Constraining Unmodeled Physics with Compact Binary Mergers from GWTC-1

Bruce Edelman,[1, *] F. J. Rivera-Paleo,[2] J.D. Merritt,[1] Ben Farr,[1] Zoheyr Doctor,[1] Jeandrew Brink,[3] Will M. Farr,[4, 5] Jonathan Gair,[6] Joey Shapiro Key,[7] Jess McIver,[8] and Alex B. Nielsen[9, 10]

[1] *Department of Physics, University of Oregon, Eugene, OR 97403, USA*
[2] *Departament d'Astronomia i Astrofsica, Universitat de Valncia, 46100, Burjassot, Spain*
[3] *Department of Mathematics and Applied Mathematics,*
University of the Free State, PO Box 339, Bloemfontein, South Africa
[4] *Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA*
[5] *Center for Computational Astrophysics, Flatiron Institute, New York NY 10010, USA*
[6] *Max Planck Institute for Gravitational Physics (Albert Einstein Institute),*
Am Mühlenberg 1, D-14476 Potsdam-Golm, Germany
[7] *University of Washington Bothell, 18115 Campus Way NE, Bothell, WA 98011, USA*
[8] *University of British Columbia, Vancouver, BC V6T 1Z4, Canada*
[9] *Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut), D-30167 Hannover, Germany*
[10] *Department of Mathematics and Physics, University of Stavanger, 4036 Stavanger, Norway*
(Dated: August 17, 2020)

We present a flexible model to describe the effects of generic deviations of observed gravitational wave signals from modeled waveforms in the LIGO and Virgo gravitational wave detectors. With the detection of 11 gravitational wave events from the GWTC-1 catalog, we are able to constrain possible deviations from our modeled waveforms. In this paper we present our coherent spline model that describes the deviations, then choose to validate our model on an example phenomenological and astrophysically motivated departure in waveforms based on extreme spontaneous scalarization. We find that the model is capable of recovering the simulated deviations. By performing model comparisons we observe that the spline model effectively describes the simulated departures better than a normal compact binary coalescence (CBC) model. We analyze the entire GWTC-1 catalog of events with our model and compare it to a normal CBC model, finding that there are no significant departures from the modeled template gravitational waveforms used.

I. INTRODUCTION

General relativity (GR) has passed a multitude of tests over the past years [1], but until the detection of gravitational waves (GWs) from binary black holes[2, 3] it had not been widely tested for strong dynamical gravitational fields. Gravitational-wave astronomy and more specifically that of compact binary coalescences (CBC) gives us access to a genuinely strong gravitational field regime to both test GR [4–7], and to provide constraints on new physics not predicted by GR modeled waveforms of these systems. Contemporary GW analyses employ matched filtering and forward modeling techniques, which both inherently rely on accurately modeled waveforms [3, 8–10]. We introduce here a model that can account for and measure in data deviations in phase and amplitude from a modeled waveform, either due to approximations inherent in the waveform calculation or a mismatch between theory (GR) and nature.

We present a parameterization that quantifies the deviations between the observed waveform ($h_{\mathrm{intrinsic}}$) and the GR waveform models (h_{model}), with few assumptions about the deviations. This provides the ability to perform additional tests of GR and also presents a generic model for describing and possibly constraining deviations from modeled gravitational waveforms, such

as higher-order modes [11] and tidal effects or deformities [12]. Quantifying such deviations is one of the major challenges in GW data analysis. The numerical method we use to parameterize deviations here is based on cubic spline interpolation in which the deviations (in phase and amplitude) are modeled as independent cubic spline functions interpolated from node points in the frequency domain. The cubic spline interpolant generates deviations that vary smoothly in frequency, but otherwise does not constrain the type or nature of deviation.

The splines employed to characterize the deviation provide a uniform way of describing GR departures rather than fitting separate parameters to the inspiral, merger, and ringdown (IMR) of the waveform separately as commonly done in GR IMR consistency tests [7, 13, 14]. In addition, IMR consistency tests have the same limitation as matched filtering, as they inherently assume perfect accuracy of the template waveform.

Another common test of GR is the parameterized test where one expands the waveform model in different regimes with post-Newtonian (PN) correction parameters added [15]. This test also builds in assumptions about how possible deviations may occur and has to fit different parts of the waveform with separate models as in the IMR consistency test. One commonly used class of model-agnostic tests of GR are residual [6] tests, where best-fit waveform models are subtracted from the data and normality tests (e.g., Anderson-Darling) are conducted on the residuals. Such tests would be sensitive to very large deviations from the signal model, but not

* bedelman@uoregon.edu

to small but correlated deviations across many frequency bins. Our proposed model differs from these tests and constraints by allowing uncertainty in the template waveforms and letting it vary smoothly across the entire frequency range. Our model is able to describe and fit the inaccuracies in waveform models. With the assumption that the template GR waveform is completely accurate, it provides a clear way of describing and constraining unmodeled physics of our waveform models or departures from GR across the frequency range of the waveform.

In Section II we describe the model and methods for incorporating it in the `LALInference` Bayesian Analysis software [10]. Our implementation is similar to the calibration spline model described in [16]. We then present simulated deviations on which we validated the performance of our model in Section III, followed by discussions and implications of the results. In Section IV we present results of this model on the entire first LIGO and Virgo Gravitational Wave Transient Catalog (GWTC-1) [17] which includes the results of the first and second observing runs of the Advanced LIGO [18] and Virgo [19] detectors. This catalog of gravitational wave events includes ten binary black hole detections and one binary neutron star detection [17, 20]. Lastly the results and conclusions are summarized and discussed in Section V.

II. METHODS

A. Waveform Representation

When a gravitational wave enters a gravitational wave detector, the detector records a data stream which we can describe in the frequency domain as $d(f) = h_{\mathrm{observed}}(f) + n(f)$, which is an additive combination of a waveform $h_{\mathrm{observed}}(f)$ and noise $n(f)$. The observed waveform in a detector can be represented as the sum of the intrinsic waveform polarizations, projected across that detector. This is done by multiplying the two (time-dependent) antenna pattern terms of that detector, F_+ and $F_\times$, to the plus and cross gravitational-wave polarizations as:

$$h_{\mathrm{observed}}(f) = h_+(f)F_+(f) + h_\times(f)F_\times(f) \quad (1)$$

with

$$h_{+,\times}(f) = h_{\mathrm{model},+,\times}(f)\left[1 + \delta A(f)\right]\exp\left[i\delta\phi(f)\right]. \quad (2)$$

Since we are searching for deviations from the coherent modeled waveforms, every detector observing a GW should see the same deviations. We model the uncertainties in the intrinsic waveform, $h_{\mathrm{model},+,\times}(f)$, as frequency-dependent amplitude and phase departures in $h_{+,\times}$ with respect to $h_{\mathrm{model},+,\times}$. This is the same technique that Farr *et al.* [16] use to model calibration errors in each detector independently. We also take the assumption that there are only two polarizations and that GWs travel at the speed of light. While our model is degenerate with the calibration model, coherent deviations observed across all detectors are modeled with a single set of

parameters for this spline model, as opposed to modeling them independently in each detector using the calibration model. Thus we expect our prior distribution to result in deviations to the intrinsic waveform being preferably described by the coherent spline, and the calibration splines to measure only detector-dependent deviations.

We assume that phase deviations are small and under this assumption, we can approach the exponential term as

$$\exp\left(i\delta\phi(f)\right) = \frac{2 + i\delta\phi(f)}{2 - i\delta\phi(f)} + \mathcal{O}(\delta\phi^3) \quad (3)$$

which is more computationally efficient [16]. Then, we can rewrite the intrinsic waveform as:

$$h_{+,\times}(f) = h_{\mathrm{model},+,\times}(f)\left[1 + \delta A(f)\right]\frac{2 + i\delta\phi(f)}{2 - i\delta\phi(f)} \quad (4)$$

This replacement agrees with the exponential term to third order for small phase deviations, and differs by 5% from the exponential term for the largest simulated deviation used in this paper of 60 degrees.

The intrinsic waveform, after being modified as Eq. 4, is then projected across the detectors to get h_{observed} in each detector as in Eq. 1. Despite the expectation that these departures are small, they have the potential to impact the measurement of all parameters of the source (masses, spins, distance, etc.) [9]. A consequence of modeling these deviations with a purely phenomenological model is that we can no longer trust the inference of astrophysically modeled parameters when deviations are present.

Under the assumption that $\delta A(f)$ and $\delta\phi(f)$ vary smoothly in frequency, they can be modeled by a spline function [16].

B. Spline Model

A spline function is a piece-wise polynomial interpolation that obeys smoothness conditions at the nodal points where the pieces connect. In the following, we use the case of cubic splines defined by 15 nodal points confined to a finite frequency interval. Formally these departures can be written as

$$\delta A(f) = \mathcal{I}_3(f; \{f_i, \delta A_i\}), \quad (5)$$
$$\delta\phi(f) = \mathcal{I}_3(f; \{f_i, \delta\phi_i\}), \quad (6)$$

where $\mathcal{I}_3$ is a cubic spline interpolant, the $\{f_i\}$ are the nodes of the spline interpolant in frequency, and $\{\delta A_i\}$ and $\{\delta\phi_i\}$ are the values of the spline at those nodal points. To better generalize the model to fit a larger variety of possible departures, we freely let the nodal points move around in frequency space during sampling (with the condition that they do not exchange orders or get too close to each other) after being initialized linearly in log-frequency space, as done in Farr *et al.* [16], Vitale *et al.* [21]. We choose the node locations to be the same

for the amplitude and phase spline functions as we expect deviations to happen at similar frequencies and to reduce our model degrees of freedom. One could choose an independent set of node locations for phase and amplitude if expecting to have deviation effects that alter the amplitude and phase at different points in frequency. We prevent nodal points from getting arbitrarily close as this causes the spline to be too sensitive to very small changes in node positions, creating extreme deviations to satisfy the smoothness conditions of the spline function. The parameters added to our inference for this model are then the $\{\delta A_i\}$, $\{\delta\phi_i\}$ and $\{f_i\}$.

With our assignment of Gaussian priors (see § II C) to the δA and $\delta\phi$ parameters, our spline model implements a Gaussian process prior for the waveform deviations from the model (c.f. [22]).

C. Statistical Framework

We place a Gaussian prior on the departure parameters, $\{\delta A_i\}$ and $\{\delta\phi_i\}$, centered around zero, with σ_A and σ_ϕ characterizing our prior uncertainties about the magnitude of the departures in amplitude and phase from the modeled waveforms.

$$p(\delta A_i) = N(0, \sigma_A) \qquad (7)$$
$$p(\delta\phi_i) = N(0, \sigma_\phi) \qquad (8)$$
$$p(f_i) = U(f_{\text{low}}, f_{\text{high}}) \qquad (9)$$

It is important however, not to think of the Gaussian priors on the node values of $\{\delta A_i\}$ and $\{\delta\phi_i\}$ as the broadband uncertainties of the interpolated spline function across the frequency range. In practice the prior widths on the nodes are wider than the broadband uncertainty resulting from sampling the prior. We constrain the spline nodes to be increasing in frequency:

$$f_{i-1} < f_i < f_{i+1}. \qquad (10)$$

The spline model introduces some challenges that need to be accounted for, the first being the freedom in the frequency values of the node points. Since we also use the node positions in frequency space as sampling parameters and we place uniform priors on the spline node locations across the frequency band from f_{low} to the Nyquist frequency, the model is degenerate under exchange of node positions. To circumvent this degeneracy, we impose that the node positions stay ordered as in Eq. 10. These parameters can then be fit and the corresponding calibration errors marginalized over during inference.

Another challenge with the spline function is that if two nodes get too close together, obeying the conditions required of the cubic spline can lead to the interpolated deviations becoming very extreme. To account for this we prevent the frequency nodes from getting closer than 4 frequency bins away from one another as shown in equation 11, where the frequency bin width (in Hz) is determined from the segment length of data we are analyzing

as $df = 1/T$ with T the segment length. That is, we reject any configurations with

$$f_{i+1} - f_i < 4 * \delta f = \frac{4}{T_{\text{obs}}}. \qquad (11)$$

Other spline based interpolation methods such as `BayesLine` combat this by keeping the spline nodes on a fixed frequency grid then turns them on/off during inference using trans-dimensional Markov-Chain Monte Carlo sampling [23]. We experimented with a few other fixes to this issue. We implemented a Gaussian prior on every frequency bin location; we also tried evaluating the Gaussian prior on some number of points between nodes of the spline interpolants to disfavor the large spline excursions as well. In practice both of these priors turned out to be too restrictive while attempting to recover the simulated deviations presented in Section III; they may still be useful when exploring astrophysical events.
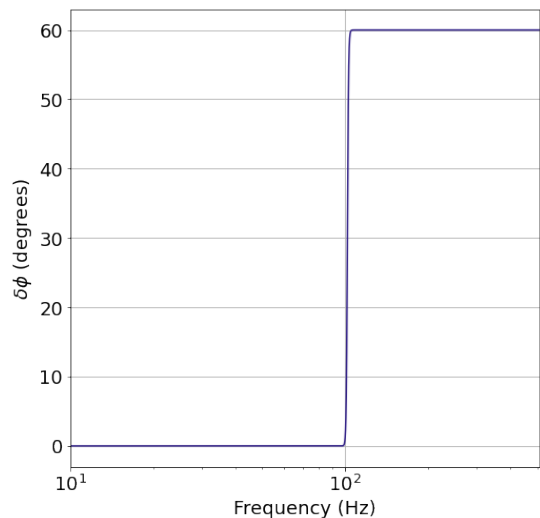

FIG. 1. Simulated phase deviation $\delta\phi$ for extreme spontaneous scalaraization toy model.

The last challenge is that our model of amplitude deviations is perfectly degenerate with the distance to the signal. The distance to the source can be increased while simultaneously producing a positive amplitude deviation across the entire frequency band of the signal to compensate. The default astrophysical prior used in `LALInference` [10] is $\propto D_L^2$ with D_L the luminosity distance. When this prior distribution is coupled with the zero-mean Gaussian priors on amplitude deviations, the strong $\propto D_L^2$ prior on distance almost always results in systematically positive amplitude deviations in the spline component of the model. However, by allowing for broadband changes to the amplitude of the signal in the first place we are no longer able to meaningfully infer the distance to the source. In other words, we use D_L (with a uniform prior) as a phenomenological parameter to fit the broad-band amplitude of the signal, and the spline
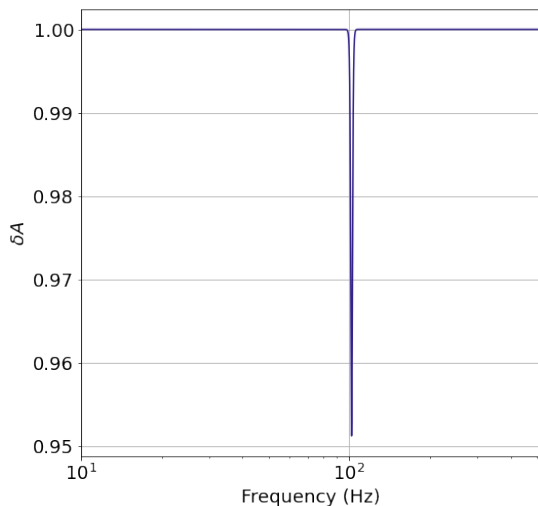
FIG. 2. Simulated amplitude deviation δA for extreme spontaneous scalaraization toy model.

model describes any frequency-dependent deviations that may be present.

With the prior assumptions and modified waveform we can now construct the posterior distribution according to Bayes' Theorem,

$$p(\boldsymbol{\theta}, \{\delta A_i\}, \{\delta\phi_i\}, \{f_i\}|d) \propto$$
$$\mathcal{L}(d|\boldsymbol{\theta}, \{\delta A_i\}, \{\delta\phi_i\}, \{f_i\})p(\boldsymbol{\theta})p(\{\delta A_i\})p(\{\delta\phi_i\})p(\{f_i\}) \tag{12}$$

then use the `LALInference` Markov Chain Monte Carlo (MCMC) algorithm to draw samples from the posterior distribution in Eq. 12, with $\boldsymbol{\theta}$ the normal CBC parameters, d the gravitational wave strain data, $\mathcal{L}$ the standard GW likelihood with the modified intrinsic waveform as shown in Eq. 4.

III. SIMULATED DEVIATIONS

In order to validate our spline model, we run on data with astrophysically motivated deviations included. We choose a toy model that presents an extreme case of spontaneous scalarization [24] which modifies a simulated high-mass BBH waveform with similar astrophysical parameters as GW150914 [2, 3]. We generate two waveforms where one has simulated modifications according to the toy model and the other is exactly as the modeled GR waveform describes. We then simulate colored Gaussian noise according to the the Advanced LIGO design sensitivity noise curve or power spectral density (PSD) [18] and add it to the simulated waveforms.

We first generate a GR-based frequency domain waveform using the IMRPhenomD [25] waveform approximant with the simulated parameters shown in Table I, then modify the same waveform according to our toy model described above extreme scalarization case [24].

m_1	36 $M_\odot$
m_2	29 $M_\odot$
$D_{\rm L}$	450 Mpc
ϕ	2.76 rad
α	1.37 rad
δ	-1.26 rad
$S_{1,z}$	0.0
$S_{2,z}$	0.0

TABLE I. Parameters for simulated validation signals

For the phase $\delta\phi$, the toy model includes an abrupt increase of 60 degrees centered at $f_z = 102$Hz, with a width of $df = 1$Hz. The amplitude temporarily drops by 5% in frequency, again centered at $f_z = 102$Hz and with a width of $df = 1$Hz. To get these modifications we used these parameters in Eqs. 13 and 14 with $dA = 0.1$, $d\phi = 60°$.

$$\delta A(f) = e^{\frac{1}{2}dA\left(\tanh\left(\frac{f-f_z}{df}\right)^2 - \tanh\left(\frac{f_{\rm ref}-f_z}{df}\right)^2\right)} \tag{13}$$

$$\delta\phi(f) = \frac{1}{2}d\phi\left[\tanh\left(\frac{f-f_z}{df}\right) - \tanh\left(\frac{f_{\rm ref}-f_z}{df}\right)\right] \tag{14}$$

A. Results on simulated signals

To compare the effect of the waveform modification and the spline model, we ran the `LALInference` parameter estimation software on both the modified and unmodified signals with our spline deviation model turned on and off. Figures 3 and 4 show the corner plots for the simulated intrinsic (component masses, spins, etc) and extrinsic (luminosity distance and sky localisation) GW parameters comparing the 1-D and 2-D marginalised posterior distributions with the spline model on or off. In figure 3 we see that there is minimal difference in posteriors for the unmodified signal for most of the parameters which is what we would expect for the case of no deviations. We also expect to see greater uncertainties on certain parameters with the spline model as there are possible degeneracies with the spline and other parameters as seen in the different 1-D spin parameter posteriors in figure 3. This means the normal CBC model or template waveform is able to explain the entire coherent signal in the data.

For the Modified case (fig 4) we see that the simulated values are included in most of the spline model posterior distributions but not the posteriors with the spline model inactive. This is because the modified signal includes the abrupt (in frequency) modifications or deviations simulated and these types of extreme abrupt deviations are very poorly described by the template waveform models
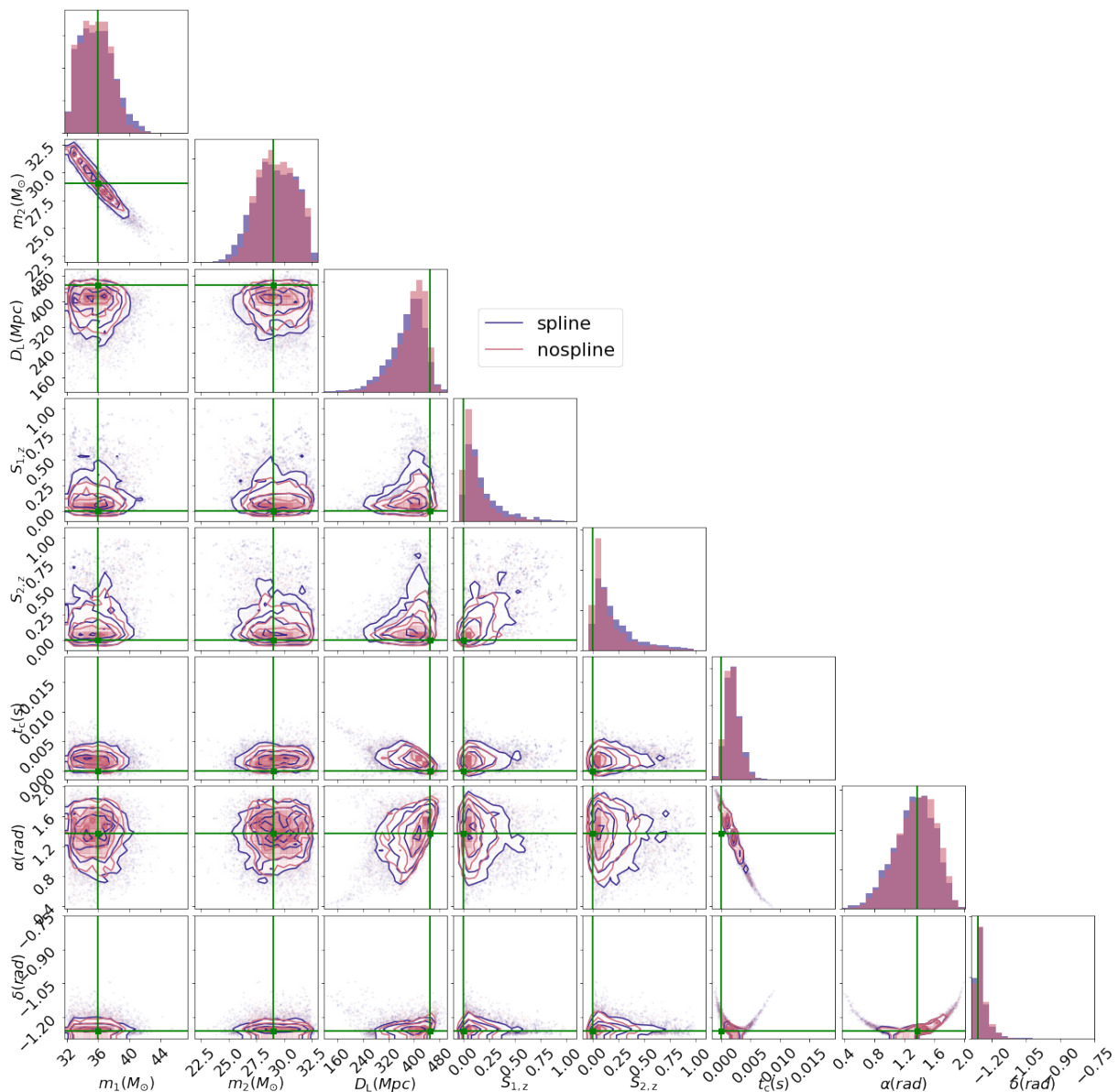
FIG. 3. Corner plots showing the 1-D and 2-D marginalised posterior distributions for simulated parameters for PE runs on the Unmodified Signal with spline model on (purple) or off (pink). This demonstrates the impact of the model flexibility on astrophysical parameter uncertainties.

and especially the IMRPhenomPv2 waveform template used in this analysis. However, we see that with our spline model turned on, the posterior distributions do encompass most of the simulated parameter values. This illustrates that parameter estimation (PE) without including deviation parameters would not be reliable at estimating the true signal parameters in cases where there may be departures from the waveform template used. In particular we see from figure 4 that the masses and trigger time posterior distributions are more consistent with the simulated values with the spline model fitting the deviations.

Figure 5 shows the interpolated spline functions for the

deviations in amplitude and phase, δA and $\delta \phi$, for both the modified and unmodified simulated signals (right, left). The plots show the 1σ and 2σ bounds in the spline interpolants along with the median. We see here that our model consistently recovers zero deviations for the unmodified signal across the frequency band. The ranges included in the 1σ and 2σ bands show the exploration of the prior bounds while sampling while also being symmetric around the median. Looking at the modified case, we see a presence of deviations away from zero in phase around 100 Hz. The phase recovery does not show the clear step function behavior that was simulated and shown in figure 1 as that extreme deviation is disfavored
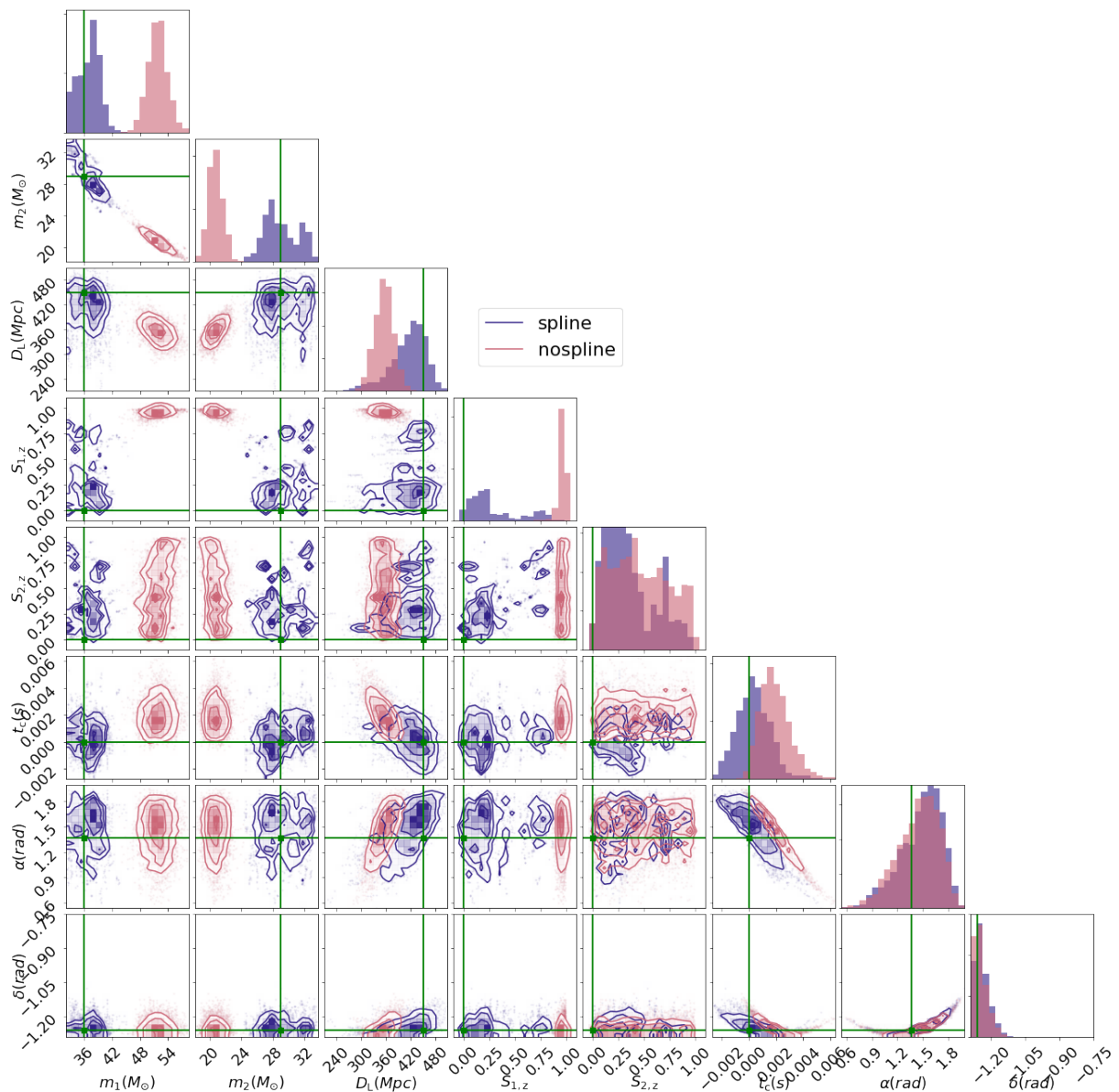
FIG. 4. Corner plots showing the 1-D and 2-D marginalised posterior distributions for simulated parameters for PE runs on the Modified Signal with spline model on (purple) or off (pink). This demonstrates inaccuracies in parameter estimation performed on signals containing deviations from the physically modeled waveforms.

by the prior, because the priors used on the node positions are Gaussian distributions centered about zero; to recover the large flat step in frequencies greater than 100 Hz would require very low prior probability at those nodes. The spline posterior does however show that there is a transition at the modified frequency (~ 100 Hz) and shows that the phase increases roughly by 60 degrees as we simulated. The model compensates for this by "ramping" down the phase modification so that it slopes back to zero after the merger frequency (~ 250 Hz for this signal) since there is no signal content to infer from after merger. The model may also be able to coherently ramp down as a result of possible degeneracies with other pa-

rameters. The amplitude recovery is less revealing since the simulated deviation is on order of the prior width along with the sharp resolution of the deviation in frequency, it is harder to clearly recover that feature in our model. The presence of deviations is corroborated by the fact that the posterior distribution on the deviations excludes zero deviation at some frequencies at the 95% level, which we do see in our modified case. To focus on this we can calculate the posterior quantile of phase and amplitude deviation that the x-axis (0-deviations) falls at for each frequency bin. This is shown in figure 6.

From this we see in the phase plot that there are significant portions of the frequency range where the
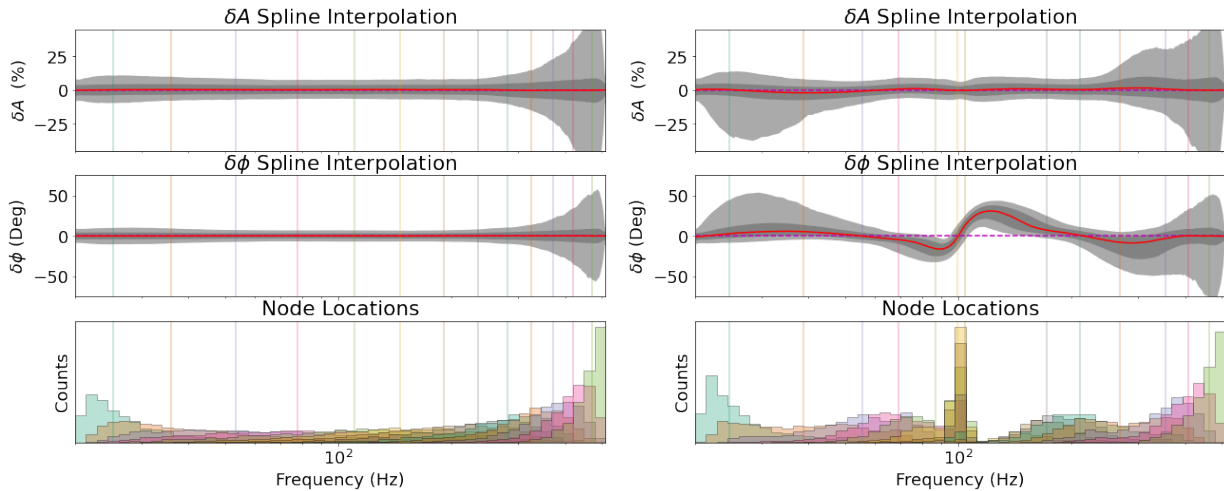
FIG. 5. Spline Interpolation of unmodified (left) and modified (right) signal deviations. 1 and 2 σ credible intervals (grey) and the median spline (red) are shown with top panel the amplitude deviations and middle panel the phase deviations. In the bottom frame we plot the node position posterior distributions, which you can see clump towards the frequency of deviations in the modified case in the bottom right most panel while they are more uniformly distributed for the unmodified case as they are exploring the prior.
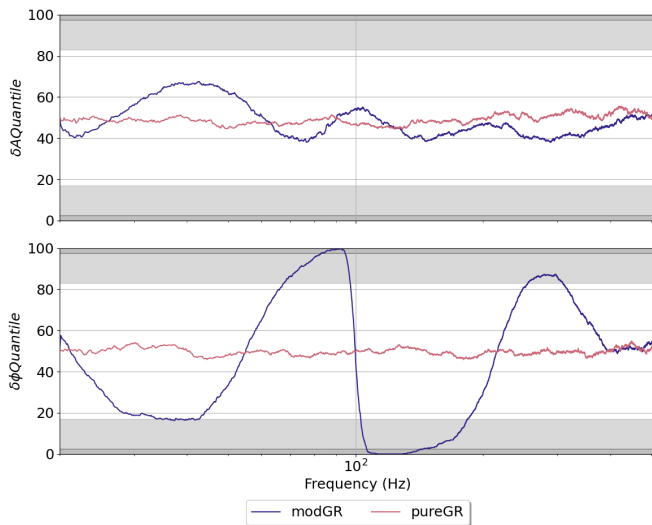

FIG. 6. Posterior quantile of the spline interpolant that 0-deviation corresponds to for the simulated unmodified (purple) and modified (pink) signals.
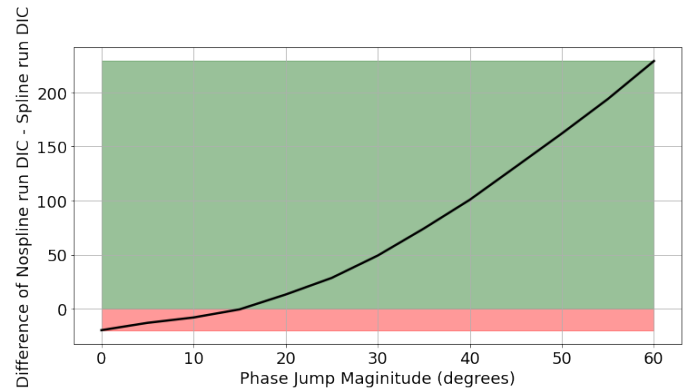

FIG. 7. The DIC value from spline model off run minus spline on run for different simulated phase deviations. Green shaded regions are where the spline model is preferred and red shaded region is where the spline model off is preferred.

B. Model Comparisons

A useful way to see how well the spline model performs on a given segment of data we can perform a model comparison between the spline on and spline off models. To do this we take an Information Theory approach by computing the Deviance Information Criterion (DIC) [26, 27]. This measure of fitness has the feature that lower values correspond to a better fit and includes an "Occam's Factor" that penalizes models with greater numbers of parameters. To effectively compare the spline on to spline off models we take the DIC value from the spline off and subtract the DIC value from the spline on model giving us a single value that is positive if the spline model is preferred and negative if the spline model is disfavored. We

phase interpolant ruled out zero deviations at greater than the 99% credible level. The amplitude recovery shows less excursion which tells us we would not be able to constrain amplitude deviations of this type and magnitude. We do see significantly more freedom in amplitude than the no-modifications case which considerably constrained the amplitude interpolants to zero, however zero deviations lies within the 1σ range across the frequency band meaning it could still be consistent with no deviations present in amplitude.

now perform a study where we run on simulated signals with modifications as before but with increasing jumps in phase from 0 to 60 degrees incrementing by 5 degrees. Looking at figure 7, we see that as the phase jump increases the spline model becomes more favored. We can compare the DIC differences from figure 7 and to the differences for the un-modified signal which has DIC difference of -74.46, showing that for the DIC the spline model off is preferred at about the same level the spline model on is preferred for a phase jump of 35 degrees.

C. Model Limitations and Alternative Parameterizations

We have attempted a few other iterations of this model to increase its flexibility and performance on simulated signals which were not included in our final analyses. The first was to parameterize the deviations in the derivative of the phase/amplitude. This was tried for the same reason that the model was insensitive to the step function deviation that we have simulated in the phase. The step function was very incompatible with the normal model priors as each node after stepping up is penalized from our priors yet if we parameterized in the derivative the step function derivative looks like a delta function which is much more compatible with zero-centered Gaussian priors on the spline nodes. In practice this brought more challenges than it solved increasing other degeneracies with parameters and did not seem to qualitatively improve our efficiency or performance of correctly fitting the simulated deviations.

As a further test of our model, we attempted to recover modifications to a neutron star-black hole (NSBH) merger waveform. In the event of tidal disruption of the neutron star, there is an expected deviation to the waveform predicted by GR for non-deformable bodies. Specifically, we constructed a toy model featuring a roll-off in amplitude beyond a disruption frequency. Physically, this corresponds to a spreading and redistribution of mass after the moment of disruption, which would decrease the intensity of GWs emitted from then on [28, 29]. We made no change to the phase. Our model was consistently unable to recover these deviations. This is likely due to more degeneracies in our parameters, and specifically in component mass. By moving to a higher mass, parameter estimation can push the amplitudes lower at high frequencies, and the additional flexibility of the spline model was unable to capture any of the deviation. Better results may be had with a more realistically modified NSBH waveform, or by changing the way we manage degeneracies in our parameters.
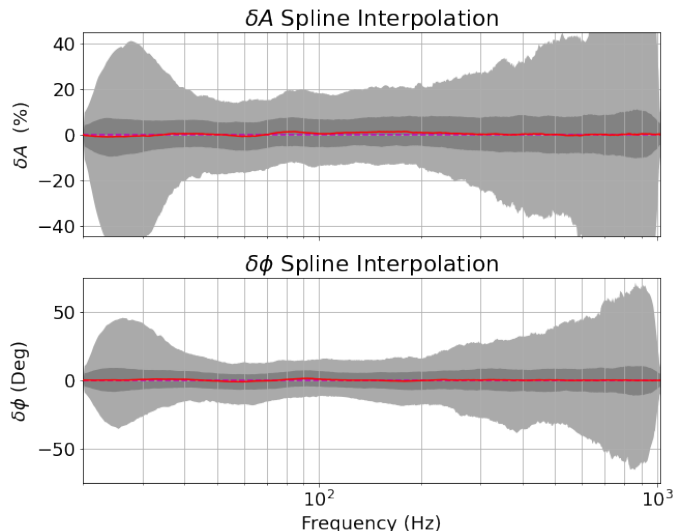

FIG. 8. Spline interpolation of GW170823 with 1 and 2 σ credible intervals (grey) and the median spline interpolant (red) shown.
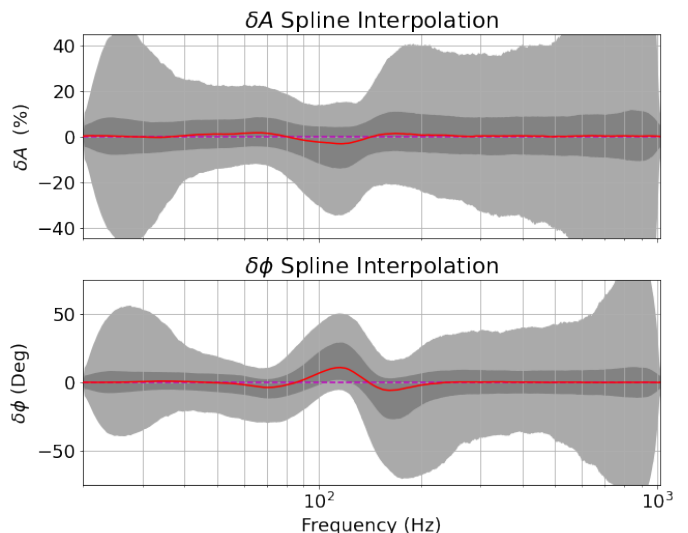

FIG. 9. Spline interpolation of GW170729 with 1 and 2 σ credible intervals (grey) and the median spline interpolant (red) shown.

IV. RESULTS FROM LIGO-VIRGO PUBLIC DATA

The LIGO-Virgo GWTC-1 catalog [17] presents a population of GW events to study deviations of GR along with a suite of data with which to validate our gravitational waveform models. We first discuss the results of our model on the ten binary black hole merger events listed with a few selected median posterior parameters in table II, and then we will discuss the one binary neutron star event, GW170817.

We performed parameter estimation on the ten binary

GWTC-1 Events							
Event	χ_{eff}	d_L (Mpc)	m_1 ($M_\odot$)	m_2 ($M_\odot$)	$\mathcal{M}$ ($M_\odot$)	ρ_{gstlal}	Source
GW150914	$-0.01^{+0.12}_{-0.13}$	$430.0^{+150.0}_{-170.0}$	$35.6^{+4.8}_{-3.0}$	$30.6^{+3.0}_{-4.4}$	$28.6^{+1.6}_{-1.5}$	24.4	BBH
GW151012	$0.04^{+0.28}_{-0.19}$	$1060.0^{+540.0}_{-480.0}$	$23.3^{+14.0}_{-5.5}$	$13.6^{+4.1}_{-4.8}$	$15.2^{+2.0}_{-1.1}$	10.0	BBH
GW151226	$0.18^{+0.2}_{-0.12}$	$440.0^{+180.0}_{-190.0}$	$13.7^{+8.8}_{-3.2}$	$7.7^{+2.2}_{-2.6}$	$8.9^{+0.3}_{-0.3}$	13.1	BBH
GW170104	$-0.04^{+0.17}_{-0.2}$	$960.0^{+430.0}_{-410.0}$	$31.0^{+7.2}_{-5.6}$	$20.1^{+4.9}_{-4.5}$	$21.5^{+2.1}_{-1.7}$	13.0	BBH
GW170608	$0.03^{+0.19}_{-0.07}$	$320.0^{+120.0}_{-110.0}$	$10.9^{+5.3}_{-1.7}$	$7.6^{+1.3}_{-2.1}$	$7.9^{+0.2}_{-0.2}$	14.9	BBH
GW170729	$0.36^{+0.21}_{-0.25}$	$2750.0^{+1350.0}_{-1320.0}$	$50.6^{+16.6}_{-10.2}$	$34.3^{+9.1}_{-10.1}$	$35.7^{+6.5}_{-4.7}$	10.8	BBH
GW170809	$0.07^{+0.16}_{-0.16}$	$990.0^{+320.0}_{-380.0}$	$35.2^{+8.3}_{-6.0}$	$23.8^{+5.2}_{-5.1}$	$25.0^{+2.1}_{-1.6}$	12.4	BBH
GW170814	$0.07^{+0.12}_{-0.11}$	$580.0^{+160.0}_{-210.0}$	$30.7^{+5.7}_{-3.0}$	$25.3^{+2.9}_{-4.1}$	$24.2^{+1.4}_{-1.1}$	15.9	BBH
GW170817	$0.0^{+0.02}_{-0.01}$	$40.0^{+10.0}_{-10.0}$	$1.46^{+0.12}_{-0.1}$	$1.27^{+0.09}_{-0.09}$	$1.186^{+0.001}_{-0.001}$	33.0	BNS
GW170818	$-0.09^{+0.18}_{-0.21}$	$1020.0^{+430.0}_{-360.0}$	$35.5^{+7.5}_{-4.7}$	$26.8^{+4.3}_{-5.2}$	$26.7^{+2.1}_{-1.7}$	11.3	BBH
GW170823	$0.08^{+0.2}_{-0.22}$	$1850.0^{+840.0}_{-840.0}$	$39.6^{+10.0}_{-6.6}$	$29.4^{+6.3}_{-7.1}$	$29.3^{+4.2}_{-3.2}$	11.5	BBH

TABLE II. Posterior estimates for LIGO's GWTC-1 catalog of events with +/- 1σ errors shown as well
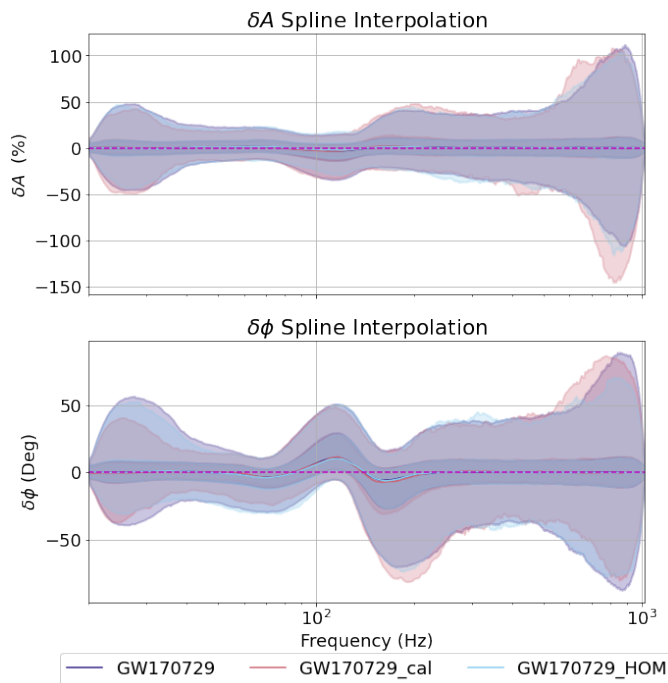

FIG. 10. Spline interpolation of GW170729 with 1σ percent credible intervals shown, comparing runs on GW170729 with calibration turned on/off and using a HOM waveform also with calibration turned off.

black hole (BBH) events with the template waveform IMRPhenomPv2 [25] with our spline model turned on and off to compare the results of each. These runs used the same prior settings detailed in Section II which is a 5% uncertainty of the amplitude and 5 degrees uncertainty on phase. First we look at the spline plot recoveries for GW170823 and GW170729 to show two different cases, both still consistent with zero deviations. In figure 8 we see the case where our model is consistent with zero deviations consistently across the frequency band with the credible intervals illustrating the prior exploration during sampling. Contrasting to this we can look at figure 9 to see a case in which our model has less posterior support for zero deviations at some frequencies. We see here that there is area on the plot, most importantly the phase portion, around 100 Hz in which zero deviations falls nearly outside our 1σ interval. The first thing one might wonder seeing deviations here is whether this can be explained by the normal calibration uncertainties of each detector being similar. We check this in fig 10 along with a run using a waveform model including higher-order modes (HOM) or greater than $\ell = 2$ modes in the spherical harmonic expansion [30], IMRPhenomPv3HM [31]. As seen in table II GW170729 is the most massive event from GWTC-1 and HOM are more important for higher mass systems along with asymmetric mass systems [11, 32]. We can clearly see a very similar spline interpolant recovery leading us to believe this is fitting features in the data that are unexplained by either a higher-order mode waveform or similar calibration errors across the network of detectors. However for each of the GW170729 results still have posterior support for zero deviations at the frequencies of largest excursions.

As in the previous section we can now look at which quantile of the spline interpolant posteriors will fall along the x-axis (corresponding to no deviation), at each frequency bin for each event. This is shown in figure 11 in which we highlight the $> 1\sigma$ and $> 2\sigma$ bands. We notice in this figure that for all ten binary black hole events, no deviations or modifications in amplitude fall outside the 1-sigma interval. We do see that for two events, namely GW170729 and GW170814, there are regions in which no deviation falls outside of the 1σ band. However there are significant portions where GW170817 also falls outside the 1σ band. Looking at the spline posterior in figure 12 for GW170817 shows the clear departures away from zero at some frequencies but overall outside of that small region looks behaved.

To further check our analysis we perform the same model comparisons described earlier in this paper on an event by event basis comparing the spline model on to
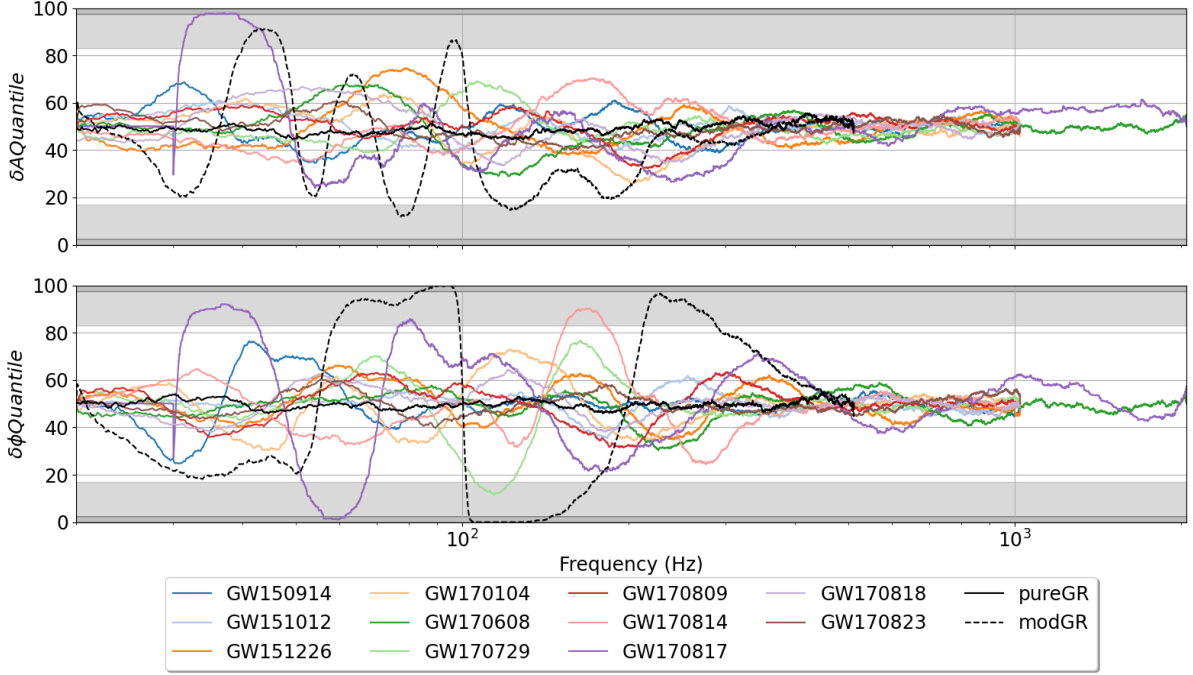
FIG. 11. Quantile of spline interpolation that 0-deviation corresponds to for GWTC-1 events. (Different choices of f_{low} and sampling rates were chosen for some events)

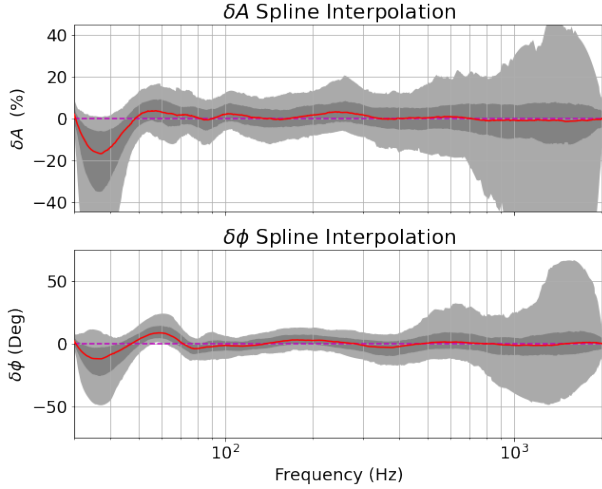FIG. 12. Spline interpolation of GW170817 with 1 and 2 σ credible intervals (grey) and the median spline interpolant (red) shown.
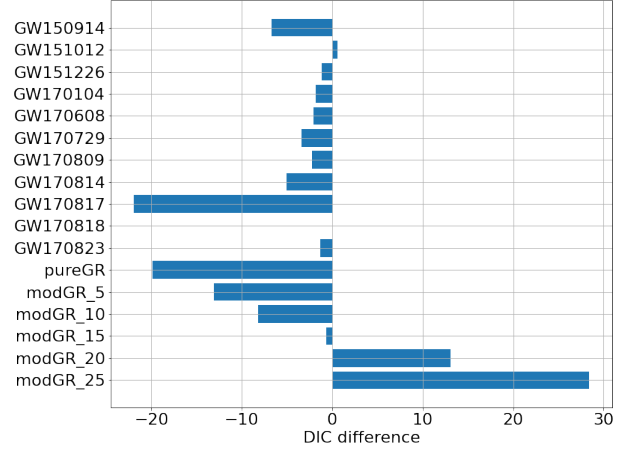

FIG. 13. DIC of spline model off run minus DIC of spline model on run for GWTC-1 events and simulated signals. Negative values correspond to spline model disfavored while positive values show the spline model favored.

the spline model off. This is seen in figure 13 with the same results from differing the phase jump as discussed in the previous section. This shows that for each event there is no preference to either model from their DIC values. Even for the events with more extreme spline interpolant posterior (i.e, GW170729 and GW170817) we still see that from a model comparison approach the spline model does not significantly describe the data better than

without the spline model. It may seem counter intuitive that the largest outlier event (GW170817) in spline interpolant posterior distributions shown in figure 11 is the event with the spline model least favored as shown in 13. When investigating the components that go into the DIC calculation the mean of the log-likelihood distributions for GW170817 with the spline model on and off are nearly the same while the variance in the log-likelihood

distributions (this is the Occam's Penalty factor for the DIC) is much greater for the spline model because of the increased degrees of freedom. This is to say that the spline model is able to fit some features in the data but it does not describe the data better than the normal CBC model enough to outweigh the greater introduced model flexibility.

V.   CONCLUSIONS

We have presented a useful model and parameterization to describe general departures or deviations from gravitational waveform models. Our model can be used to look for departures of any of LIGO's modeled waveforms by generically fitting the entire frequency band at once with spline functions. We find that for the 11 gravitational wave events of both BBH and BNS origin from LIGO's GWTC-1 catalog, are all consistent with the IMRPhenomPv2 waveform. Shown in figure 11 there are two events that we consider "outliers" with some portion of the phase deviation outside of the 1σ range but still lying well within the 2σ bounds of no deviations, which for a sample size of 11 events would not be unexpected, even with no deviations present.

Currently, more investigation into possible degeneracies of our model would be necessary to vet any significant sign of deviation. Based on the generally larger uncertainties in amplitude parameters and the ampli-

tude/luminosity distance degeneracy future work may choose to only model phase deviations as this reduces the model degrees of freedom by a third and may provide a stronger constraint on possible deviations. Further studies also need to be done to evaluate effects of detector sensitivity on our model, expand the validation of our model on other physically motivated deviations that can be simulated, and possibly incorporating information from proposed alternatives to GR into the priors. However, the model presented in this paper can be used as a model agnostic test to look for first signs of departures in the modeled waveforms. With more events and increased detector sensitivity, we will be able to better constrain any general deviations with our model while at the same time giving us a better testing set to look for hidden degeneracies between our model and other parameters.

VI.   ACKNOWLEDGMENTS

[1] C. M. Will, Living Rev. Relativity **17**, 0510072 (2014), arXiv:1403.7377 [gr-qc].
[2] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al., Physical Review Letters **116**, 061102 (2016), arXiv:1602.03837 [gr-qc].
[3] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al., Physical Review Letters **116**, 241102 (2016), arXiv:1602.03840 [gr-qc].
[4] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al., Physical Review Letters **116**, 221101 (2016), arXiv:1602.03841 [gr-qc].
[5] B. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, V. Adya, and et al., Physical Review Letters **123** (2019), 10.1103/physrevlett.123.011102.
[6] B. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, V. Adya, and et al., "Tests of general relativity with the binary black hole signals from the ligo-virgo catalog gwtc-1," (2019), arXiv:1903.04467 [gr-qc].
[7] A. Ghosh, A. Ghosh, N. K. Johnson-McDaniel, C. K. Mishra, P. Ajith, W. Del Pozzo, D. A. Nichols, Y. Chen, A. B. Nielsen, C. P. L. Berry, and L. London, Phys. Rev. D **94**, 021101 (2016).

[8] C. J. Moore and J. R. Gair, Physical Review Letters **113** (2014), 10.1103/physrevlett.113.251101.
[9] R. J. E. Smith, K. Cannon, C. Hanna, D. Keppel, and I. Mandel, Phys. Rev. D **87**, 122002 (2013).
[10] J. Veitch, V. Raymond, B. Farr, W. Farr, P. Graff, S. Vitale, B. Aylott, K. Blackburn, N. Christensen, M. Coughlin, W. Del Pozzo, F. Feroz, J. Gair, C.-J. Haster, V. Kalogera, T. Littenberg, I. Mandel, R. O'Shaughnessy, M. Pitkin, C. Rodriguez, C. Röver, T. Sidery, R. Smith, M. Van Der Sluys, A. Vecchio, W. Vousden, and L. Wade, Phys. Rev. D **91**, 042003 (2015).
[11] K. Chatziioannou, R. Cotesta, S. Ghonge, J. Lange, K. K. Ng, J. Caldern Bustillo, J. Clark, C.-J. Haster, S. Khan, M. Prrer, and et al., Physical Review D **100** (2019), 10.1103/physrevd.100.104015.
[12] B. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, V. Adya, and et al., Physical Review Letters **122** (2019), 10.1103/physrevlett.122.061104.
[13] A. Ghosh, N. K. Johnson-McDaniel, A. Ghosh, C. K. Mishra, P. Ajith, W. D. Pozzo, C. P. L. Berry, A. B. Nielsen, and L. London, Classical and Quantum Gravity **35**, 014002 (2017).
[14] M. Breschi, R. O'Shaughnessy, J. Lange, and O. Birnholtz, "Inspiral-merger-ringdown consistency tests with higher modes on gravitational signals from the second observing run of ligo and virgo," (2019), arXiv:1903.05982

[gr-qc].

[15] T. Li, W. Del Pozzo, S. Vitale, C. Van Den Broeck, M. Agathos, J. Veitch, K. Grover, T. Sidery, R. Sturani, and A. Vecchio, Phys. Rev. D **85**, 082003 (2012), arXiv:1110.0530 [gr-qc].

[16] W. M. Farr, B. Farr, and T. Littenberg, *Modelling Calibration Errors In CBC Waveforms*, Tech. Rep. LIGO-T1400682 (LIGO Project, 2015).

[17] B. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt, and et al., Physical Review X **9** (2019), 10.1103/physrevx.9.031040.

[18] J. Aasi, B. P. Abbott, R. Abbott, T. Abbott, M. R. Abernathy, K. Ackley, C. Adams, T. Adams, P. Addesso, and et al., Classical and Quantum Gravity **32**, 074001 (2015).

[19] F. Acernese, M. Agathos, K. Agatsuma, D. Aisa, N. Allemandou, A. Allocca, J. Amarni, P. Astone, G. Balestri, G. Ballardin, and et al., Classical and Quantum Gravity **32**, 024001 (2014).

[20] B. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, V. Adya, and et al., Physical Review Letters **119** (2017), 10.1103/physrevlett.119.161101.

[21] S. Vitale, W. Del Pozzo, T. G. F. Li, C. Van Den Broeck, I. Mandel, B. Aylott, and J. Veitch, Phys. Rev. D **85**, 064034 (2012), arXiv:1111.3044 [gr-qc].

[22] G. S. Kimeldorf and G. Wahba, Ann. Math. Statist. **41**, 495 (1970).

[23] T. B. Littenberg and N. J. Cornish, Physical Review D **91** (2015), 10.1103/physrevd.91.084034.

[24] L. Sampson, N. Yunes, N. Cornish, M. Ponce, E. Barausse, A. Klein, C. Palenzuela, and L. Lehner, Phys. Rev. D **90**, 124091 (2014).

[25] S. Khan, K. Chatziioannou, M. Hannam, and F. Ohme, Physical Review D **100** (2019), 10.1103/physrevd.100.024059.

[26] D. Spiegelhalter, N. Best, B. Carlin, and A. Linde, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **76** (2014), 10.1111/rssb.12062.

[27] D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. Van Der Linde, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **64**, 583 (2002), https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/1467-9868.00353.

[28] F. Pannarale, E. Berti, K. Kyutoku, B. D. Lackey, and M. Shibata, Physical Review D **92** (2015), 10.1103/physrevd.92.084050.

[29] B. D. Lackey, K. Kyutoku, M. Shibata, P. R. Brady, and J. L. Friedman, Physical Review D **85** (2012), 10.1103/physrevd.85.044061.

[30] L. London, S. Khan, E. Fauchon-Jones, C. García, M. Hannam, S. Husa, X. Jiménez-Forteza, C. Kalaghatgi, F. Ohme, and F. Pannarale, Phys. Rev. Lett. **120**, 161102 (2018).

[31] S. Khan, F. Ohme, K. Chatziioannou, and M. Hannam, Physical Review D **101** (2020), 10.1103/physrevd.101.024056.

[32] C. Kimball, C. Berry, and V. Kalogera, Research Notes of the AAS **4**, 2 (2020).

[33] A. Nitz, I. Harry, D. Brown, C. M. Biwer, J. Willis, T. D. Canton, C. Capano, L. Pekowsky, T. Dent, A. R. Williamson, G. S. Davies, S. De, M. Cabero, B. Machenschalk, P. Kumar, S. Reyes, D. Macleod, F. Pannarale, dfinstad, T. Massinger, M. Tpai, L. Singer, S. Khan, S. Fairhurst, S. Kumar, A. Nielsen, shasvath, I. Dorrington, A. Lenon, and H. Gabbard, "gwastro/pycbc: Pycbc release 1.16.4," (2020).

[34] D. Macleod, A. L. Urban, S. Coughlin, T. Massinger, M. Pitkin, paulaltin, J. Areeda, E. Quintero, T. G. Badger, L. Singer, and K. Leinweber, "gwpy/gwpy: 1.0.1," (2020).

[35] J. D. Hunter, Computing in Science & Engineering **9**, 90 (2007).

[36] D. Foreman-Mackey, The Journal of Open Source Software **1**, 24 (2016).