ell m n} = 2\pi f_{\ell m n} - i/\tau_{\ell m n}$, are functions of the remnant BH mass M_f (detector frame) and dimensionless spin χ_f. Additionally, each QNM is further described by an amplitude, A_n, and a phase, ϕ_n.

Second, we describe the inspiral-merger part of the model. This is modeled as a truncated sum of W wavelets. At early times the model takes the form

$$h^{\mathrm{IM}}(t) = h^{\mathrm{IM}}_+(t) + ih^{\mathrm{IM}}_\times(t)$$
$$= \sum_{w=1}^{W} \mathcal{A}_w \exp\left[-2\pi i\nu_w(t-\eta_w) \quad (2)\right.$$
$$\left. -\left(\frac{t-\eta_w}{\tau_w}\right)^2 - i\varphi_w\right], \quad t < t_0.$$

Again, the minor differences in sign conventions compared to Ref. [30] come from fixing the inclination angle to be face-off. The wavelets are each described by five parameters: $\mathcal{A}_w$ and φ_w are the wavelet amplitudes and phases, τ_w are the wavelet widths, ν_w are the wavelet frequencies, and η_w are the wavelet central times. In this paper we use $W = 3$ (three wavelets) in our model. This number was empirically found to be sufficient (see the appendix of Ref. [30], where the number of wavelets was varied for a GW150914-like injection).

The full signal model is given by discontinuously joining the inspiral-merger to the ringdown at t_0,

$$h(t) = h^{\mathrm{IM}}(t) + h^{\mathrm{R}}(t). \quad (3)$$

Finally, the detector response must be considered. We project the waveform polarizations onto each interferometer (IFO) with the antenna patterns, $F_{+,\times}^{\mathrm{IFO}}$. The detector response for each IFO $\in \{\mathrm{H}, \mathrm{L}\}$ is given by

$$h^{\mathrm{IFO}}(t) = F_+^{\mathrm{IFO}}(\alpha, \delta, \psi)\, h_+(t+\Delta t_{\mathrm{IFO}})$$
$$+F_\times^{\mathrm{IFO}}(\alpha, \delta, \psi)\, h_\times(t+\Delta t_{\mathrm{IFO}}), \quad (4)$$

where α, δ are the source right ascension and declination, and ψ is the GW polarization angle. The time delay $\Delta t_{\mathrm{IFO}}(\alpha, \delta)$ accounts for the different signal arrival times at the detectors and is also a function of the

source sky location. Throughout this paper we quote times in the Hanford frame. So, in particular, t_0 refers to the ringdown start time in Hanford. By definition, $h_+(t) = \mathrm{Re}\{h(t)\}$, and $h_\times(t) = \mathrm{Im}\{h(t)\}$.

B. Data and Priors

We use the GW150914 strain data sampled at $4096\,\mathrm{Hz}$ for both the Hanford and Livingston interferometers, which was obtained from [32, 33]. A total of $4096\,\mathrm{s}$ of data around the event was downloaded, from which the mean was subtracted (this is effectively equivalent to applying a $\sim 1\,\mathrm{Hz}$ highpass filter). Pre-computed power spectral densities (PDSs) associated with GW150914 from the GWTC-1 release were used [34]. It has been verified our results are insensitive to the exact noise PSD used; for example, our results are unchanged when using a PSD estimated from a length of off-source data. The analysis data consists of $4\,\mathrm{s}$ of data centered on the event GPS time ($1126259462.4\,\mathrm{s}$), and a Tukey window with an alpha parameter of 0.2 was applied to this analysis data. The Bayesian analysis used the standard frequency-domain log-likelihood function (see, e.g., Eq. 1 in Ref. [30]), with the limits of the frequency integration between 20 and $1000\,\mathrm{Hz}$. The choices of sampling rate and upper limit of frequency integration are discussed further in appendix A.

All the model parameters described in Sec. II were sampled over as part of a Bayesian analysis. For the wavelet parameters, uniform priors are used for the amplitudes ($\mathcal{A}_w \in [0, 10^{-20}]$), phases ($\varphi_w \in [0, 2\pi]$), frequencies ($\nu_w \in [20, 200]\,\mathrm{Hz}$), and widths ($\tau_w \in [4, 80]\,\tilde{M}_f$, or equivalently $\sim [1.4, 27]\,\mathrm{ms}$). Here, $\tilde{M}_f = 68.779 M_\odot = 0.33875\,\mathrm{ms}$ is a fixed point estimate of the final, detector-frame mass (obtained using the median value from Ref. [35]) and should not be confused with the varying model parameter M_f. The label-switching ambiguity among the wavelets was removed by enforcing the ordering $\nu_w \leq \nu_{w+1}$ via the *hypertriangulation* transformation described in Ref. [36]. We sample over the wavelet central times (η_w) using a Gaussian prior in the Hanford frame with a width of $50\,\tilde{M}_f$ ($\sim 17\,\mathrm{ms}$) centered on $t_{\mathrm{ref}} = 1126259462.423\,\mathrm{s}$. This choice was found to be sufficiently flexible, whilst at the same time encouraging the wavelets to accurately model the signal near the peak (see the discussion in Ref. [30]).

For the ringdown, uniform priors are used for the amplitudes ($A_n \in [0, 10^{-19}]$), phases ($\phi_n \in [0, 2\pi]$), remnant mass ($M_f \in [40, 100]\,M_\odot$), remnant spin ($\chi_f \in [0, 0.99]$) and ringdown start time ($t_0 - t_{\mathrm{ref}} \in [-15, 15]\,\tilde{M}_f$, which in SI units corresponds to $\sim [-5.1, 5.1]\,\mathrm{ms}$). We use a uniform prior on t_0 so that the samples can be easily reweighted in post-processing (see Sec. II C). For the remaining parameters, we used a uniform prior over the sphere of the sky (parameterized using α and δ) for the source location and a flat, periodic prior on the polarization angle ψ in the range 0 to π.
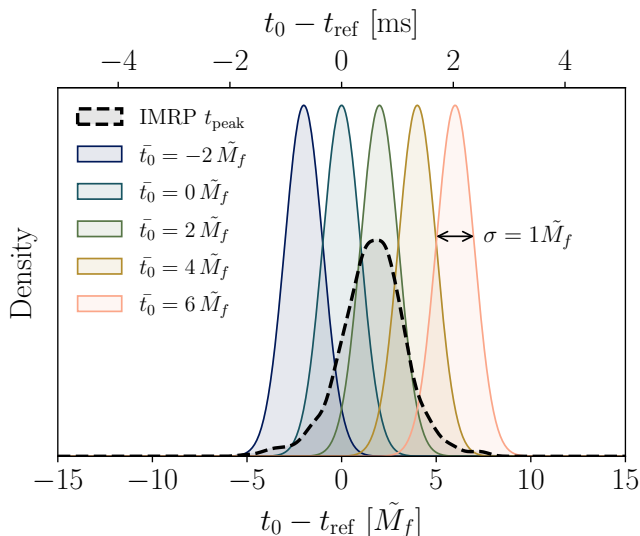
FIG. 1. Our ringdown inference is run initially using a flat, uniform prior on the ringdown start time, t_0, over the plot range $\pm 15\tilde{M}_f$ relative to $t_{\rm ref}$ (Hanford frame). In post processing, the posterior samples can be reweighted to a different choice of prior on t_0 (see Sec. II C). The different prior choices used in this paper are shown in this figure. We use a sequence of narrow Gaussian priors (with different means $\bar{t}_0$ defined relative to $t_{\rm ref}$ and fixed standard deviation, $\sigma = 1\tilde{M}_f$) as well as using the posterior on the time of peak strain from a full IMR analysis as a prior.

The nested sampling [37] algorithm as implemented in DYNESTY [38] was used to sample the posterior with 4000 live points and using the random walk sampling method with a walk length parameter of 2000.

C. Reweighting

An ever present issue in ringdown analyses is the choice of ringdown start time, t_0, and this choice is closely related to the issue of the presence of an overtone. To address this issue, previous time-domain analyses [9, 27, 28] perform large numbers of Bayesian analysis runs with different choices of start time.

One key conceptual benefit of the frequency-domain approach of Ref. [30] is that the ringdown start time enters as a parameter of the model and can therefore be easily marginalized over, instead of simply being fixed (although, see Ref. [8] where the ringdown start time was varied in a time-domain analysis). However, it is necessary to choose an informative (narrow) prior for the parameter t_0.

A related computational benefit of our approach is that we can do a *single* Bayesian analysis run with a broad uniform prior on t_0. We can then explore different, narrower priors by reweighting the results in post processing. This is an example of *importance sampling* (see, for example, [39]) and is the approach adopted here. This

removes the need to perform the large number of runs used to explore the effect of varying the ringdown start time when performing time-domain ringdown analyses.

Given a model that depends on parameters $\vec{\theta}$, a likelihood $\mathcal{L}(\text{data}|\vec{\theta})$, and a prior $\pi(\vec{\theta})$, nested sampling can be used to draw a large number of samples $\vec{\theta}_i$ from the posterior, which is given by Bayes' theorem $P(\vec{\theta}|\text{data}) \propto \mathcal{L}(\text{data}|\vec{\theta})\pi(\vec{\theta})$. Samples from the posterior have associated weights w_i (samples may often be equally weighted with $w_i = 1$, but we do not require this to be the case). Such samples can be used to approximate integrals via a Monte-Carlo sum; $\int \mathrm{d}\vec{\theta} \, P(\vec{\theta}|\text{data})f(\vec{\theta}) = \sum_i w_i f(\vec{\theta}_i)/W$, where $W = \sum_i w_i$. If we choose a new prior $\hat{\pi}(\vec{\theta})$, then the Bayesian posterior is given instead by $\hat{P}(\vec{\theta}|\text{data}) \propto \mathcal{L}(\text{data}|\vec{\theta})\hat{\pi}(\vec{\theta})$. We can define the new weights via

$$\hat{w}_i = w_i \frac{\hat{\pi}(\vec{\theta}_i)}{\pi(\vec{\theta}_i)}. \tag{5}$$

In this way the same samples can be used to approximate integrals of the form $\int \mathrm{d}\vec{\theta} \, \hat{P}(\vec{\theta}|\text{data})f(\vec{\theta})$ via the Monte-Carlo sum $\sum_i \hat{w}_i f(\vec{\theta}_i)/\hat{W}$, where $\hat{W} = \sum_i \hat{w}_i$.

It is also possible to reweight the Bayesian evidence for the new choice of prior. In a GW context this approach has been used previously for inference with higher-order modes [40]. The Bayesian evidence (i.e. the normalization denominator in Bayes' theorem) under the original prior is given by $Z = P(\text{data}) = \int \mathrm{d}\vec{\theta} \, \mathcal{L}(\text{data}|\vec{\theta})\pi(\vec{\theta})$. The Bayesian evidence under the new prior, $\hat{\pi}(\vec{\theta})$, is $\hat{Z} = \int \mathrm{d}\vec{\theta} \, \mathcal{L}(\text{data}|\vec{\theta})\hat{\pi}(\vec{\theta})$. Using the reweighted samples to approximate the integral, it can be shown that the new evidence is given by

$$\hat{Z} = Z \frac{\hat{W}}{W}. \tag{6}$$

The process of reweighting to the new, target prior reduces the effective number of posterior samples available. For this not to be a problem, we require the original prior to have significant support across the target prior. Here, we reweight on just a single parameter, the ringdown start time t_0. As described above, we use a uniform prior on t_0 as the original prior, π, in our analyses. For the target prior we use a variety of different choices, this removes the need for performing a large number of runs with different start times. Our prior choices are plotted in Fig. 1. Narrow Gaussians centered at different start times are used to explore the start time dependence on the results, and we use the notation $\bar{t}_0$ to indicate the mean of the Gaussian relative to $t_{\rm ref}$. For more details on the t_0 reweighting, see appendix B.

We also use the posterior on $t_{\rm peak}$ from a full inspiral-merger-ringdown (IMR) analysis from Ref. [28], obtained with the IMRPHENOMPV2 (IMRP) waveform model [41], as another prior on t_0. Our aim in doing this is to marginalize over our uncertainty on the ringdown start
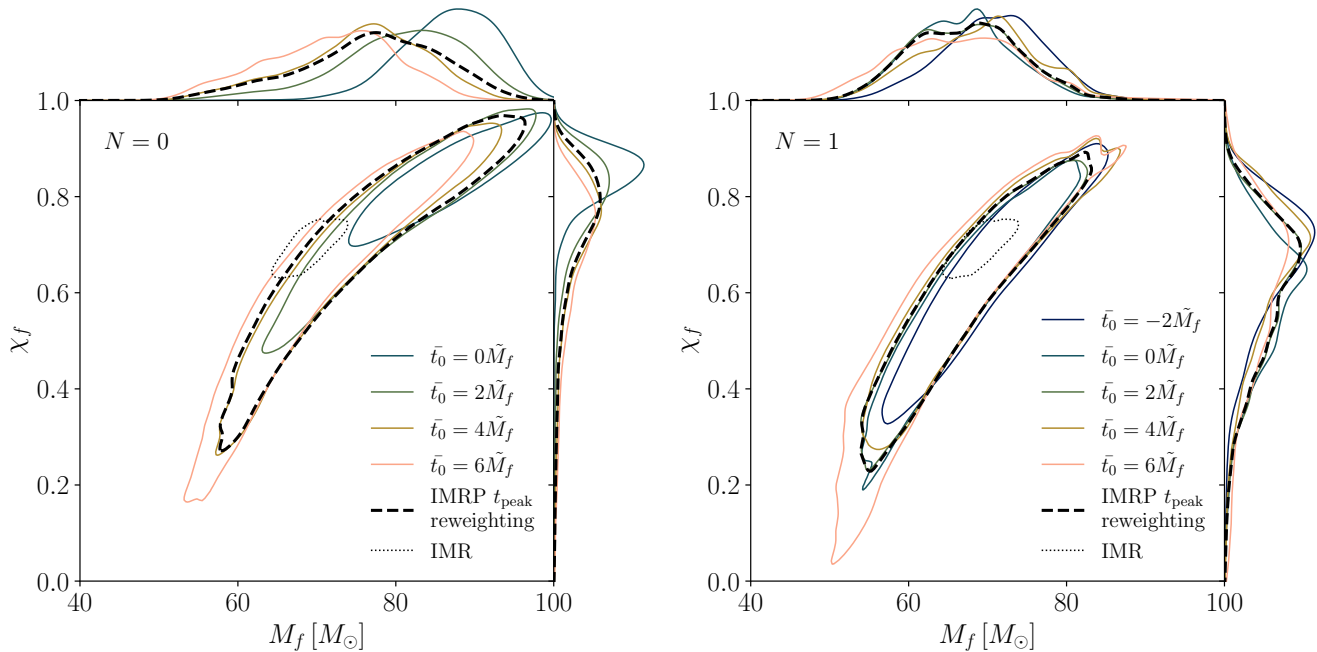
FIG. 2. Posterior distributions on the remnant mass, M_f, and dimensionless spin, χ_f, for different choices of t_0 prior (the colors and line styles correspond to those used in Fig. 1). *Left:* the results from the $(2, 2, 0)$ fundamental-mode-only analysis (i.e. $N = 0$). *Right:* the results from the overtone analysis including the $(2, 2, 0)$ and $(2, 2, 1)$ modes (i.e. $N = 1$). Each line corresponds to a different choice of t_0 prior. Colored lines correspond to Gaussians with widths of $1\tilde{M}_f$ and means $\bar{t}_0$ (see Fig. 1). The dashed black line corresponds to using the posterior on time of peak strain (from a full IMR analysis) as our prior, which marginalizes over uncertainty on the time of peak strain. Also shown for reference (dotted line) is the posterior from a full IMR analysis. The main panel shows the 90% confidence contours while the side panels show the one-dimensional marginalized posteriors.

time, t_0. We emphasize that this is achieved here by using the posterior on the time of peak strain as a prior on t_0; this is motivated by the observations of Refs. [17–26] described above, which show that generically the ringdown can be considered to start at around this time.

III. RESULTS

There are several ways to investigate and quantify the evidence for additional QNMs in the ringdown. Sec. III A contains the results of a series of analyses designed to study the presence of a possible overtone in GW150914. Sec. III B contains the results of a series of analyses designed to test whether or not what has been detected really is an overtone and is not the accumulation of other effects. Sec. III C describes further checks on the stability of the results, and Sec. III D contains some additional results that further demonstrate the capabilities of the frequency-domain approach to ringdown analysis.

Throughout Secs. III A and III B, we compare our results with those in Refs. [27] and [28]. This is done in the hope of helping to resolve the controversy over the evidence for a ringdown overtone in GW150914. However, it should be stressed that our results are produced

using a very different method and care should therefore be taken in making direct comparisons. Although the frequency-domain analysis is formally equivalent to the time-domain analysis in a particular limit (as discussed in the introduction, and in more detail in Ref. [30]) we do not take this limit in a practical analysis. Furthermore, the frequency-domain analysis is further generalized with respect to the time-domain analysis in that it marginalizes over parameters such as the sky position and ringdown start time (which are fixed in the analyses of Refs. [27, 28]). Results from our frequency-domain analyses should therefore not be expected to agree perfectly with those from previous time-domain analyses.

A. Presence of an overtone

In order to investigate the presence of an overtone in the GW150914 ringdown, we initially perform two analyses using the model described in Sec. II A: one analysis uses only the fundamental QNM ($N = 0$) and the other includes the first overtone ($N = 1$). Aside from the inclusion of the overtone in the ringdown (which introduces two additional parameters: an amplitude and a phase), these two analyses are otherwise identical.
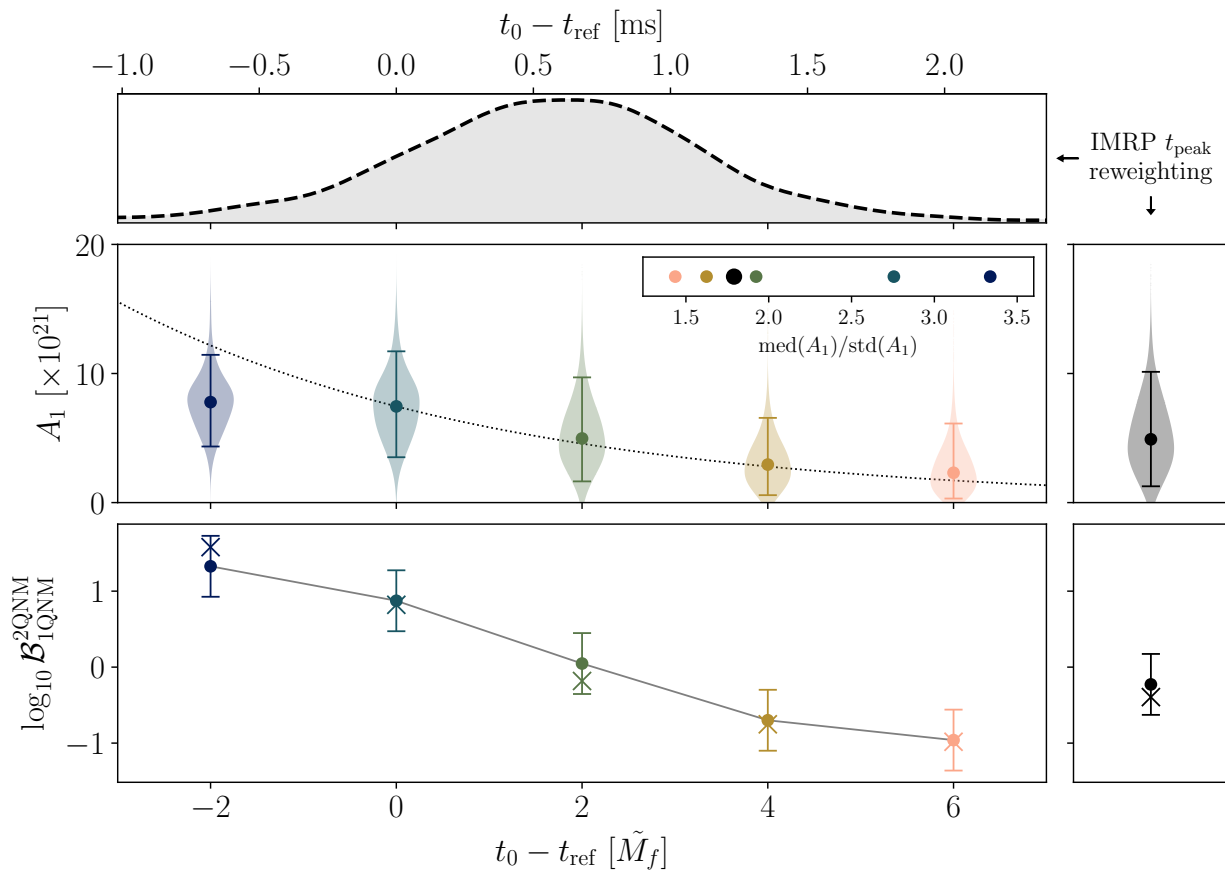
FIG. 3. Posteriors on the overtone amplitude, and Bayes' factors in favor of the overtone model for different choices of t_0 prior (the colors and line styles correspond to those used in Fig. 1). *Top:* Posterior on the time of peak strain in the Hanford frame, from a IMRPHENOMPV2 analysis, as in Fig. 1 (and originally from Ref. [28]). *Middle:* Overtone amplitude posteriors for different choices of t_0 prior. The left panel corresponds to Gaussian priors with standard deviation $1\tilde{M}_f$, centered at the time they are plotted. The dotted line indicates the expected exponential decay of the A_1 mode; this is included merely to guide the eye and was produced using the median mass and spin values from the full IMR analysis and the median value of A_1 from the $\tilde{t}_0 = 0$ prior. The right panel corresponds to using the IMRPHENOMPV2 time of peak strain as a prior. For earlier start times the posteriors on the amplitude are peaked further away from zero; this is quantified in the inset plot where the ratio of the median to the standard deviation of the A_1 posterior is plotted. *Bottom:* the Bayes' factor in favor of the overtone model for each prior choice; circles with error bars show the Bayes' factor calculated from nested sampling (with errors estimated by the sampler) while the crosses show the results calculated using the Savage-Dickey density ratio.

In Fig. 2 we plot the posterior distributions on the remnant BH mass, M_f, and dimensionless spin, χ_f, for both of these analyses. Results are shown for the different choices of the prior on the ringdown start time shown in Fig. 1 (these results were obtained by reweighting the samples obtained with a flat prior using the approach described in Sec. II C). The earliest start time ($\tilde{t}_0 = -2\tilde{M}_f$) is omitted from the fundamental-only ($N = 0$) plot in the left-hand panel of Fig. 2 because of a low number of posterior samples at these time (see appendix B). Also shown for comparison are the much tighter constraints resulting from the full IMR analysis. These IMR posterior samples were obtained from Ref. [42], which (as detailed in Refs. [9, 28]) are obtained from applying fitting formulas to the samples available at Ref. [43]. When only the fundamental QNM is used ($N = 0$), and when the anal-

ysis is started at early times (e.g. $t_0 - t_{\rm ref} \lesssim -2\tilde{M}_f$) our posteriors on the remnant parameters are biased towards high values of M_f and χ_f. This behavior is expected; a single QNM is only able to model the ringdown signal starting well after the time of peak strain. Including an overtone ($N = 1$) allows the ringdown analysis to start at earlier times, as can be seen by the removal of the bias in the right panel. This improvement is suggestive that the data supports the inclusion of an overtone. Additionally, using an earlier ringdown start time increases the SNR in the ringdown and reduces the posterior width; this effect can be seen in both the $N = 0$ and $N = 1$ analyses.

Our results in Fig. 2 can be compared to the corresponding results of the time-domain analyses shown in Fig. 1 from Cotesta et al. [27] and Figs. 4 and 5 from Isi & Farr [28]. In general terms, there is broad agreement

between all three sets of results. In particular, all three sets of authors find that the overtone analyses ($N = 1$) always gives results that are more consistent with the IMR result and get increasingly broader for later choices of the ringdown start time. All sets of authors also find that for the fundamental-only analysis ($N = 0$) starting at early times (i.e. $t_0 - t_{\rm ref} \lesssim 0$) leads to posteriors that are inconsistent with the IMR result. However, there are subtle differences between the various results. Our results with $N = 0$ and early start times gives posteriors biased to large values of M_f and χ_f; this is also seen in Ref. [28], but not in Ref. [27] (where the posterior consistently reaches lower values of χ_f). Our results with $N = 0$ and late start times (i.e. $t_0 - t_{\rm ref} \gtrsim 4\tilde{M}_f$) are partially consistent with the IMR results; this is also seen in Ref. [27], but not in Ref. [28] who never find consistency with the IMR result for any choice of start time. Finally, when including the overtone ($N = 1$) and starting at late times, Ref. [27] find results that are consistent with $\chi_f = 0$ (i.e. a Schwarzschild BH) at 90% confidence, in stark disagreement with Ref. [28] who find $\chi_f \gtrsim 0.2$. Our results are in better agreement with those of Ref. [28].

In the middle panel of Fig. 3 we investigate our $N = 1$ overtone analysis further by plotting the 1-dimensional marginalized posteriors for the amplitude, A_1, of the QNM overtone. An amplitude posterior peaked away from zero has been suggested (particularly by Ref. [9]) as one good indication for the presence of an overtone in the data. As expected, the QNM overtone decays quickly and when starting at later times we find a small value for the amplitude. The degree to which the A_1 posterior is peaked away from zero can be quantified using the ratio between the median and standard deviation; this is plotted in the inset of the middle panel of Fig. 3. For values of $\bar{t}_0$ between $-2\tilde{M}_f$ and $+6\tilde{M}_f$, we find posteriors on A_1 that are peaked away from zero at between 1.44 and 3.34σ. If we reweight using the IMRP $t_{\rm peak}$ prior, we find a posterior peaked away from zero at 1.79σ.

Our results in the middle panel of Fig. 3 can be compared to the corresponding results of the time-domain analyses shown in Fig. 1 of Ref. [28] and Fig. 2 of Ref. [27]. All three sets of authors find values of A_1 that are smaller at later times, consistent with the expected exponential decay of the overtone, but they disagree on the absolute value of the amplitude and the significance with which a zero amplitude can be excluded. Refs. [9, 28] find the largest values; they report a posterior peaked 3.6σ away from zero. Ref. [27] finds much smaller values which are consistent with zero for many choices of start time. These analyses use essentially the same method and should therefore agree exactly. Our result, produced using a different method, lies somewhere in between; we do find non-zero values are preferred for a range of start times, but only with a modest significance of $\sim 1.79\sigma$ for our preferred IMRP $t_{\rm peak}$ prior which we consider to be the best description of our uncertainty on the ringdown start time.

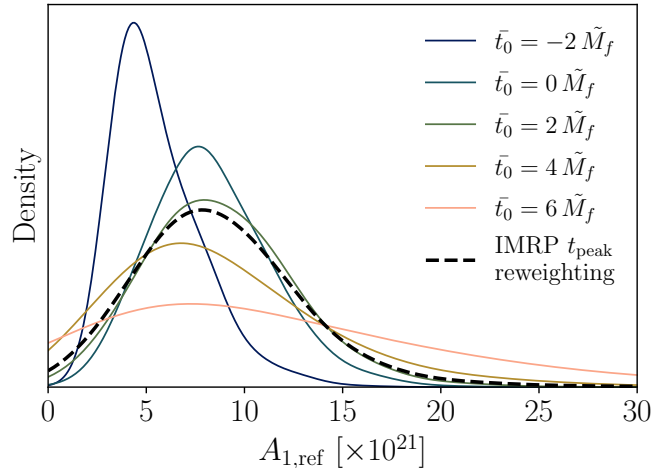The comparison of our results with those of Refs. [9, 27,

FIG. 4. Posteriors on the overtone amplitude from our $N = 1$ overtone analysis, re-scaled to a fixed reference time of $t_{\rm ref}$. The re-scaling does not significantly affect the significance with which the posteriors are peaked away from zero. The colors and line styles indicate the prior used on t_0 and correspond to those used in Fig. 1.

28] is complicated by the fact that we use subtly different definitions for the amplitude. The time-domain analyses naturally define the mode amplitudes at a fixed time, usually t_0. Our frequency-domain analysis also defines the mode amplitudes at t_0, but this start time is then varied as part of the analysis, blurring the exact time at which the amplitude is defined. This is a fairly small effect for the narrow Gaussian priors, but more significant for the wider IMRP $t_{\rm peak}$ prior. We can correct for this effect by re-scaling all the overtone amplitudes to any fixed reference time (here we use $t_{\rm ref}$) using the known decay rate for the QNMs;

$$A_{1,\rm ref} = A_1 \exp\left(\frac{t_0 - t_{\rm ref}}{\tau_{221}(M_f, \chi_f)}\right), \qquad (7)$$

where $\tau_{221}(M_f, \chi_f)$ is the exponential decay time of the $(2, 2, 1)$ QNM and is a function of the remnant mass and spin. This re-scaling can be done for any QNM and the resulting amplitude parameters $A_{\ell mn,\rm ref}$ are more directly comparable with the amplitudes used in time-domain analyses. Posteriors on $A_{1,\rm ref}$ are shown in Fig. 4.

In the bottom panel of Fig. 3 we plot the Bayes' factors between the fundamental only ($N = 0$) and overtone ($N = 1$) analyses. This is defined as $\mathcal{B}^{\rm 2QNM}_{\rm 1QNM} = Z_{N=1}/Z_{N=0}$. The Bayes' factor has been suggested (particularly by Ref. [27]) as another good way for quantifying the support for an overtone in the data. The Bayes' factor was computed in two different ways. Firstly, DYNESTY was used to calculate the evidences $Z_{N=0}$ and $Z_{N=1}$ for both the analyses described above, and these were reweighted to the desired t_0 prior using Eq. 6. Nested sampling also returns an estimate for the error on the evidences, and these are used to plot the error
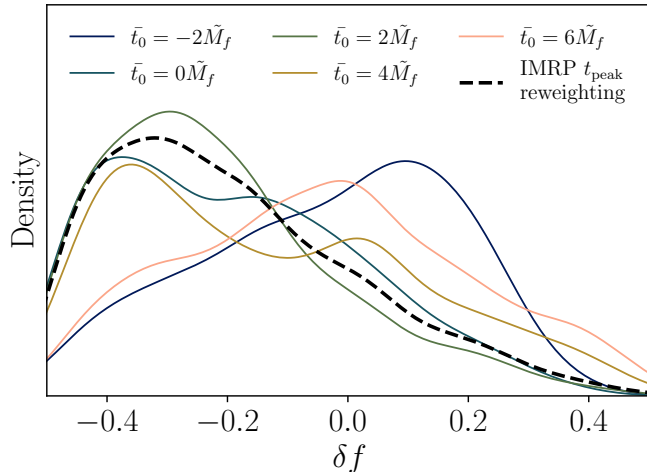
FIG. 5. Posteriors on the deviation parameter from the Kerr value for the real part of the overtone frequency. The colors and line styles distinguish different choices for the t_0 prior and correspond to those used in Fig. 1. The mode frequency is given by $f_{221}^{\mathrm{Kerr}} \exp(\delta f)$, so that $\delta_f = 0$ is the expected result for the Kerr metric. For all choices of t_0 prior the data is consistent with $\delta f = 0$.

bars in Fig. 3. Secondly, exploiting the fact that the $N = 0$ model is nested within the $N = 1$ model, the Bayes' factors were computed using the posterior on A_1 from the $N = 1$ analysis to find the Savage-Dickey density ratio [44].

Our results in the bottom panel of Fig. 3 can be compared to the corresponding results of the time-domain analyses shown in Fig. 7 of Ref. [28] and Fig. 2 of Ref. [27]. Ref. [27] computes the Bayes' factors using the ratio of evidences evaluated with nested sampling, whereas Ref. [28] computes Bayes' factors using Savage-Dickey density ratios. All sets of authors find Bayes' factors that decrease for later ringdown start times, although they disagree on the exact value. Ref. [28] finds the strongest log-evidence of ~ 1.7 at $t_0 - t_{\mathrm{ref}} \sim 0$. Ref. [27] finds slightly *negative* log-evidence starting at this time. Again, our result lies somewhere in between, we find a moderate log-evidence of ~ 1.0 when marginalizing over a narrow prior on t_0 centered at this time. If we instead marginalize over the time of peak strain using the broader IMRP t_{peak} prior, the evidence is slightly negative. However, as discussed in Sec. IV below, we consider the actual values of the Bayes factors to be less important than their trend with varying start time.

B. The nature of the overtone

The results of the previous section show that there is tentative evidence for something beyond the fundamental $(2, 2, 0)$ mode in the GW150914 data. In the previous section it was assumed that this is the $(2, 2, 1)$ QNM overtone; this is motivated by our expectations from numer-

ical relativity experiments (see, for example, Ref. [17]). In this section, we address this assumption by measuring the frequency and amplitude of the QNM overtone and comparing with the expectations from GR.

Fig. 5 shows the results of a third ringdown analysis that also includes two QNMs. In this analysis the complex frequency of the second QNM is allowed to deviate from the Kerr overtone value. This differs from the $N = 1$ overtone analysis described above, where the frequency of the overtone was fixed by the remnant mass and spin to the Kerr value, $\omega_{221} = 2\pi f_{221} - i/\tau_{221}$. Recovering a value of δf consistent with zero has been suggested (particularly by Ref. [28]) as further evidence for the presence of an overtone; otherwise, it might be expected that the extra parameters would fit to the noise and would not recover the Kerr value. We use the parameterization from Ref. [28]; the complex frequency of the second QNM is now $\omega_{221} = 2\pi f - i/\tau$, where $f = f_{221} \exp(\delta f)$ and $\tau = \tau_{221} \exp(\delta \tau)$. This introduces the two new dimensionless parameters δf and $\delta \tau$ into the model, for which we use uniform priors in the range $[-0.5, 0.5]$. The $\delta \tau$ parameter is not well constrained, therefore we focus initially on δf. We find posteriors on δf consistent with zero for all choices of t_0 prior with standard deviations ~ 0.2. This is consistent with what was found in Ref. [9] and can be viewed as a test of the no-hair theorem at the $\sim 20\%$ level.

Our results in Fig. 5 can be compared with Fig. 2 of Ref. [28]. Our preferred run, using the IMRP t_{peak} prior on t_0, is broadly consistent with that result. However, what is notable about our results is that we do not find
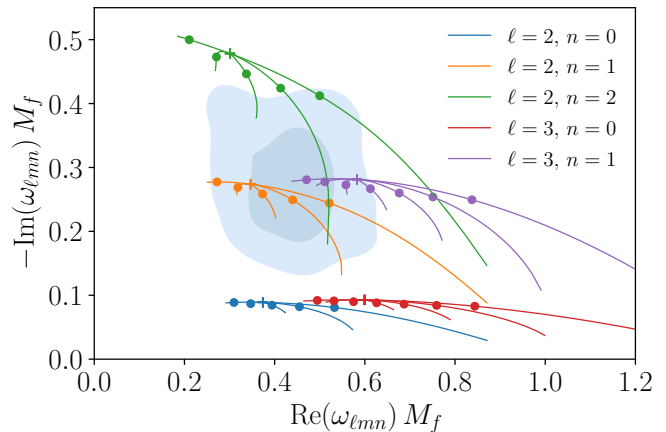

FIG. 6. The posterior on the dimensionless complex frequency of the second QNM (50% and 90% regions), assuming the first is the fundamental $(\ell, |m|, n) = (2, 2, 0)$ mode. Lines indicate the Kerr frequencies parameterized by the remnant spin; dots and crosses indicate points with $\chi_f = 0.7$ and 0 respectively. Lines are colored according to their ℓ and n indices and the m index increases left to right in each set. The frequency of the second QNM is consistent with the expected $(2, 2, 1)$ overtone, but also with several other modes. However, all fundamental modes (those with $n = 0$) are excluded.
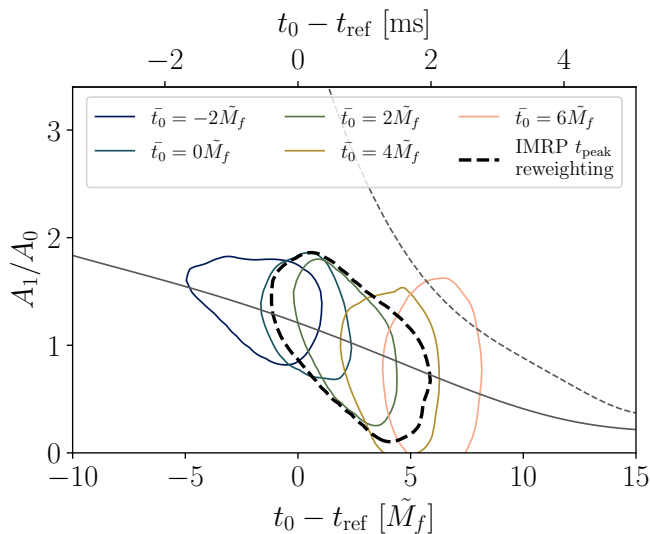
FIG. 7. Posteriors on the amplitude ratio A_1/A_0 from our $N = 1$ overtone analysis. The 90% contours are plotted, with the colors and line styles indicating the t_0 prior and correspond to those used in Fig. 1. The solid gray curve shows the results of a two-QNM fit to the numerical relativity simulation SXS:BBH:0305 which has parameters consistent with GW150914. The dashed gray curve shows the results of a multi-QNM fit to SXS:BBH:0305 which follows closely the expected exponential decay rate for the amplitude ratio.

a significant broadening of the posterior for later choices of the start time. This was found by Ref. [28] and would be expected if an overtone was present, as both the overtone amplitude and ringdown SNR decaying with later ringdown start times.

To investigate this further, we use the results of the ringdown analysis where the frequency of the overtone is allowed to vary freely to address another important question. If the data does indeed contain a second QNM, can we determine which mode it is? Theoretical studies of numerical relativity simulations suggest that the $(\ell, |m|, n) = (2, 2, 1)$ will be the next most prominent, especially for early start times [17]. In Fig. 6 we plot the posterior on the *dimensionless* complex frequency (allowing both the real and imaginary parts to vary freely) of the second QNM, ωM_f. This plot uses the value for M_f calculated from the complex frequency of the first QNM, assuming this is the expected $(2, 2, 0)$ fundamental mode of Kerr. We find that we can confidently conclude that the second mode is an overtone ($n \geq 1$) but that it is not possible to say from the data alone exactly which overtone. For example, the modes $(2, 2, 1)$ and $(2, 1, 1)$ are both equally compatible with the data. In general, when searching for additional QNMs it is necessary to be guided by our prior expectations regarding which modes are expected to be excited with the highest amplitudes.

We now turn our attention to the measured amplitude A_1 and whether this matches the theoretical expectations for the $(2, 2, 1)$ overtone. For convenience, we choose to

work with the amplitude ratio A_1/A_0 which eliminates factors common to all modes, such as the distance to the source. Because the two QNMs decay exponentially at different rates, the amplitude ratio depends strongly on the chosen ringdown start time. Our two-dimensional posteriors on the amplitude ratio and ringdown start time are plotted in Fig. 7. As expected we find that the amplitude ratio decreases for later start times, and the error on the amplitude ratio increases for later start times because of the decreasing SNR in the ringdown.

In order to check whether this is consistent with the theoretical expectation for an overtone we compare with fits to the numerical relativity simulation SXS:BBH:0305 [45, 46] which has parameters consistent with GW150914. Fixing the remnant mass and spin to the values reported in the simulation metadata, we perform QNM least-squares fits to this simulation for a range of ringdown start times using the code previously developed in Ref. [23]. Results are shown in Fig. 7 for two such fits. Firstly, we performed a two-QNM fit intended to mimic the analysis of the real GW150914 data described in Sec. III A above. In this analysis the $_{-2}Y_{22}$ spherical harmonic mode of the simulation is modeled as a sum of the $(2, 2, 0)$ and $(2, 2, 1)$ QNMs and the amplitude ratio is recorded. The results from this two-QNM fit agree very well with what is seen in the real data giving us further confidence that there is nothing unexpected present in the data and that our results are not unduly affected by noise fluctuations (see also the discussion in Sec. III C). Secondly, we perform a full multi-QNM fit to all the spherical harmonic modes (up to and including $\ell = 8$) with a ringdown model that includes all QNMs (including both prograde and retrograde modes) up to $\ell = 8$ and $n = 7$ (1232 QNMs in total). The ratio of the amplitudes of the $(2, 2, 0)$ and $(2, 2, 1)$ prograde modes from this fit behaves very differently; the ratio follows very closely a exponential time evolution which can be understood in terms of the difference between the two QNM decay times.

The fact that the two-QNM analysis gives a very different amplitude ratio compared to the full multi-QNM analysis for ringdown start times near the peak strain is related to the extreme destructive interference observed in the QNM overtone fits of Refs. [17–26] with large values of N. This shows that the amplitude A_1 recovered from a two-QNM analysis is not purely the amplitude of the first overtone but also includes significant contributions from higher overtones and other harmonics. However, absorbing these contributions into the first overtone introduces a systematic bias in the remnant properties that is smaller than the statistical uncertainty; this can be seen in, for example, Fig. 2 and Sec. IV C of Ref. [17]. For this reason, it still makes sense to describe the results of the two-QNM analysis as a measurement of the overtone, even though there are undoubtedly other contributions present in the signal.

C. The effect of noise and sampling rates

One of the key claims made in Ref. [27] was the overtone detection was highly sensitive to noise fluctuations. This was disputed by Ref. [28]. In order to address this issue, we performed a noise injection study mirroring closely what was done in Ref. [27]. The results of this injection study are presented in appendix C. As expected, the results of injecting into different noise realizations show some scatter. However, this scatter is not larger than expected and we are unable to reproduce the claim in Ref. [27] with our (very different) analysis method.

It has been suggested [47] that the results of ringdown analyses, in particular those including highly damped overtones that contain significant power at high frequencies, might be sensitive to aliasing effects when using downsampled strain data. This is discussed in more detail in appendix A where it is argued that our results are insensitive to changes in both the sampling rate of the data and the value of f_{high}.

D. Other results

One important benefit of the frequency-domain approach to ringdown data analysis introduced in Ref. [30] and used here is that it naturally allows us to search (and hence to numerically marginalize) over source sky position and ringdown start time. This should be contrasted with the treatment of these parameters in most time-domain analyses where these parameters are fixed, potentially biasing the results. (Although it is technically possible to search over the sky in a time-domain analysis [8, 11], this is rarely done in practice.) To emphasize this, we plot the posterior on the sky location of GW150914 from our $N = 1$ overtone analysis reweighted to the IMRP t_{peak} prior on the ringdown start time. This can be compared with the publicly available LIGO sky posterior for GW150914 obtained using sample from [48]. This is shown in Fig. 8. As discussed in [30], it should be emphasized that this sky posterior is not a ringdown-only result because much of the information is also coming from the wavelets used to model the inspiral-merger portion of the signal.

Because the inspiral and merger parts of the signal are being modeled using truncated wavelets as part of the frequency-domain ringdown analysis, this allows us to plot a full waveform reconstruction from our results. This reconstruction is shown in Fig. 9 for our $N = 1$ overtone analysis reweighted to the IMRP t_{peak} prior. The full waveform model used in our analysis is discontinuous at t_0. However, as discussed in Ref. [30], the whitened waveform reconstruction plotted here is smooth; this is a result of marginalizing over the location of the discontinuity at t_0, the waveform model "learning" the continuity from the data, and the whitening process used to make the figure. This waveform reconstruction uses the posterior on all of the model parameters, including those for
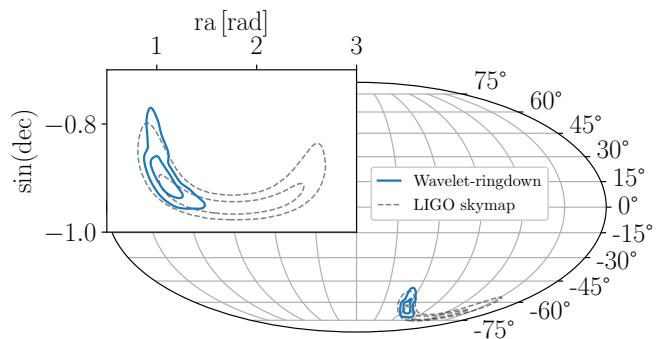

FIG. 8. Posterior on the source sky position using geocentric coordinates in Mollweide projection. Shown in blue is the results from the $N = 1$ overtone analysis using the IMRP t_{peak} reweighting for the prior on the ringdown start time. The LIGO skymap for this event is shown by the dashed black line for comparison. The inset plot shows a zoomed in map plotted using right ascension and the sine of the declination. In both cases, 50% and 90% contours are plotted.

the wavelets; more details on these parameters are given in appendix D.

IV. DISCUSSION AND CONCLUSIONS

The main motivation for this work comes from the ongoing discussion in the literature about whether a ringdown overtone can be confidently detected in the GW150914 data. In particular, the detection claim made in Ref. [9] was disputed by Ref. [27] where a nearly identical time-domain analysis was re-performed (see also the reply Ref. [28]). Applying the frequency-domain ringdown analysis originally presented in Ref. [30], we contribute to this discussion with a thorough reanalysis of the GW150914 data. This includes performing analyses with and without an overtone while considering different ringdown start times, as well as performing a noise injection study and studying the effects of different data sampling rates and frequency integration limits on our results. Although the method used here differs significantly from previous time-domain analyses, we present our results in a way that makes it as easy as possible to compare with earlier work. In conclusion, we do find tentative evidence for a ringdown overtone, but not at the high level of significance originally claimed in Ref. [9].

In order to be more quantitative, it is first necessary to be able to say clearly what it even means to "detect a overtone". Although intuitively obvious, it is not clear how to make this notion precise (this issue has previously been discussed in Ref. [28]). Several approaches have been suggested: looking to see if including the overtone improves the posterior on the remnant parameters (see Fig. 2); looking at the posterior on the overtone amplitude for a range of start times (see middle panel of Fig. 3); computing the Bayes' factor in favor of an overtone (see
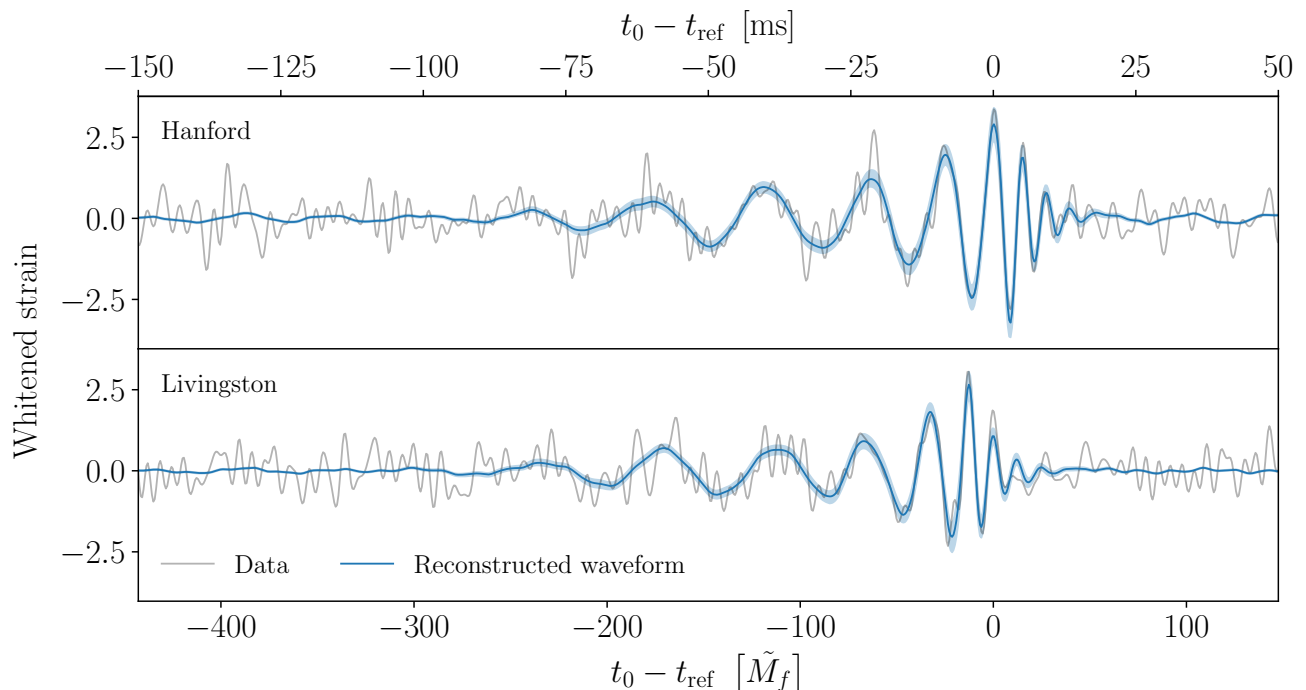
FIG. 9. Posterior on the reconstructed whitened waveform. Shown in gray is the strain data from both LIGO interferometers (*top*: Hanford, *bottom*: Livingston) whitened according to the noise amplitude spectral density in the detector and bandpass filtered between 32 and 512 Hz for clarity. Shown in blue is the waveform reconstruction from the $N = 1$ overtone analysis with the IMRP t_{peak} reweighting for the prior on the ringdown start time. The blue lines and shaded regions indicate median and the 90% credible interval. The signal is plotted as a function of time from t_{ref} using both SI and natural units on the upper and lower x-axis respectively.

bottom panel of Fig. 3); and allowing the frequency of the second QNM to vary freely to see if the data prefers, or at least is consistent with, the expected Kerr value (see Figs. 5 and 6). Although these are not all independent from one another, they all help shed light on which QNMs are present. The results of all of these tests can also be compared to results from a noise injection study.

As well as not being completely independent of each other, none of these tests are, by themselves, sufficient to justify a claim of a detection. For example, one issue that has been raised is that the Bayes' factor can be made to take any value with a suitable adjustment to the prior range. There are also conceptual problems regarding what it means to compare two models, neither of which is expected to fully describe the data. Here we are comparing the fundamental-only mode model (with a single QNM) to the overtone model (with two QNMs) when our firm prior belief is that the true signal should contain an infinite number of QNMs plus additional corrections (e.g. from nonlinearities in the merger, tails, and memory effects).

From the above discussion, it is clear that ringdown analyses are rather subtle. We think our frequency-domain method has some important advantages over what has been done before. For example, it marginalizes over the ringdown start time and sky position which

is preferable to fixing these parameters (which potentially introduces systematic biases). Ideally, we should also marginalize over the uncertainties in the noise power spectral density (see, e.g., Ref. [49]) and detector calibration (see, e.g., Ref. [50]) as part of a ringdown analysis. The ability to do this is, in principle, another benefit of the frequency-domain analysis approach used here as this can be done using techniques that are standard in the field.

We stress that while our results have been compared with those of previous time-domain studies, our frequency-domain method is rather different and therefore we do not expect to find perfect agreement. In contrast, the results of Refs. [9, 27, 28] are produced using essentially identical methods and should therefore be expected to agree exactly. The reason for the disagreement that is seen there is currently unknown and the subject of an ongoing investigation by both sets of authors. It is vitally important for QNM science that all results are reproducible. To that end we have made all our data products and plotting scripts publicly available at Ref. [31].

If QNMs are going to fulfill their promise for testing GR, fundamental physics and the Kerr metric hypothesis, then the community must be able to agree on standards for what it means to detect them and to be able to robustly quantify their significance. This field is still

very young, and that there is already significant controversy regarding the QNM content of GW150914 and GW190521 is concerning, and we risk the situation becoming more confused with many more suitable events expected in O4. And, as discussed in the introduction, this is a conceptual issue that will not be resolved with more observations, even at higher SNRs. This issue needs input from the whole community; however, we suggest that (as a minimum) future claims of an overtone detection are accompanied by the investigations in Fig. 2, both panels of Fig. 3 and Fig. 5. That is, posteriors on the remnant properties with and without the overtone, posteriors on the overtone amplitude, a study of the Bayes' factor trends for different start times, and posteriors on deviations from Kerr when the overtone frequency is allowed to vary.

[1] O. Dreyer, B. J. Kelly, B. Krishnan, L. S. Finn, D. Garrison, and R. Lopez-Aleman, Class. Quant. Grav. **21**, 787 (2004), arXiv:gr-qc/0309007.

[2] R. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. D **103**, 122002 (2021), arXiv:2010.14529 [gr-qc].

[3] R. Abbott *et al.* (LIGO Scientific, VIRGO, KAGRA), (2021), arXiv:2112.06861 [gr-qc].

[4] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. Lett. **116**, 061102 (2016), arXiv:1602.03837 [gr-qc].

[5] J. Aasi *et al.* (LIGO Scientific), Class. Quant. Grav. **32**, 074001 (2015), arXiv:1411.4547 [gr-qc].

[6] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. Lett. **116**, 221101 (2016), [Erratum: Phys.Rev.Lett. 121, 129902 (2018)], arXiv:1602.03841 [gr-qc].

[7] P. Whittle, Journal of the Royal Statistical Society. Series B (Methodological) **19**, 38 (1957).

[8] G. Carullo, W. Del Pozzo, and J. Veitch, Phys. Rev. D **99**, 123029 (2019), [Erratum: Phys.Rev.D 100, 089903 (2019)], arXiv:1902.07527 [gr-qc].

[9] M. Isi, M. Giesler, W. M. Farr, M. A. Scheel, and S. A. Teukolsky, Phys. Rev. Lett. **123**, 111102 (2019), arXiv:1905.00869 [gr-qc].

[10] 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].

[11] M. Isi and W. M. Farr, (2021), arXiv:2107.05609 [gr-qc].

[12] E. Berti, V. Cardoso, and C. M. Will, Phys. Rev. D **73**, 064030 (2006), arXiv:gr-qc/0512160.

[13] S. Gossan, J. Veitch, and B. S. Sathyaprakash, Phys. Rev. D **85**, 124056 (2012), arXiv:1111.5819 [gr-qc].

[14] R. Brito, A. Buonanno, and V. Raymond, Phys. Rev. D **98**, 084038 (2018), arXiv:1805.00293 [gr-qc].

[15] M. Isi, W. M. Farr, M. Giesler, M. A. Scheel, and S. A. Teukolsky, Phys. Rev. Lett. **127**, 011103 (2021), arXiv:2012.04486 [gr-qc].

[16] R. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. Lett. **125**, 101102 (2020), arXiv:2009.01075 [gr-qc].

[17] M. Giesler, M. Isi, M. A. Scheel, and S. Teukolsky, Phys. Rev. X **9**, 041060 (2019), arXiv:1903.08284 [gr-qc].

[18] S. Bhagwat, X. J. Forteza, P. Pani, and V. Ferrari, Phys. Rev. D **101**, 044033 (2020), arXiv:1910.08708 [gr-qc].

[19] I. Ota and C. Chirenti, Phys. Rev. D **101**, 104005 (2020), arXiv:1911.00440 [gr-qc].

[20] G. B. Cook, Phys. Rev. D **102**, 024027 (2020), arXiv:2004.08347 [gr-qc].

[21] X. Jiménez Forteza, S. Bhagwat, P. Pani, and V. Ferrari, Phys. Rev. D **102**, 044053 (2020), arXiv:2005.03260 [gr-qc].

[22] A. Dhani, Phys. Rev. D **103**, 104048 (2021), arXiv:2010.08602 [gr-qc].

[23] E. Finch and C. J. Moore, Phys. Rev. D **103**, 084048 (2021), arXiv:2102.07794 [gr-qc].

[24] X. J. Forteza and P. Mourier, Phys. Rev. D **104**, 124072 (2021), arXiv:2107.11829 [gr-qc].

[25] A. Dhani and B. S. Sathyaprakash, (2021), arXiv:2107.14195 [gr-qc].

[26] L. Magaña Zertuche *et al.*, Phys. Rev. D **105**, 104015 (2022), arXiv:2110.15922 [gr-qc].

[27] R. Cotesta, G. Carullo, E. Berti, and V. Cardoso, (2022), arXiv:2201.00822 [gr-qc].

[28] M. Isi and W. M. Farr, (2022), arXiv:2202.02941 [gr-qc].

[29] J. C. Bustillo, P. D. Lasky, and E. Thrane, Phys. Rev. D **103**, 024041 (2021), arXiv:2010.01857 [gr-qc].

[30] E. Finch and C. J. Moore, Phys. Rev. D **104**, 123034 (2021), arXiv:2108.09344 [gr-qc].

[31] E. Finch and C. J. Moore, "Data release for "Searching for a Ringdown Overtone in GW150914,"" https://doi.org/10.5281/zenodo.6554513 (2022).

[32] LIGO Scientific and Virgo Collaboration, "Gravitational wave open science center," (2022).

[33] R. Abbott *et al.*, SoftwareX **13**, 100658 (2021), arXiv:1912.11716 [gr-qc].

[34] LIGO Scientific and Virgo Collaboration, "Power Spectral Densities (PSD) release for GWTC-1: LIGO document P1900011-v1," (2019).

[35] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. X **9**, 031040 (2019), arXiv:1811.12907 [astro-ph.HE].

[36] R. Buscicchio, E. Roebber, J. M. Goldstein, and C. J. Moore, Phys. Rev. D **100**, 084041 (2019), arXiv:1907.11631 [astro-ph.IM].

[37] J. Skilling, Bayesian Analysis **1**, 833 (2006).

[38] J. S. Speagle, Mon. Not. Roy. Astron. Soc. **493**, 3132 (2020), arXiv:1904.02180 [astro-ph.IM].

[39] C. Robert and G. Casella, *Monte Carlo statistical methods*, Springer texts in statistics (Springer, 2013).

[40] E. Payne, C. Talbot, and E. Thrane, Phys. Rev. D **100**, 123017 (2019), arXiv:1905.05477 [astro-ph.IM].

[41] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, Phys. Rev. Lett. **113**, 151101 (2014), arXiv:1308.3271 [gr-qc].

[42] M. Isi and W. M. Farr, "Data release for "revisiting the ringdown of gw150914"," (2022).

[43] LIGO Scientific and Virgo Collaboration, "Parameter estimation sample release for gwtc-1r," (2019).

[44] J. M. Dickey, The Annals of Mathematical Statistics **42**, 204 (1971).

[45] G. Lovelace *et al.*, Class. Quantum Grav. **33**, 244002 (2016), arXiv:1607.05377 [gr-qc].

[46] SXS collaboration, https://data.black-holes.org/waveforms/index.html.

[47] W. M. Farr and M. Isi, "private communication," (2022).

[48] The LIGO Scientific and Virgo Collaborations, "Data release for "population properties of compact objects from the second ligo-virgo gravitational-wave transient catalog"," (2021).

[49] N. J. Cornish, T. B. Littenberg, B. Bécsy, K. Chatziioannou, J. A. Clark, S. Ghonge, and M. Millhouse, Phys. Rev. D **103**, 044006 (2021), arXiv:2011.09494 [gr-qc].

[50] C. Cahillane *et al.*, Phys. Rev. D **96**, 102001 (2017), arXiv:1708.03023 [astro-ph.IM].

[51] F. Crameri, "Scientific colour maps," (2021).

Appendix A: Sampling Frequency and Integration Limits

Overtones ($n \geq 1$) have roughly the same real part of the frequency as the corresponding (i.e. with the same ℓ and m) fundamental model, but they have a shorter damping time. See, for example, Fig. 6. This means that if a single, isolated, mode is viewed in Fourier space, the power spectrum is broader and contains significant power at higher frequencies.

It has been suggested [47] that overtone studies using highly downsampled data might suffer from aliasing effects due to the reduced Nyquist frequency. Early ringdown studies, including those in Refs. [6, 9, 28], generally used strain data that had been downsampled to 2 kHz. This was done for convenience and computational speed and was not originally anticipated to be a problem because the merger of GW150914 occurs at ~ 200 Hz, safely below the Nyquist frequency.

In this paper the 4 kHz data is used for all the analyses in the main text. Additionally, a frequency-domain log-likelihood with an upper integration limit of $f_{\text{high}} = 1000$ Hz was used. Our method is very different from the time-domain analyses, and do not expect our results to be sensitive to small changes in these choices. To check that this is the case we have repeated the $N = 1$ overtone analysis using the 16 kHz sampled data (obtained from [32]) and we find no significant changes in our results. Using reweighting techniques (this time applied to the likelihood) we have also investigated the effect of changing the upper limit of integration in the likelihood. By re-evaluating the $N = 1$ posterior chain on a likelihood with $f_{\text{high}} = 1500$ Hz and 2000 Hz, and then reweighting, we again find no significant changes in our results.

Appendix B: Posteriors on the ringdown start time

As described in Sec. II C, we initially perform Bayesian inference on the ringdown using a broad, flat prior on the ringdown start time parameter t_0. The posterior on t_0 from both the $N = 0$ and $N = 1$ analyses are shown in Fig. 12. We do not consider these posteriors to be physically meaningful results because they were obtained with a prior that does not correctly describe our state of knowledge about when the ringdown should start. These results are produced merely as an intermediate step in our analysis, before the reweighting was applied, and are shown here only to further illustrate the reweighting procedure described in Sec. II C.

In order for the subsequent reweighting step to be accurate, it is necessary for the posterior chains (particularly for the $N = 1$ overtone analysis) to contain samples across the range of start times that we consider. For this reason, the DYNESTY sampler settings described in Sec. II B were chosen to ensure a large number of posterior samples were produced; we obtained 203697 and 218882 posterior samples from the $N = 0$ and $N = 1$
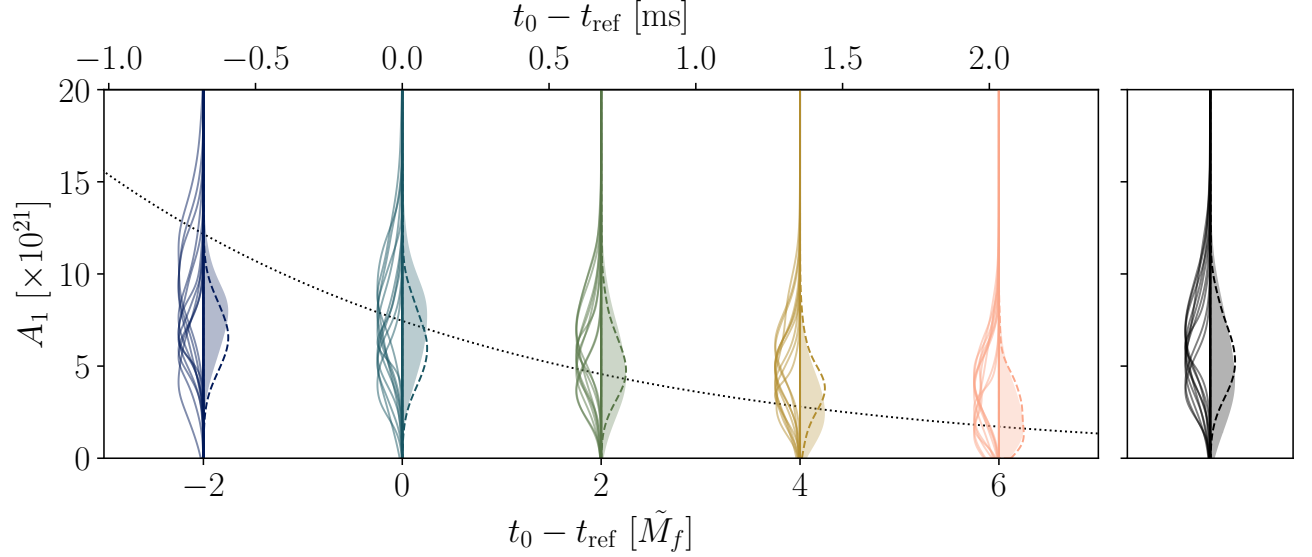
FIG. 10. This is similar to the middle panel of Fig. 3 in the main text, but shows the posteriors on the overtone amplitude from the noise injection study. The different violin plots are for the different priors on the ringdown start time and the colors are the same as those used in Figs. 1 and 3. On the right-hand side of each set of violin plots, the filled posterior shows the result obtained using the real GW150914 data (this is the same as what is plotted in Fig. 3). On the left-hand side are all the posteriors from the injection campaign, which indicate the spread in results due to different noise realizations. Finally, the dashed lines on the right-hand side are the posteriors from the zero-noise injection. This plot is intended to be compared to Fig. 2 of Ref. [27], and Fig. 6 of Ref. [28].
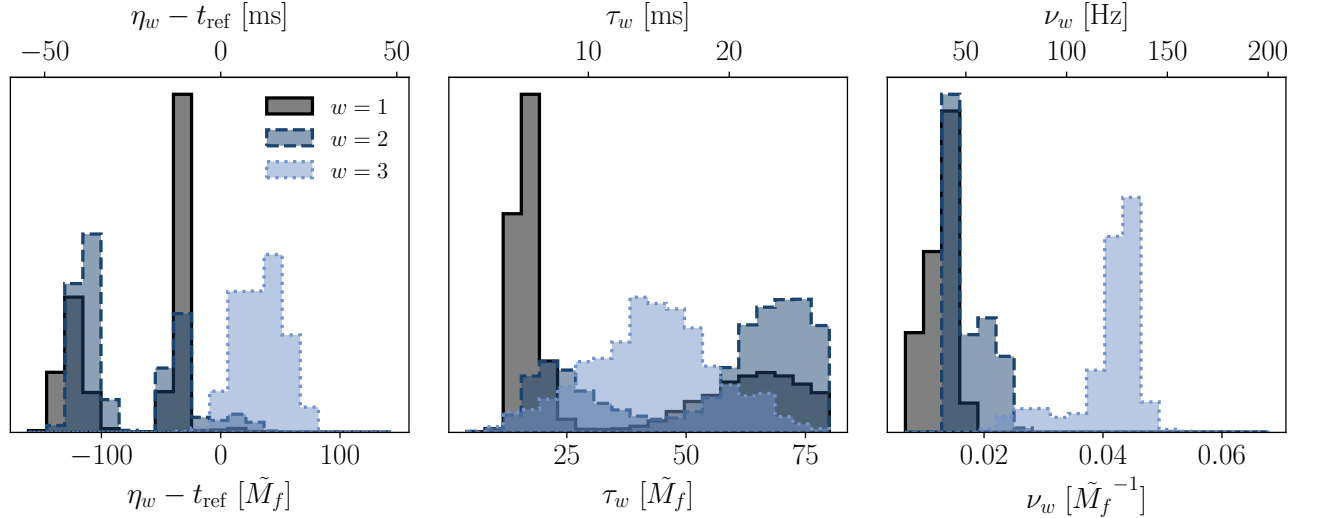

FIG. 11. Posteriors on selected parameters for the $W = 3$ wavelets used in the $N = 1$ overtone analysis, reweighted using the IMRP t_{peak} prior on the ringdown start time. *Left:* the wavelet central times, η_w. *Middle:* the wavelet widths, τ_w. *Right:* the wavelet frequencies, ν_w. The index runs over values $w = 1$, 2 and 3, where the numbering of the wavelets is chosen to enforce the ordering $\nu_w < \nu_{w+1}$. All plots use SI units on the upper x-axis and natural units on the lower x-axis.

analyses respectively. This is sufficient for the reweighting to remain accurate everywhere except for the earliest start time in the $N = 0$ analysis. This is the reason why this result is omitted from Fig. 2.

Appendix C: Injection study

Closely following the injection study performed in Ref. [27], we inject GW150914-like signals in the instrumental noise surrounding the true GW150914 event and re-perform our overtone analysis ($N = 1$).
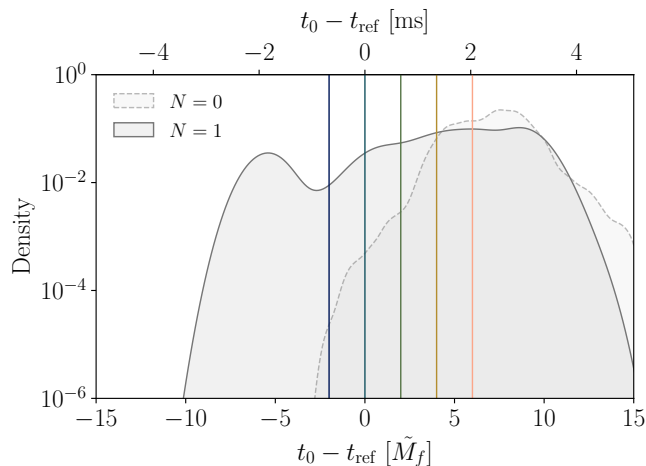
FIG. 12. Posteriors on the ringdown start time obtained from our initial analysis using a flat prior over the range shown in the plot. Results are shown for the fundamental only ($N = 0$) and overtone ($N = 1$) analyses. Vertical colored lines show the locations of the means $\bar{t}_0$ of the narrow Gaussian priors used for the subsequent reweighting (see Fig. 1). The $N = 1$ posterior has ample support across the entire range of interest, as required for the reweighting to remain accurate. The $N = 0$ posterior has enough support everywhere except the $\bar{t}_0 = -2\ \tilde{M}_f$ prior.

The $\ell = 2$ spin-weighted spherical harmonic of the numerical relativity simulation SXS:BBH:0305 [45, 46] was used as the mock signal, scaled to a total mass of $72\,M_\odot$ and injected with a face-off orientation at a luminosity distance of $410\,\mathrm{Mpc}$. The sky position was taken to be $\alpha = 1.95\,\mathrm{rad}$, $\delta = -1.27\,\mathrm{rad}$. This signal was injected into the data surrounding GW150914, such that the peak of the absolute value of the strain occurred at times $[-20, -15, -10, 5, 15, 20, 25, 30, 35, 40]\,\mathrm{s}$ relative to t_{ref}. These choices ensure the mock signal does not overlap with the real event. Additionally, a zero-noise injection was performed for comparison.

We performed the frequency-domain ringdown analysis on these mock datasets using the same setup as was used for the real data and as described in Sec. II B. This

includes using the same PSD in the likelihood for all datasets. We plot the resulting posteriors on the overtone amplitudes in Fig. 10. As with the real data, prior reweighting (see Sec. II C) has been used to show results for different choices of the ringdown start time prior. We also investigated the Bayes' factors and found the same declining trend.

As expected, different noise realizations introduce some scatter into the results and we observe a spread in the locations of the maximum posterior values for the overtone amplitudes. However, this spread is consistent with the width of the posterior. The analysis described in the main text found only tentative evidence for the overtone, but there is no indication that this is overly effected by noise fluctuations.

Appendix D: Wavelet posteriors

The frequency-domain ringdown analysis method described in [30] and used here marginalizes over the early-time inspiral-merger signal using a flexible combination of sine-Gaussian wavelets (see Eq. 2). In the GW150914 analyses presented in this paper $W = 3$ wavelets were used. This choice was found empirically to be large enough to model the inspiral-merger signal without biasing the ringdown inference. We have also verified that no strong correlations are observed between the wavelet and QNM parameters, and that further increasing the number of wavelets does not significantly affect the results for physically meaningful parameters (such as remnant parameters M_f and χ_f). These tests are described further in appendix A of Ref. [30].

The whitened strain posterior on the sum of these wavelets, together with the QNMs, can be seen in the early-time signal in Fig. 9 where the fit to the data is seen to be excellent. The wavelet parameters themselves are not physical; the wavelets are being used here solely to marginalize out the inspiral-merger. Nevertheless, in this appendix we show some additional posterior plots on the wavelet parameters, see Fig. 11. As expected, in order to describe the "chirping" inspiral signal, the wavelets naturally order themselves with their amplitudes and frequencies increasing with time.