

Systematic errors due to quasiuniversal relations in binary neutron stars and their correction for unbiased model selection

Rahul Kashyap^{1,*}, Arnab Dhani¹, and Bangalore Sathyaprakash^{1,2,3}

¹*Institute for Gravitation and the Cosmos, Department of Physics,
The Pennsylvania State University, University Park, Pennsylvania 16802, USA*

²*Department of Astronomy and Astrophysics, The Pennsylvania State University,
University Park, Pennsylvania 16802, USA*

³*School of Physics and Astronomy, Cardiff University, Cardiff CF24 3AA, United Kingdom*



(Received 14 September 2022; accepted 20 October 2022; published 1 December 2022)

Inference of the equation of state (EOS) of dense nuclear matter in neutron-star cores is a principal science goal of x-ray and gravitational-wave observations of neutron stars. In particular, gravitational-wave observations provide an independent probe of the properties of bulk matter in neutron star cores that can then be used to compare with theoretically derived equations of state. In this paper, we quantify the systematic errors arising from the application of EOS-independent *quasiuniversal relations* in the estimation of neutron star tidal deformabilities and radii from gravitational-wave measurements and introduce a strategy to correct for the systematic biases in the inferred radii. We apply this method to a simulated population of events expected to be observed by future upgrades of current detectors and the next generation of ground-based observatories. We show that our approach can accurately correct for the systematic biases arising from approximate universal relations in the mass-radius curves of neutron stars. Using the posterior distributions of the mass and radius for the simulated population we infer the underlying EOS with a good degree of precision. Our method revives the possibility of using the universal relations for rapid Bayesian model selection of the dense matter EOS in gravitational-wave observations.

DOI: [10.1103/PhysRevD.106.123001](https://doi.org/10.1103/PhysRevD.106.123001)

I. INTRODUCTION

Gravitational wave observations of compact binary coalescences over three observing runs [1–5] of the Advanced Laser Interferometer Gravitational-wave Observatory (LIGO) [6] and Advanced Virgo [7] have determined that binary neutron star mergers are among the most abundant sources of transient gravitational waves in the Universe [8]. Imprinted in the gravitational waves from binary neutron stars (BNS) is the bulk deformation of the stars due to the tidal field of their companions, quantified in terms of their tidal deformability parameter Λ . The measurement of Λ from gravitational-wave observations can provide insight into the thermodynamic properties of the high-density nuclear matter in their cores [9–11] as described by their equation of state (EOS) [12].

A zero temperature EOS is a curve with a functional relationship between the pressure p and energy density ϵ , i.e., $p = p(\epsilon)$. Several formulations of the EOS have been used in the literature such as parametrizations of p as a function of ϵ in the form of piecewise polytropes (see, e.g., [13,14]), spectral representations [15,16], or the speed of

sound [17]. In this work, we use functional relationships among the mass, radius, and tidal deformability of a nonspinning neutron star (NS), obtained by solving the Tolman-Oppenheimer-Volkoff (TOV) equations [18,19], for a given pressure-density curve $p(\epsilon)$, as a mathematical model for an EOS. The magnitude of the tidal deformability is determined by the neutron star mass and their EOS.

Tidal effects alter the orbital dynamics of a binary neutron star and hence the emitted gravitational waves. The phase evolution of gravitational waves depends predominantly on a certain linear combination of the individual tidal deformabilities Λ_i of the companion stars called the *reduced tidal deformability* $\tilde{\Lambda}$ [see Eq. (9) for a definition of $\tilde{\Lambda}$ in terms of individual tidal deformabilities]. The reduced tidal deformability is a measure of the sum of the individual tidal deformabilities and enters the phase evolution of gravitational waves at the fifth post-Newtonian (PN) order [9]. The dual parameter $\delta\tilde{\Lambda}$, which is a measure of the difference in individual tidal deformabilities, enters the phase evolution at a higher, sixth PN order [9,11,20,21]. The high PN orders imply that the tidal effects contribute significantly to the phase evolution at gravitational-wave frequencies greater than about 100 Hz [22] and most significantly just before the two neutron stars merge at frequencies of 1–1.5 kHz,

*rkk5314@psu.edu

depending on the EOS [23]. The noise spectral density of the current detector network is at its lowest around 200 Hz and raises quadratically at larger frequencies [24]. Consequently, there is currently no hope of accurately measuring both the parameters but only the reduced tidal deformability, that too with significant errors. While future observatories might measure symmetric mass ratio [$\eta = m_1 m_2 / (m_1 + m_2)^2$], chirp mass [$\mathcal{M} = \eta^{3/5} (m_1 + m_2)$] and $\delta\tilde{\Lambda}$ better than now, the sensitivity will not be good enough to accurately measure both of the tidal parameters (see, e.g., Smith *et al.* [25]). The individual tidal deformabilities of the component NSs can be determined only if both $\tilde{\Lambda}$ and $\delta\tilde{\Lambda}$ are measured. In the absence of such measurements, Yagi and Yunes [26] found approximate correlations between combinations of the individual tidal deformabilities that depend only on the mass ratio of the companion stars and are largely independent of the equations of state (EOSs). These quasiuniversal relations can be used to infer the individual tidal deformabilities Λ_k , ($k = 1, 2$) from a measurement of $\tilde{\Lambda}$ alone.

In addition to the above universal relations, Damour and Nagar [27] discovered a strong sensitivity of the quadrupolar tidal deformability on NS compactness $\mathcal{C}_k = GM_k/c^2$, irrespective of the properties of the nuclear matter making up the cores. New universal relations have been proposed recently by Saes and Mendes [28], where the ratio of central pressure and densities are found to be correlated with the compactness. It should be pointed out that these universal relations are also “quasi” in nature because the correlations among the various quantities are only approximately true across the different EOS models.

The universal relations are currently one way to deduce the radii of component NSs without making any assumption about the underlying EOS. This is because the measurement of the radius requires knowledge of the EOS which is currently unknown. Universal relations were used to infer the radii of companion neutron stars in GW170817 [1,2,29]—the first BNS merger ever observed. The residuals on the systematic errors due to the approximate nature of the universal relations were marginalized, assuming a Gaussian distribution of the residuals [30–33]. Kumar and Landry [31] construct their own universal relation for combining constraints obtained from multiple events. We note an important limitation of model selection based on universal relations. A set of TOV sequences for individual EOSs were used to find quasiuniversal features. It, then, seems contradictory to calculate the evidence for the same set of models using the thence inferred quantities. The resolution of such a fallacy is that the use of universal relations is only as good as the fits for them which are 20% and 2% accurate for the universal relations given in Eqs. (8a) and (8b), respectively [34]. The procedure of marginalizing over the residuals was meant to alleviate some of the systematic errors present in the universal relations. However, this process is not infallible since marginalization can bring its own set of systematic errors

in the presence of any non-Gaussian feature in the residuals. Moreover, the residuals are not random numbers, they are known quite precisely. Hence, assuming that residuals are described by a normal distribution and marginalizing over it can potentially introduce systematic errors. Abbott *et al.* [35] avoid these systematic biases sampling tidal deformability of both components independently without using universal relations. The inferred posteriors of the mass-tidal deformability of both neutron stars were then compared with a wide collection of EOSs.

A more direct approach to model selection is to parametrize the $p = p(\epsilon)$ curve to obtain the posterior in the space of an assumed set of (e.g., piecewise polytropic, spectral or nuclear) parameters, capable of describing a wide variety of *ab initio* zero temperature EOS models. The posterior probability of these parameters is inferred for a set of events, which are then combined to develop an effective model selection method. Lackey and Wade [36] have used a piecewise polytropic method while Wade *et al.* [21] argue that spectral parametrization provides a better constraint when stacking multiple events in model selection. Biswas [33], on the other hand, used a set of nuclear parameters to constrain and combine events for the same purpose.

While we break the degeneracy contained in the posterior probability density function (PDF), $p(\mathcal{M}, \eta, \tilde{\Lambda})$ before model selection, one can also obtain the evidence of an EOS model directly from such a PDF. Among such approaches, Pacilio *et al.* [37] uses the joint posterior PDF to calculate the Bayes factor for each EOS model against a particular EOS by integrating the TOV equations for each value of central densities in a prior range. Ghosh *et al.* [38] follows the method similar to [35] but, in the space of $(\mathcal{M}, \eta, \tilde{\Lambda})$ rather than $(m_1, m_2, \Lambda_1, \Lambda_2)$. However, we believe that there are larger degeneracies among model EOSs that might be affecting the approach to model selection using the effective tidal deformability compared to breaking the degeneracy to individual tidal deformabilities.

In this study, we will describe a different and improved method of taking the systematic errors of universal relations into account for model selection which works equally well for all binary configurations. Our proposal is as follows. We recognized that at the time of model selection, we have an estimate for the component masses and the assumption of a fiducial model. This gives us the expected distribution of the tidal/radial parameters. Together, given a measurement of the tidal/radial terms, we do a piecewise shift of the samples using the precisely known residuals with respect to the fiducial model to construct an unbiased inference. The evidence in favor of a model is calculated from this distribution. We find that if systematic errors are unaccounted for, it can lead to an incorrect model being preferred. Predictably, correcting for them results in an unbiased model selection.

The rest of this paper is organized as follows. In Sec. II, we describe the BNS population assumed in this study and

present the distributions of errors in key inferred parameters for one of the samples from our simulation. In Sec. III, we describe the universal relations and the EOS-agnostic analysis pipeline that is currently used in the study of dense matter EOSs. We also show that the use of universal relations in inferring neutron star tidal deformabilities and radii lead to systematic biases and propose a simple algorithm to correct for the bias during model selection. Section IV summarizes the main results of this study concluding that the next-generation gravitational-wave observatories will have the ability to precisely measure the dense matter EOS. Section V concludes with a brief summary of the paper and future directions.

II. NETWORKS AND POPULATION

In this section, we will outline the two gravitational-wave detector networks considered in this study and describe the population of BNS mergers accessible to them. We will then discuss the parameter estimation capabilities of these networks for gravitational waves from binary neutron star mergers, and in particular, how well they can measure the chirp mass, mass ratio, and tidal deformability that are relevant to measuring the dense matter EOS.

A. Networks

We consider two different detector networks in this study and they are acronymed and elucidated as follows:

- (i) *HLVKI+*: This is a planned global network of five gravitational-wave detectors operating at A+ sensitivity. It consists of three LIGO detectors [39] at Hanford (H), Livingston (L), and India (I) operating at A+ sensitivity, the Virgo (V) detector at AdVirgo+ [40] sensitivity, and the KAGRA detector at KAGRA+ [41] sensitivity.
- (ii) *ECS*: This is a proposed next-generation gravitational-wave detector network consisting of two Cosmic Explorer detectors [42], one in the U.S. and the other in Australia, each with an arm length of 40 km, optimized for low frequency, and an Einstein Telescope [43] in Europe.

Additional details about the networks such as the technologies to be used in them and their sensitivities to different populations of compact binaries can be found in Borhanian and Sathyaprakash [44] and the references therein. We do not consider the detector noise realizations which could further impact the recovery of correct parameters.

B. Population characteristics

1. Redshift distribution

We simulate a population of BNS mergers up to a redshift of $z = 1$. The redshift distribution of the population is given by

$$p(z) = \frac{R_z(z)}{\int_0^{10} R_z(z) dz}, \quad (1)$$

where $R_z(z)$ is the merger rate density in the observer frame and can be expressed as

$$R_z(z) = \frac{R_m(z)}{1+z} \frac{dV(z)}{dz}. \quad (2)$$

Here $dV(z)/dz$ is the comoving volume element and $R_m(z)$ is the merger rate per comoving volume in the source frame which, in turn, is assumed to be proportional to the star formation rate (SFR), $R_f(t)$ and takes the form [45]

$$R_m(z) = \int_{t_{\min}}^{t_{\max}} R_f(t(z) - t_d) P(t_d) dt_d, \quad (3)$$

where the t_z is given by

$$t(z) = \frac{1+z}{H_0} \int_z^{z_f} \frac{dz'}{E(z')}. \quad (4)$$

This equation signifies that the binaries that form at time $t(z) - t_d$ merge at time $t(z)$ (i.e., redshift z) after a delay time t_d . Here, we choose the cosmic SFR to follow Vangioni *et al.* [46]. The probability distribution for the time to coalesce for a binary after formation is taken to be $P(t_d) \propto 1/t_d$ [47] with a minimum merger time of $t_{\min} = 20$ Myr and a maximum of $t_{\max} = 1/H_0 = 14.4$ Gyr using $H_0 = 67.7$ km s⁻¹ Mpc⁻¹ [48] ($\approx z = 30$). Using the local merger rate $R_m(z=0)$ from the second LIGO-Virgo Gravitational-Wave Transient Catalog, *GWTC-2* [49] to be

$$R_m(z=0) = 320 \text{ Gpc}^{-3} \text{ yr}^{-1}, \quad (5)$$

TABLE I. The number of BNS mergers with an SNR greater than a threshold of 10, 30, and 100 out of a simulated population of 80,000 events expected to be detected annually by next generation GW detectors. We show the number of events in the two detector networks and the three EOS considered.

EOS	Network		
		HLVKI+ events per year	ECS events per year
SNR = 10	ALF2	260	74304
	APR3	327	75868
	APR4	299	75191
SNR = 30	ALF2	10	22565
	APR3	19	27620
	APR4	9	25252
SNR = 100	ALF2	...	638
	APR3	...	846
	APR4	...	767

we find the total number of BNS mergers up to $z = 1$ comes out to be approximately 80,000 per year and we take this to be the annual rate.

2. Equations of state simulated

We simulate the gravitational-wave signal for our population of BNS for three different EOSs: ALF2 [50], APR3, APR4 [51]—with sufficiently different mass-radius curves.

These three EOSs are representative of different regions of the mass-radius space as well as a range of maximum masses for NSs as shown in Fig. 9. In addition to these we use several other EOSs from literature (BHB [52], DD2 [53], H3 [54], H4 [54], LS220 [55], SFHo [56], SLy [57]) for the model selection process to be discussed in Sec. III. Two EOSs have been constructed by randomly selecting piecewise polytropic indices from [58] marked here as PP2 and PP5.

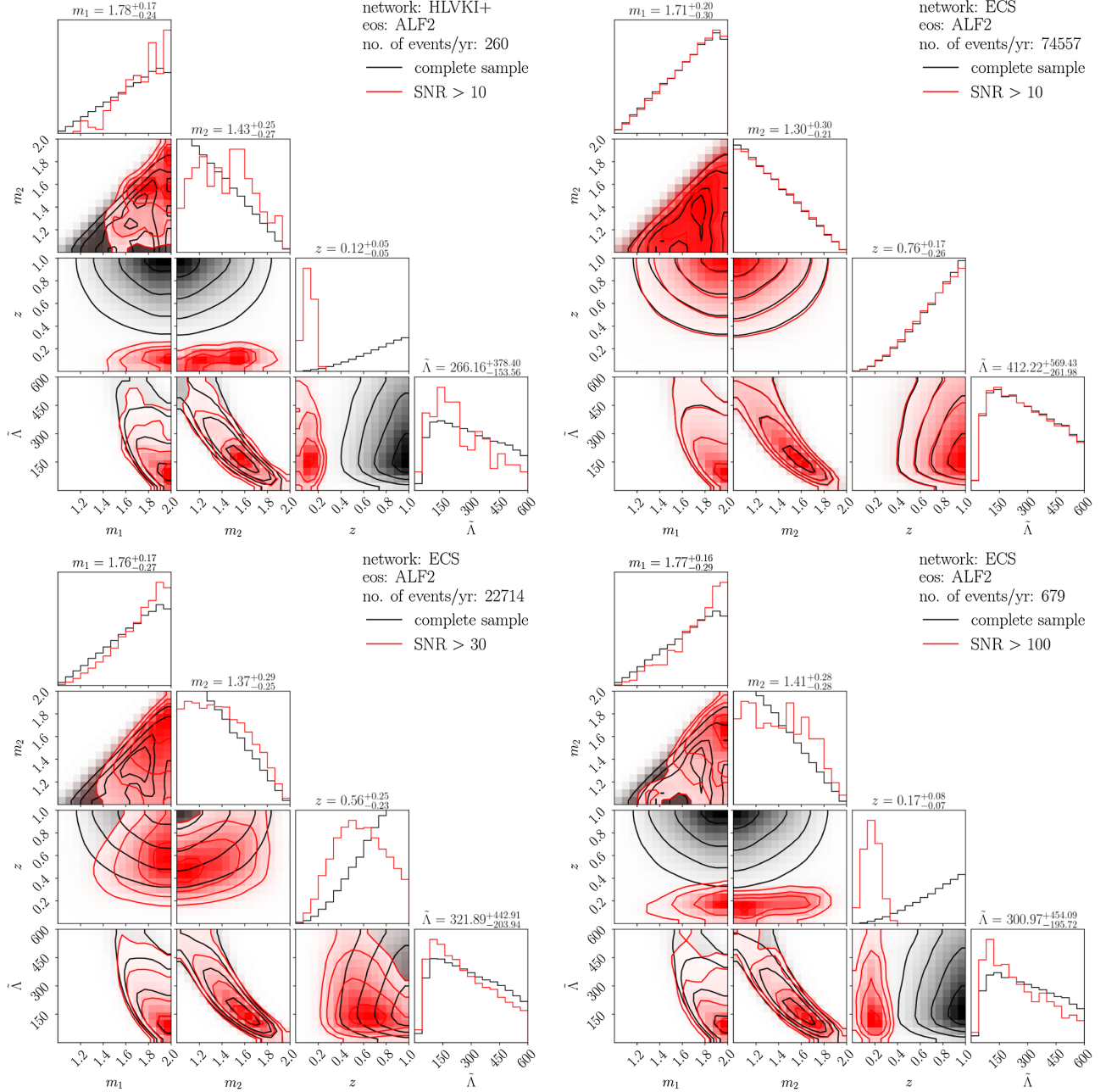


FIG. 1. Distribution of selected parameters in our population for different SNR thresholds in the two detector networks considered. These are the subpopulations that are used in the bias correction analysis. The HLVKI+ network is not expected to observe a significant number of events above an SNR of 30. The median values for the observed population are also quoted on top of each column.

3. Distributions of intrinsic and extrinsic parameters

We consider the individual neutron stars to be non-spinning and distributed uniformly in masses between $1 M_\odot$ and the maximum mass allowed by the corresponding EOS. The extrinsic parameters—cosine of the inclination angle $\cos \iota$, location of the source in the sky (cosine of the declination angle $\cos \delta$ and right ascension α), polarization angle ψ , and the phase of coalescence ϕ_0 , of the fiducial BNS population are drawn from a uniform distribution across their domains. The luminosity distances are calculated from the redshifts of the binaries using *Planck18* [59] cosmology.

In Table I, we show the number of BNS mergers per year above a certain signal-to-noise ratio (SNR) for our simulated populations in the two detector networks and the three EOSs considered in this study [44]. The variation of the detection rate across different EOS is directly related to the maximum mass allowed by the corresponding EOS. Note that for an SNR threshold of 10, which is a typical detection criterion for an event, the next-generation network detects almost all BNS mergers within a redshift of $z = 1$ [44]. In Fig. 1, we also depict the distributions of various parameters of the binary given a threshold SNR. This shows the parameter space probed by the observed population compared to the astrophysical distribution of sources.

C. Parameter estimation

We simulate gravitational waves from the population of binaries distributed as detailed in the present section using the TaylorF2-tidal waveform model [21]. The signal spans a

frequency range of $[f_{\min}, f_{\max}]$ where $f_{\min}$ is 5 Hz for ECS while 10 Hz for A+ detectors and $f_{\max} = \min(f_{\text{ISCO}}, 1024 \text{ Hz})$. $f_{\text{ISCO}} = 1/(6^{3/2}\pi M)$ is the orbital frequency of the innermost stable circular orbit for a point particle in an effective black hole spacetime, where M is the total mass of the binary [60]. The signal is truncated at 1024 Hz because signals at higher frequencies do not contribute to the SNR [44].

The errors on the parameters of the gravitational-wave signal are calculated using the Fisher approximation for the likelihood of a signal using the publicly available code GWBENCH [61]. The Fisher matrix is defined using frequency domain inspiral gravitational waveform, $h(f)$, as

$$\mathcal{F}_{ij} = \left\langle \frac{\partial h(f)}{\partial \theta^i}, \frac{\partial h(f)}{\partial \theta^j} \right\rangle \quad (6)$$

with a vector in this space given by $\vec{\theta} = (\mathcal{M}, \eta, \tilde{\Lambda}, D_L, \iota, \alpha, \delta, \psi, \phi_0, t_c)$ where $\mathcal{M}$, η , $\tilde{\Lambda}$, D_L and t_c are the chirp mass, symmetric mass ratio, combined tidal deformability, luminosity distance and the time of coalescence, respectively, and the other parameters are defined in the previous section. The covariance matrix for the errors on these parameters is then $C_{ij} = \mathcal{F}_{ij}^{-1}$. The inner product in the above expression is given by

$$\langle a(f), b(f) \rangle = 2 \int_{f_{\min}}^{f_{\max}} \frac{\tilde{a}(f)^* \tilde{b}(f) + \tilde{a}(f) \tilde{b}(f)^*}{S_n(f)} df, \quad (7)$$

where $S_n(f)$ is the noise spectral density of the detectors.

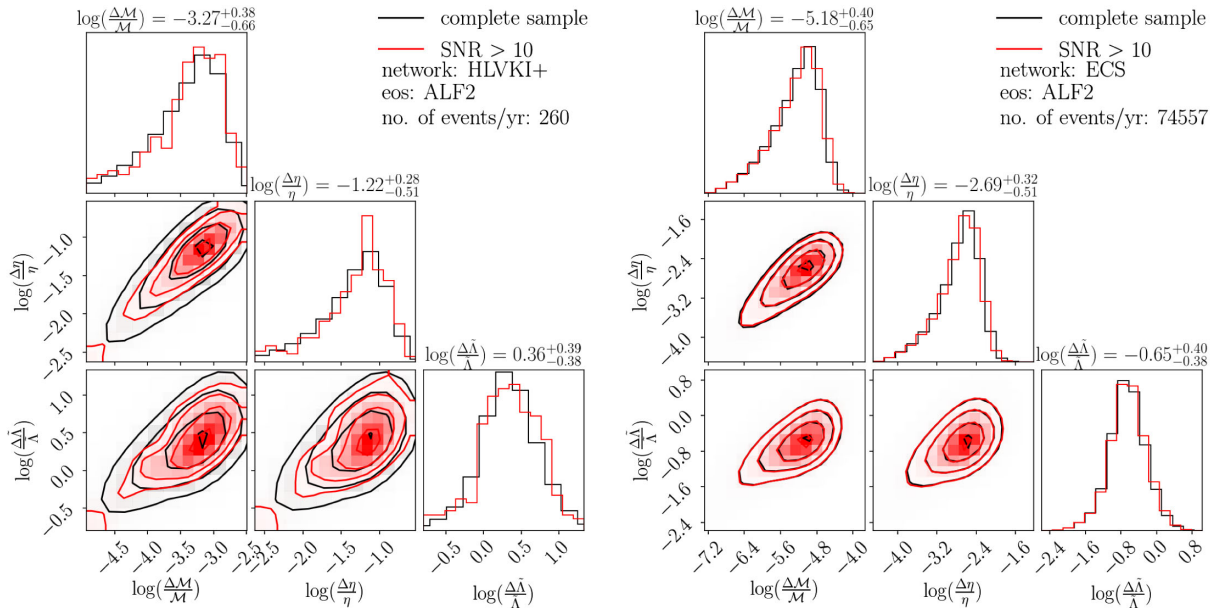


FIG. 2. Distribution of relative errors on log scale for the key quantities for our population of BNS systems with ALF2 EOS. The full population is depicted in black while the sample with SNR greater than 10 is shown in red. The right panel contains the results for ECS while the left panel is for the HLVKI+ network. The median values of error distributions in each of the quantities are mentioned on the top of each column.

In Fig. 2, we show the normalized distribution of 1σ statistical errors for our population in the parameters $\mathcal{M}$, η , and $\tilde{\Lambda}$. The error distributions are shown for only these parameters because we will use them to calculate the individual tidal deformabilities and radii of the NSs. Furthermore, only the population having ALF2 as its EOS is shown and they are similar for the other EOS. The results for the full population are shown in black while the subpopulation with a minimum SNR of 10 is shown in red. The right panel shows the capabilities of the next-generation network of ECS whereas the left panel depicts the HLVKI+ network. The 1D histograms show that the error distributions for the detected population and the full population are similar in the parameters shown with the median errors for the detected population tabulated above the histograms corresponding to the relevant parameters. This implies that the errors on these parameters are weak functions of the SNR of the signal, which is used to demarcate a detection from a nondetection. The ellipses show how the errors on the different parameters are correlated. As expected, we see a positive correlation between the error distributions of the intrinsic parameters signifying that if one of the parameters is well measured then so are the others. Nevertheless, it can also be seen that the correlation of $\mathcal{M}$ and η with $\tilde{\Lambda}$ is less than that between them. This is because the $\tilde{\Lambda}$ errors are determined by the high-frequency sensitivity of the network while the $\mathcal{M}$ and η errors are primarily controlled by the low-frequency sensitivity [23,62]. Note that the ellipses are not the correlations among the different parameters for a given event but rather the distribution of the errors for a population of events.

III. USE OF QUASIUNIVERSAL RELATIONS IN MODEL SELECTION

This section begins by discussing the two universal relations that relate the symmetric and asymmetric combinations of the tidal deformabilities on the one hand and the compactness and tidal deformability on the other. This is followed by a brief outline of the current method for inferring the tidal deformabilities and radii of individual neutron stars from gravitational-wave observations [29]. We point out how this approach can lead to biased estimation of the tidal deformability parameters and the radii due to systematic errors in the universal relations and hence lead to erroneous EOS model selection. For the ECS network of next-generation observatories, the systematics can dominate over statistical uncertainties even for individual events. For the HLVKI+ network, the systematics for individual events are smaller than statistical uncertainties; however, model selection with a population of 30 events or more can be biased even for this network. We describe how to rectify systematic uncertainties at the time of model selection and introduce several statistical

measures to show that the bias-corrected estimates of the tidal deformability and radius converge to the correct EOS. Marginalizing over the errors due to universal relations can account for the biases but that comes at the expense of increased errors in the inferred quantities.

A. Universal relations and residuals

1. Universal relation between tidal deformabilities of a pair of neutron stars

The structure of neutron stars is determined by their nuclear EOS via the TOV equations [18,19]. The EOS of neutron stars is currently unknown and there are numerous models describing the pressure-density (equivalently, mass radius) curves of neutron stars (for a review see Lattimer and Prakash [63] and references therein). Although x-ray and gravitational-wave observations severely constrain the family of viable EOS models, many of them are still consistent with data.

First universal relation.—In spite of the huge variation in the relationship between the masses and radii of neutron stars amongst different EOS models, Yagi and Yunes [34] found the remarkable result that the asymmetric combination Λ_a of the tidal deformabilities of two neutron stars, defined by $2\Lambda_a \equiv (\Lambda_2 - \Lambda_1)$, is uniquely related to the symmetric combination, Λ_s , defined by $2\Lambda_s \equiv (\Lambda_2 + \Lambda_1)$, depending only on the ratio of their masses $q \equiv m_2/m_1 \leq 1$. This universal relation is given by [30,64]

$$\Lambda_a = F_n(q)\Lambda_s \frac{1 + \sum_{i,j}^{3,2} b_{ij}q^j/\Lambda_s^{i/5}}{1 + \sum_{i,j}^{3,2} c_{ij}q^j/\Lambda_s^{i/5}}, \quad (8a)$$

$$F_n(q) = \frac{1 - q^{10/(3-n)}}{1 + q^{10/(3-n)}}, \quad (8b)$$

where Λ_1 and Λ_2 are the individual tidal deformabilities of the companion stars, b_{ij} , c_{ij} and n are the fitting parameters given in Table 3 of Yagi [64] (see also Chatzioannou *et al.* [30] and Godzieba *et al.* [58]).

Residuals in the first universal relation.—The universal relation between Λ_a and Λ_s is plotted in the top left panel of Fig. 3 for three different values of the mass ratio $q = 0.5$, 0.8 and 0.9. Solid and dot-dashed curves use the fitting formulas in Yagi [64] and Godzieba *et al.* [58], respectively. The two versions of the universal relations agree with each other pretty well over a wide range of Λ_s for $q = 0.5$ but less so for larger masses (smaller Λ_s) and larger mass ratios.

The bottom left panel plots the residual of the fits with respect to the exact TOV sequences for the EOS models used in this work. The residuals remain below $\sim 10\%$ for many EOS models and over a wide range of Λ_s ; the fits of Godzieba *et al.* [58] have smaller residuals than those of Yagi [64] but can still be as large as 25% when the tidal

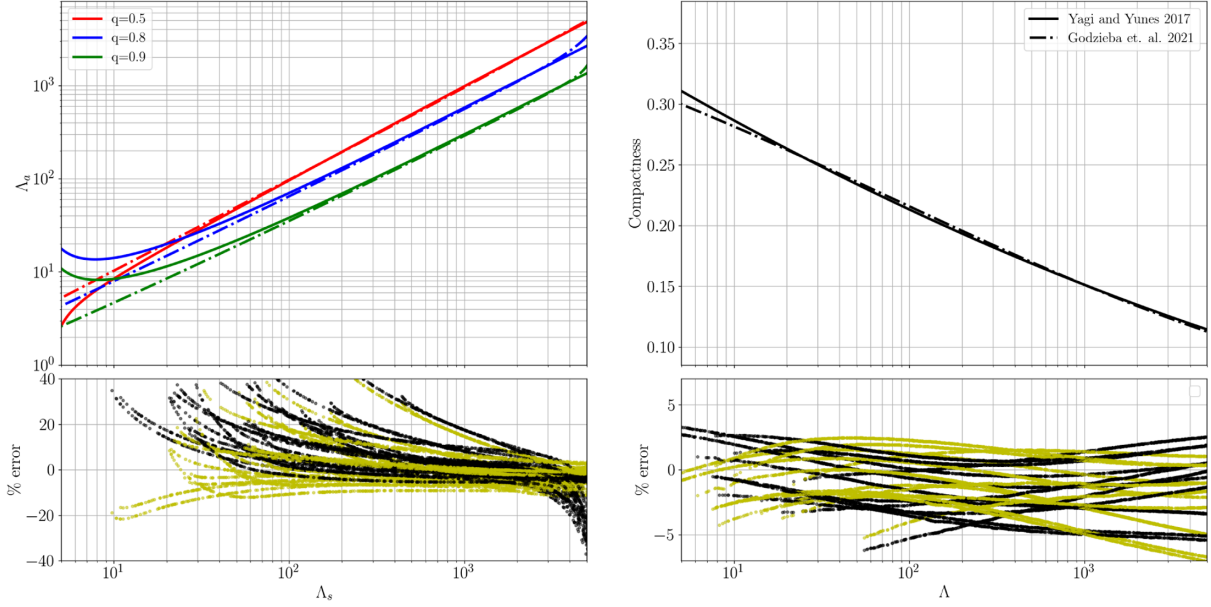


FIG. 3. The top left panel plots the universal relation between the asymmetric and symmetric combinations, Λ_a and Λ_s , respectively, of the individual tidal deformabilities Λ_1 and Λ_2 for several EOS as constructed in Yagi and Yunes [34] and Godzieba *et al.* [58]. The (small) difference in the two versions of the universal relations is due to the error in the fitting function and the EOS data. In the lower left panel, we plot the deviation of the universal relations from the TOV sequences corresponding to the EOS models used in this work provided by Godzieba *et al.* [58] (in black) and Yagi and Yunes [34] (in yellow), respectively. These residuals can be as large as 25% for the largest and smallest tidal deformabilities (Λ_s) corresponding to smallest and largest neutron star masses, respectively. The top right panel plots the second universal relation between a neutron star's compactness C and its tidal deformability Λ as constructed by the same two authors [34,58]. Please note that the left panel shows universal relations for a pair of NSs while the universal relation shown in the right panel is for single NSs. We use these exactly known deviations to correct the radius and tidal deformability posteriors at the time of model selection.

deformabilities are small, corresponding to heavier neutron stars and/or softer EOSs.

Individual tidal deformabilities.—In spite of the residuals, Eq. (8) is still very useful in inferring the individual tidal deformabilities and hence assist in the process of EOS model selection. The PN expansion of the gravitational-wave phase contains the individual tidal deformabilities as linear combinations in the PN coefficients at the fifth and sixth PN orders [i.e., corrections in the phase evolution at orders $\mathcal{O}(v/c)^{10}$ and $\mathcal{O}(v/c)^{12}$ beyond the leading order quadrupole term], respectively. Tidal effects are encoded in parameters $\tilde{\Lambda}$ and $\delta\tilde{\Lambda}$ defined by

$$\tilde{\Lambda} = \frac{1}{26} \left[(1 + 12q)(\Lambda_s - \Lambda_a) + \left(1 + \frac{12}{q}\right)(\Lambda_s + \Lambda_a) \right] \quad (9a)$$

$$\delta\tilde{\Lambda} = \sqrt{1 - 4\eta} \left(1 - \frac{13272}{1319}\eta + \frac{8944}{1319}\eta^2 \right) \Lambda_s + \left(1 - \frac{15910}{1319}\eta + \frac{32850}{1319}\eta^2 + \frac{3380}{1319}\eta^3 \right) \Lambda_a. \quad (9b)$$

The high PN orders at which they appear imply that the effect of $\tilde{\Lambda}$ is only important for frequencies $f_{\text{GW}} \gtrsim 100$ Hz and that of $\delta\tilde{\Lambda}$ at even larger frequencies [22] (see also

Dietrich *et al.* [23]). Thus, only the final few cycles of the waveform before the merger contain significant tidal effects and only $\tilde{\Lambda}$ is measurable to a good accuracy even when the SNRs are ~ 100 [25]. Thus, there is no hope of inferring the individual tidal deformabilities from gravitational-wave observations alone. But universal relation Eq. (8) can be of help albeit inferred Λ values will be biased.

Systematic errors in Λ .—Suppose $(q, \tilde{\Lambda})$ are known exactly. We can numerically solve the pair of Eqs. (8) and (9a) for Λ_s and Λ_a and hence infer the individual tidal deformabilities Λ_k , $k = 1, 2$. The numerical inversion algorithm is found to be accurate to a relative error of $\sim 10^{-8}$ while the interpolation of TOV tables are accurate to better than $\sim 10^{-4}$. Since the universal relations are not exact, however, the inferred values of Λ_k will not be the same as the ones that went into computing $\tilde{\Lambda}$. We assess the systematics incurred using a population of BNS systems uniformly sampled in the $m_1 - m_2$ plane.

For each pair of masses, and for a given EOS, we find Λ_1 and Λ_2 by solving the TOV equations using the TOV solver developed in Damour and Nagar [27], Bernuzzi and Nagar [65],¹ which are then used to obtain $\tilde{\Lambda}$. For a given EOS, we

¹<https://bitbucket.org/bernuzzi/tov/src/master/>.

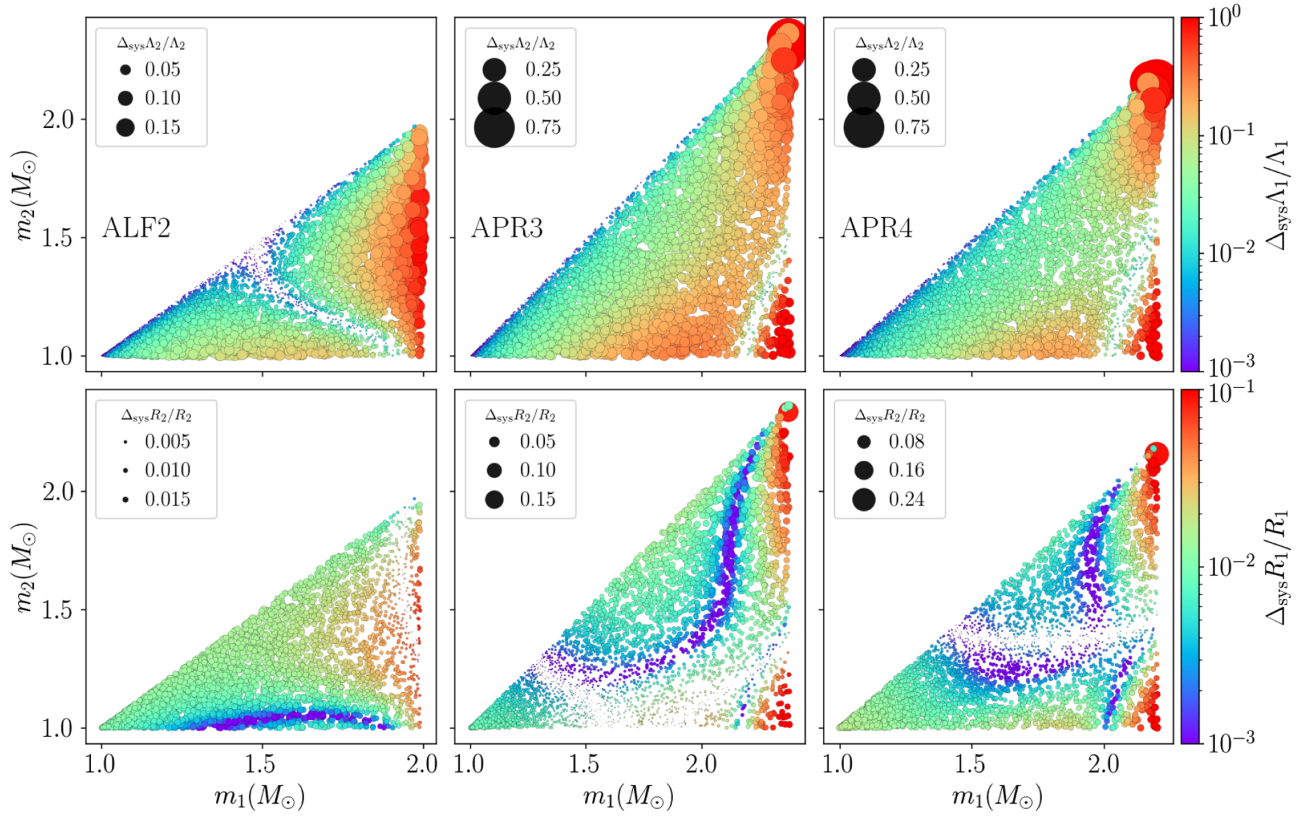


FIG. 4. This figure shows the relative error in the reconstruction of individual tidal deformabilities of NSs in the space of component masses of BNS systems for three example EOSs. The component masses are sampled uniformly over the range allowed by the respective EOS without imposing any restriction on the mass ratio. Colors represent the systematic error in the tidal deformability of the primary companion (i.e., the heavier NS) while the size of the circle represents the systematic error in the tidal deformability of the lighter companion. The errors are greater for more massive NSs and softer EOS.

solve TOV equations for a discrete set of masses to obtain the tidal deformabilities and radii and interpolate the solutions to obtain these parameters for arbitrary NS masses. The interpolated values of tidal deformabilities and radii agree with the exact numerical solution of TOV equations to within a fractional error of 10^{-4} , which, as we shall see below, is far smaller than the systematic errors due to universal relations.

For each binary pair in the population, the bottom panels of Fig. 4 show the fractional difference in the true and reconstructed tidal deformabilities $|\Delta_{\text{sys}}\Lambda_{1,2}|/\Lambda_{1,2}$, the colorbar representing the systematic error in Λ_1 (i.e., the tidal deformability of the heavier companion) and the size of the circles representing the systematic error in Λ_2 (i.e., the tidal deformability of the lighter companion). The biases are shown for three example EOSs: ALF2, APR3, and APR4, corresponding to increasingly softer EOS from left to right. The biases are particularly large ($\sim 75\%$) for softer EOS and heavier NSs whose Λ values are around few tens to hundreds. For intermediate masses and stiffer EOSs, the errors are lower but could still be few to ten percent. This is consistent with the fact that the fit residuals of the universal relation (8) are smaller for stiffer EOSs [34].

Thus, while the tidal deformabilities obtained using universal relations are EOS agnostic the inferred values are biased. Consequently, a model selection algorithm that compares the mass-tidal deformability curve obtained from a collection of BNS events with the ones obtained from solving TOV equations could lead to a greater evidence for an incorrect EOS model.

2. Universal relation between tidal deformability and compactness of a single neutron star

The first universal relation helps infer the individual tidal deformabilities of neutron stars from gravitational-wave observations but this does not help in inferring their radii. This is because the tidal deformability of a neutron star is related to its radius via

$$\Lambda \equiv \frac{2}{3}k_2 \left(\frac{R}{m}\right)^5, \quad (10)$$

where k_2 , R and m are the star's tidal Love number, radius and mass [9]. A knowledge of the Love number k_2 , also determined by the EOS, is necessary to deduce the radius of

a neutron star from its mass and tidal deformability. Although k_2 varies quite a bit from one EOS to another for a given neutron-star mass, there seems to be a second universal relation [27,34] that connects Λ to the star's compactness defined by $\mathcal{C} \equiv GM/(Rc^2)$:

$$\mathcal{C}(\Lambda) = \sum_{k=0}^2 a_k (\ln \Lambda)^k, \quad (11)$$

where a_k s are the fitting coefficients in Godzieba *et al.* [58]. The right panel of Fig. 3 plots this second universal relation constructed in Godzieba *et al.* [58] and Yagi and Yunes [34]. Both versions agree with each other pretty well over a wide range of tidal deformability. The residual between the universal relation and the true dependence of the compactness on the tidal deformability for the collection of EOS considered in this work (bottom right panel) is at most 5% over the entire range of Λ . This error is tolerable given the large measurement uncertainties of reduced tidal deformability expected in the near future (i.e., for the HLVKI+ network). However, an error of 5% is too large for the ECS network which has the potential to measure the radii and deformabilities at the subpercent level.

This second universal relation, together with the posterior probabilities of the component masses and tidal deformabilities, allows the construction of the posterior probabilities of the radii. While the Λ posteriors are only affected by the systematics in the first universal relation, the radii posteriors are affected by the systematics of both universal relations.

In the next section, we summarize the analysis pipeline used to infer the mass-tidal deformability and mass-radius curves using the two universal relations. A similar pipeline was applied to GW170817 to obtain EOS-independent posterior distribution of masses and radii [29].

B. Inference of tidal deformability parameters and neutron star radii without the residuals

The current approach for the calculation of EOS-agnostic radii of neutron stars from gravitational-wave observations of BNS mergers, as described in Abbott *et al.* [29], involves the use of the two universal relations given in Eqs. (8) and (11). We developed an alternative, but equivalent, approach, the flowchart for which is shown in Fig. 5.

- (1) In the first step, posterior distributions are obtained for the chirp mass, symmetric mass ratio and $\tilde{\Lambda}$ for a catalog of expected events as shown in the box at the top left of the flowchart. To this end, we use the GWBENCH implementation of Fisher information matrix [61] for fast computation of the posteriors but in a real data analysis problem posteriors would be obtained using a Bayesian inference algorithm.
- (2) From the posterior distribution of the chirp mass and the symmetric mass ratio we can derive the posterior

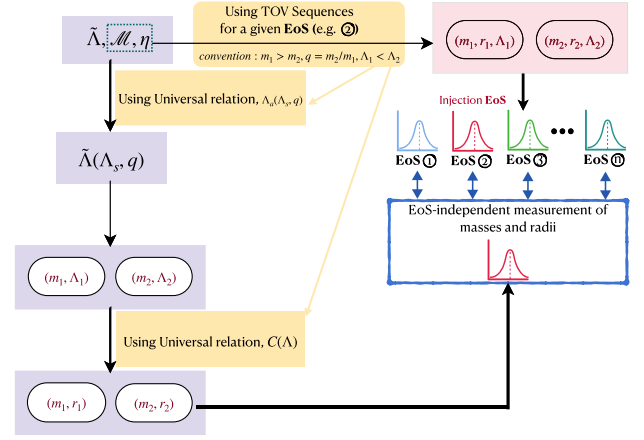


FIG. 5. This is a sketch of the data analysis pipeline. The posterior samples for $M_c, \eta, \tilde{\Lambda}$ are created using the Fisher information matrix which is in the left-uppermost box. Yellow boxes represent information coming out from the construction of TOV sequences using multiple choice of EOS. We use quasi-universal relations, (8) and (11) to obtain the component mass-radii posterior probability distribution, which is then compared to a set of EOS models as represented by a vertical double arrow.

distribution of the companion masses m_1 and m_2 and the mass ratio q .

- (3) In the next step (first downward arrow from top left), the first universal relation Eq. (8a) is used to eliminate Λ_a from Eq. (9a) to arrive at an expression for $\tilde{\Lambda}$ that depends only on the symmetric combination of the tidal parameters and the mass ratio q , namely $\tilde{\Lambda} = \tilde{\Lambda}(\Lambda_s, q)$.
- (4) The transcendental equation $\tilde{\Lambda} = \tilde{\Lambda}(\Lambda_s, q)$, together with the posterior distributions of q and $\tilde{\Lambda}$, is then solved to obtain the posterior of Λ_s and the first universal relation is deployed once again to derive the posterior of Λ_a . From the Λ_s and Λ_a posteriors, it is straightforward to deduce the posteriors of the individual tidal deformabilities (second downward arrow from top left).
- (5) In the next step, the second universal relation Eq. (11), together with the m_1 and m_2 posteriors, is used to infer the posteriors of the neutron star radii (third downward arrow from top left).
- (6) The above steps essentially give us posteriors in the mass-tidal deformability ($m - \Lambda$) or mass-radius ($m - R$) plane (bottom, horizontal, upward turning arrow).
- (7) Given an EOS, the mass-radius and mass-tidal deformability curves can also be obtained by solving the TOV equations (top horizontal arrow) and can be compared with the posterior distributions.

The above pipeline essentially compares the measured parameters (e.g., the mass-radius curve) with the predictions of a set of models to choose the best EOS consistent with the data. However, the systematic biases in the tidal parameters

(arising from the approximate nature of the first universal relation) and radii (arising from both the first and second universal relations) could favor the wrong model. We next discuss a strategy to remedy the biases by incorporating the residuals as part of the universal relation. We differ from Ref. [35] in that we do not sample the tidal deformability of the companions but instead only the reduced tidal deformability. Additionally, instead of marginalizing over the residuals while inferring radii and individual tidal deformabilities we correct for them at the time of model selection, which differs in one key aspect from Ref. [29]. Our method is equivalent to calculating the $\Lambda_s - \Lambda_a$ relation for each EOS.

C. Mitigating systematic errors in the tidal deformability and radius

In this section, we elucidate our method to correct for the systematic biases incurred in the estimation of the individual tidal deformabilities and radii due to the use of quasiuniversal relations. The biases, being systematic and not statistical, do not asymptote to zero with increasing SNR but become the dominant source of error as the statistical errors decrease.

Using residuals to correct systematic biases.—Our proposal is to correct for systematic errors at the time of model selection. To do so, we begin with a specific EOS model and compare its predictions of the mass- Λ or mass-radius curve with the posterior distributions of the same but obtained from gravitational-wave observations, typically computing the χ^2 or Bayesian evidence for the model given the data. At this point, we know not only the *prior probability* for the model but also the *residuals of the model* with respect to the universal relations. We can, therefore, subtract the known residuals for the model from the measured tidal deformabilities and radii before calculating the χ^2 or evidence for the model. We emphasize that the residuals *cannot* be used to obtain EOS-agnostic tidal deformabilities or radii as the corrections are specific to the EOS model chosen.

In more detail, we start with an EOS-agnostic estimate as calculated using the procedure outlined in the previous section. Given an EOS model and the measured masses from gravitational wave observations, one can infer the corresponding distributions of Λ s and radii using the $m - \Lambda$ or the $M - R$ curves for the EOS model. In the presence of systematic biases these curves would shift from those inferred using universal relations, which also broaden due to the statistical uncertainties in their measurements. Depending on the SNR of an event, statistical uncertainties can be larger than systematic biases. In that case, correcting for the systematics would not significantly improve the evidence for the correct model. But as we will see, the ECS network will observe events where such a correction would be critical. Moreover, as mentioned before, biases

accumulate as the evidences from multiple events are combined and, therefore, the correction would be important even for the less sensitive HLVKI+ network.

Bias correction for a specific event.—In Fig. 6 we show the impact of bias correction for a fiducial event in our population. Here, an event with the ALF2 EOS is chosen from the ECS network. The component masses of the binary system are $m_1 = 1.63 M_\odot$ and $m_2 = 1.11 M_\odot$ and the event had an SNR of 103.

The green horizontal and vertical lines show the true tidal deformabilities and radii corresponding to these masses. The EOS agnostic measurements for the event are depicted by solid lines in red and blue for the universal relations proposed in Chatziioannou *et al.* [30] and Yagi and Yunes [34] and Godzieba *et al.* [58], respectively. The corrected distributions are shown with the same colors but in dashed lines. Evidently, the shift in the estimation of the radii with respect to their true values vanishes upon correction.

Bias correction for a population of events.—In the top row of Fig. 7, we show in gray filled circles the median values of the radii (left panel) and tidal deformabilities

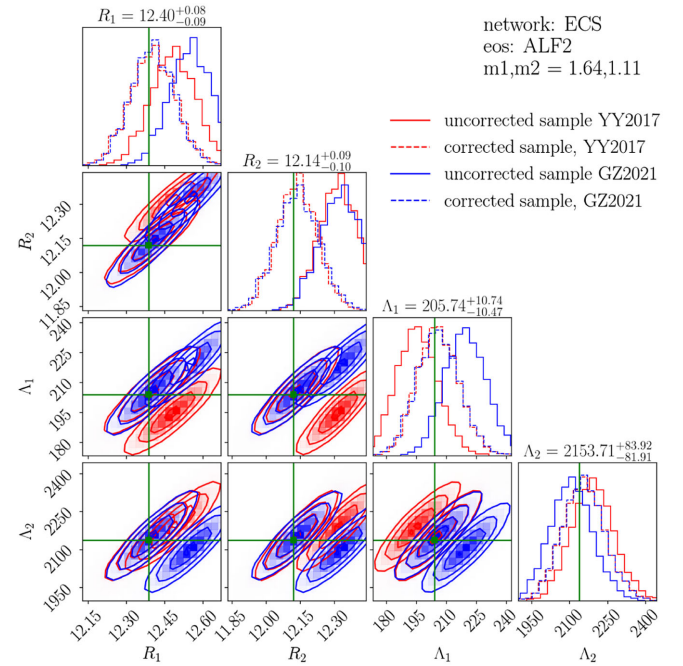


FIG. 6. The comparison between the uncorrected and corrected distributions of radii as well as the tidal deformabilities of the component neutron stars are shown for a fiducial event in our population. The uncorrected and respective corrected distributions are shown with respect to two universal relations shown in the figure—YY2017 [34] and GZ2021 [58]. The event was simulated in the ECS network with the ALF2 EOS. The true values corresponding to the injected masses of $m_1 = 1.64 M_\odot$ and $m_2 = 1.11 M_\odot$ are shown with green horizontal and vertical lines and values on top of each subplot are the median values of the corresponding distributions. Note that both components need not be massive for the bias to be significant.

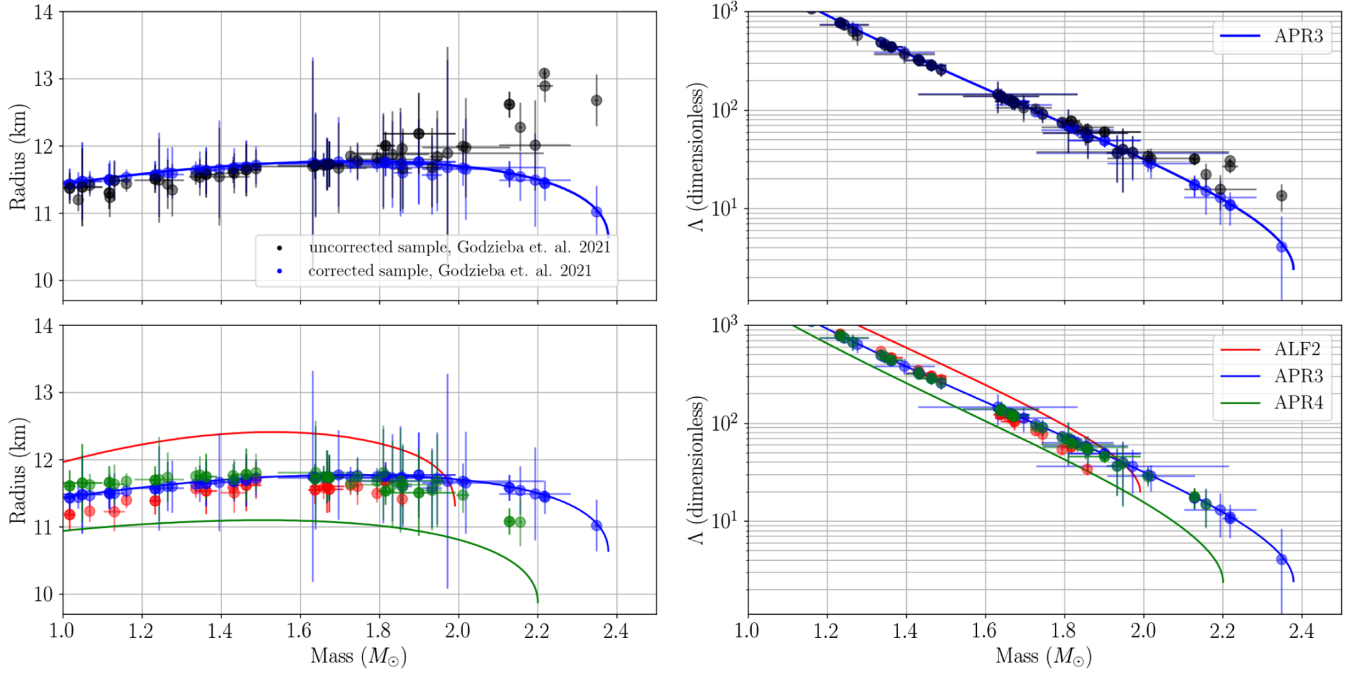


FIG. 7. Inferred radii and Λ before and after bias correction for 30 random events with SNR values larger than 100 in the ECS network, where injection EOS is APR3. Such corrections can also be included separately for two universal relations for tidal deformability and compactness correction. However we find the second universal relation to be quite accurate. Hence, we report the joint correction due to the relevance of radii as a key parameter constraining EOS. We find the recovery of $\tilde{\Lambda}$ to be quite accurate, however individual tidal deformabilities show larger variance. We report the impact of correction on individual Λ .

(right panel) as a function of the component mass, for a randomly selected population of 30 events with SNRs greater than 100 in the ECS network. Tidal deformabilities and radii for this set of events were generated for the APR3 EOS model shown as blue solid lines. Note that the median values of Λ and radii differ significantly from the corresponding EOS curves for the largest and the smallest masses as expected from the systematic bias plot in Fig. 4. Note that the smallest masses are part of a small mass ratio system.

Correcting the radii and tidal deformabilities using residuals appropriate for APR3 yield filled blue circles. Note that the corrected values are now much closer to the true EOS curves. Since the SNRs are pretty large, statistical uncertainties for these events are far smaller than the systematic errors, demonstrating the extent of the biases and effectiveness of the corrections. Note also that systematic biases in the case mass-radius curve, affected by two universal relations, are far greater than the mass- Λ curve, affected by only the first universal relation.

D. Model selection criteria

We now describe our model selection method for a population of events satisfying a given SNR threshold. As before, the population contains 50 events each having an SNR of at least 100 and the events were generated assuming APR3 to be the true EOS. To select a model among a set of EOSs, say, ALF2, APR3, and APR4, we correct for biases

in EOS-agnostic distributions using residuals appropriate to each EOS in the set. The bottom panels of Fig. 7 show how uncorrected median values (gray circles in the top panels) shift to new positions when bias corrections appropriate for the ALF2 (red circles), APR3 (blue circles) and APR4 (green circles) models were applied. Also shown are the mass- Λ (right panel) and mass-radius (left panel) curves corresponding to the three models. Evidently, the population matches better with an EOS model when the bias correction corresponds to the true EOS.

This procedure can be repeated for radii and tidal deformabilities sampled from the posterior distributions of tidal deformabilities and radii, which can then be used in computing the χ^2 relative to each model (see below). We emphasize, however, that in this process we did not consider the errors in the distributions of masses as they are negligibly small but it is straightforward to account for statistical uncertainties in masses.

Model selection with chi square.—Next, to quantify how well the inferred mass-radius curve matches with an EOS model we calculate the χ^2 between the two using

$$\chi^2_{k,M} = \frac{1}{N} \sum_{n=1}^N \frac{(x_n^k - x_n^M)^2}{\sigma_n^2}, \quad (12)$$

where x_n^M is the tidal deformability (or the radius) of the EOS model M corresponding to the n th event and x_n^k is a

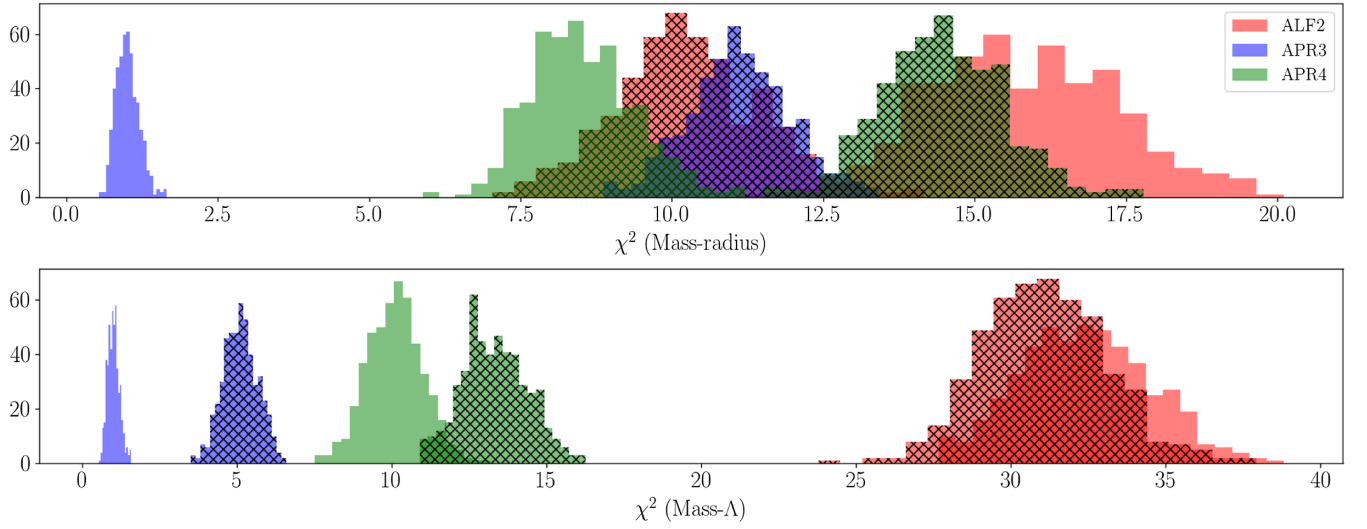


FIG. 8. Distribution of χ^2 for a sample of 30 random events with SNR larger than 100 in the ECS network where the true EOS model is APR3. Bias-corrected χ^2 distributions are shown as unhatched histograms while uncorrected ones are shown as hatched histograms, for radii (top panel) and Λ (bottom panel). Low values of χ^2 at the left end of the panels for bias-corrected χ^2 distributions imply that the reference EOS model is distinguishable from the rest. Similar values of χ^2 for the reference and alternative EOS models for bias-uncorrected χ^2 distributions in the middle of the panels implies that it is hard to distinguish the reference EOS model from an alternative. Note that the distributions are not normalized which accounts for why some look thinner than others.

sample drawn from the bias-corrected posterior distribution (for model M) of the n th event whose standard deviation is σ_n . The set of values $\{m_n, x_n^k\}$, $n = 1, \dots, N$ forms the k th realization of the mass- Λ (or the mass-radius) curve, giving the chi-square value $\chi_{k,M}^2$ for model M . We compute χ^2 for 1000 realizations of the mass- Λ (or the mass-radius) curves to obtain the χ^2 distribution for a given model.

We construct such χ^2 distributions for each model in the set of all EOSs. For an unbiased model selection, the χ^2 distribution constructed from a population of bias-corrected posteriors should have the smallest χ^2 value for the true EOS. In Fig. 8, we plot χ^2 histograms before (hatched) and after (unhatched) bias correction. APR3 was chosen as the true EOS model for the simulated population in the ECS

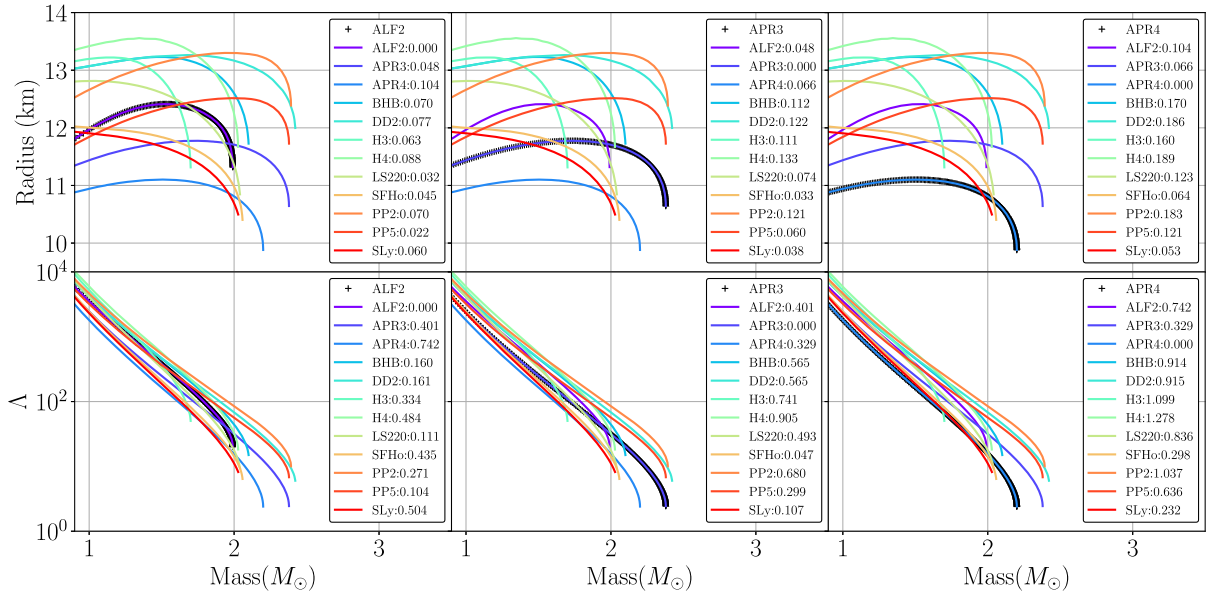


FIG. 9. In this figure we plot mass-radius (top panels) and mass- Λ (bottom panels) for a selection of 12 different EOS models and their L_2 distance from a reference EOS model thick black line, ALF2 (left panels), APR3 (middle panels) and APR4 (right panels). The range of integration in the L_2 norm is taken from $1 M_\odot$ up to the smaller of the maximum mass of the two EOSs.

network and 30 events with SNR larger than 100 were chosen at random from the full population for model selection. It is evident from the plots that bias correction vastly improves model selection, giving lower χ^2 values for the correct EOS model for both the mass-radius (top panel) and mass- Λ comparisons. Furthermore, for this sample population of events model selection without bias correction will either be inconclusive or lead to the selection of an incorrect model as the true model, once again highlighting the efficacy of bias correction.

Kolmogorov-Smirnov (KS) test for model selection.—We use the directional Kolmogorov-Smirnov test statistic [66]

to distinguish two (near-by) χ^2 distributions. Let $\{\chi^2_{k,T}\}$ and $\{\chi^2_{k,A}\}$ be the χ^2 distributions for two EOS models T and A , respectively, and let Φ_T and Φ_A be the corresponding cumulative distribution functions. We have chosen T to be the true model and A to be an alternative. Our null ($\mathcal{H}_0$) and alternative ($\mathcal{H}_1$) hypotheses are defined as

- (i) $\mathcal{H}_0: \Phi_T(\chi^2) \leq \Phi_A(\chi^2)$ for all χ^2 , and
- (ii) $\mathcal{H}_1: \Phi_T(\chi^2) > \Phi_A(\chi^2)$ for at least one χ^2 .

This implies that for an unbiased model selection, where the χ^2 values for the true EOS are expected to be smaller and mostly nonoverlapping with the χ^2 values for the false EOS, the KS statistic will be close to 1. On the contrary, if

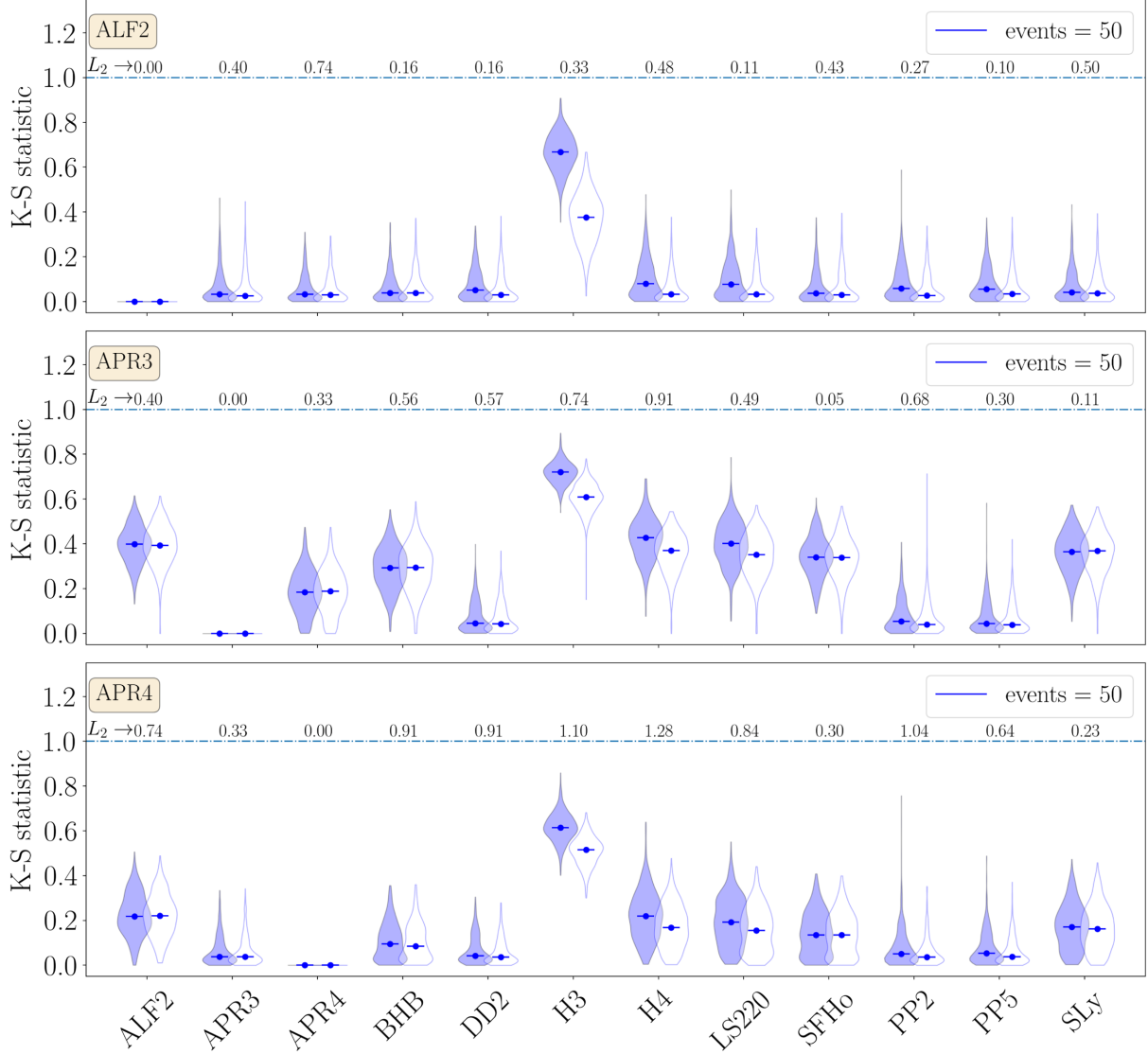


FIG. 10. The plot shows the distribution of the directional KS statistics between the reference model [ALF2 (top), APR3 (middle) and APR4 (bottom)] and an alternative listed along the horizontal axis for mass- Λ curves. The statistic is computed for 50 simulated events with SNR > 10 in the HLVKI+ network. Also listed along the top of each panel are the L_2 distances between the reference model and the alternatives. Blue horizon lines in violin plots are the median values of the KS statistic. Filled (empty) violins are the KS statistic for bias-corrected (uncorrected) posteriors of the tidal deformability. The median posterior is close to zero when the alternative model is the same as the reference model. However, the medians are small for many alternative models. Thus, the HLVKI+ network will not have the sensitivity to determine the right EOS.

two χ^2 distributions are indistinguishable or the model selection is biased meaning that the χ^2 distribution for the incorrect EOS has lower values, the KS statistic will approach 0. In order to not be misled by the fluctuations of a specific realization of a population of events, we bootstrap over 500 distinct realizations of the population. We verified that this is large enough to describe the variation in different realizations of the population.

E. L_2 distance between EOS models

We exemplify the limitations of any model selection method in the following. First, the mass-radius curves of EOSs are not unique for all masses of NSs, but only a subset of them (see Fig. 9). Many of the mass-radius or

mass-tidal deformability EOS pairs intersect each other making those pairs identical at and around the point of intersections (for example ALF2-SLy around $1.4 M_\odot$). Some pairs of EOSs are similar, or even identical, over a large range of masses, which means they are distinguishable only for events in the nonoverlapping region (for example, BHB and DD2 overlap for masses $< 1.5 M_\odot$ in mass-radius plane). Hence, identification of the correct model also depends on the component masses of NSs in the catalog of BNS events, and only events at which a pair of EOSs do not intersect can distinguish them. Luckily, there are many EOS pairs that do not intersect and differ over the entire allowed range of masses. It is easier to discriminate between such pairs.

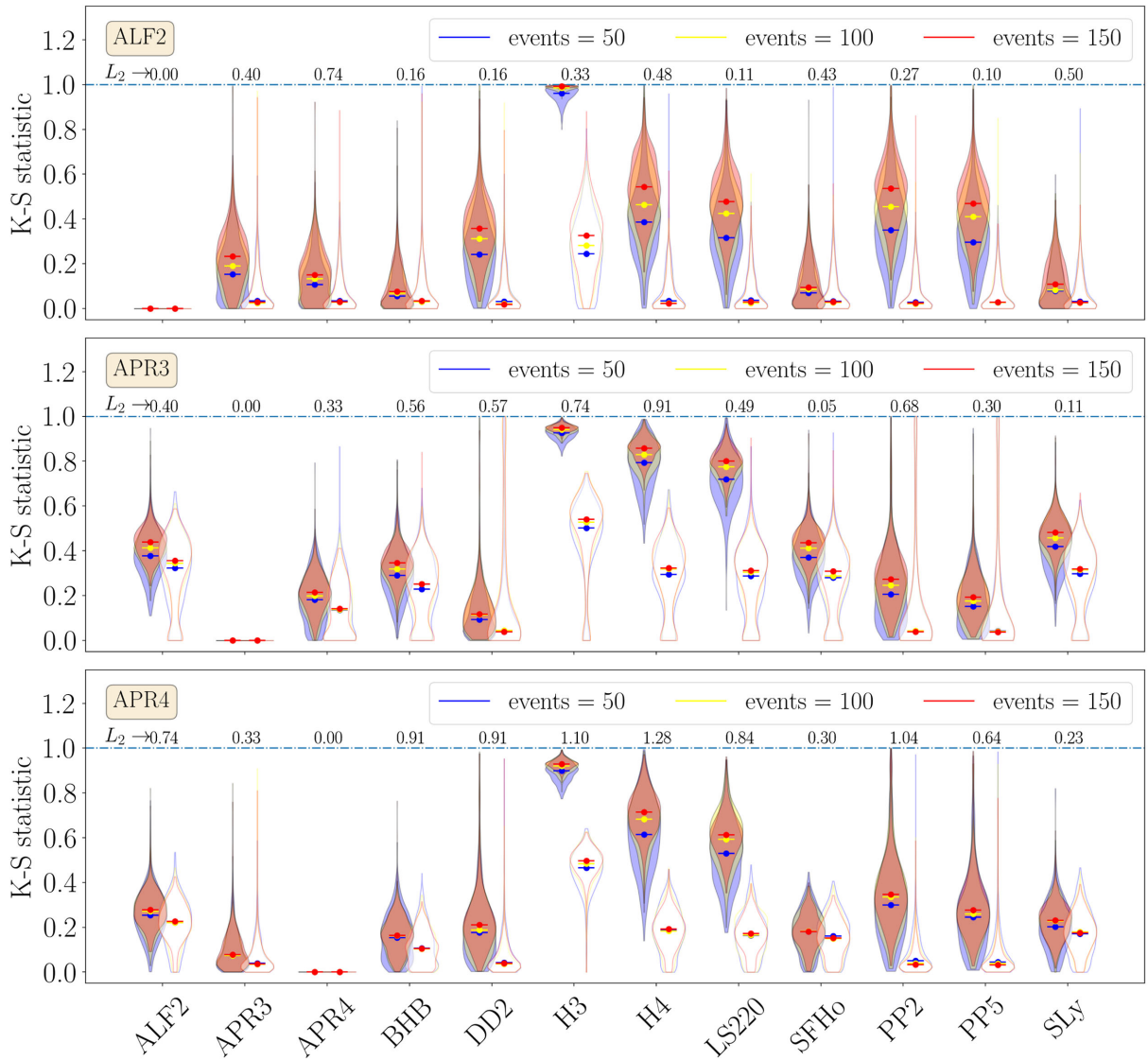


FIG. 11. Same as Fig. 10 but for the ECS network with a random sample of 50 (blue violins), 100 (yellow violins) and 150 (red violins) events with SNR > 10 . It is evident that bias correction increases the KS statistic for incorrect models leaving the statistic unchanged for the reference EOS. As expected, the statistic has greater discriminatory power between different models for a larger number of events.

The L_2 distance between the mass-radius or mass- Λ curves can be used as a measure of the distance between a pair of EOS models. We propose that two models are more easily distinguishable greater is the L_2 distance between the corresponding curves defined as

$$L_{2,R}(A, B) \equiv N_R \int_{m_l}^{m_u} [R_A(m) - R_B(m)]^2 dm \quad (13)$$

$$L_{2,\Lambda}(A, B) \equiv N_\Lambda \int_{m_l}^{m_u} [\Lambda_A(m) - \Lambda_B(m)]^2 dm, \quad (14)$$

where N_R and N_Λ are normalization constants to render the distances dimensionless chosen to be

$$N_R \equiv \left[\int_{m_l}^{m_u} R_A^2 dm \int_{m_l}^{m_u} R_B^2 dm \right]^{-1/2},$$

$$N_\Lambda \equiv \left[\int_{m_l}^{m_u} \Lambda_A^2 dm \int_{m_l}^{m_u} \Lambda_B^2 dm \right]^{-1/2},$$

$R_A(m)$ and $\Lambda_A(m)$ [$R_B(m)$ and $\Lambda_B(m)$] are the mass-radius and mass- Λ curves corresponding to model A (B), m_l is the smallest NS mass in the observed population and m_u is the smaller of the maximum mass allowed by models A and B . Figure 9 plots mass-radius (top panels) and mass- Λ (bottom panels) curves for 12 different EOS models and their L_2 distance from ALF2 (left panels), APR3 (middle panels) and APR4 (right panels).

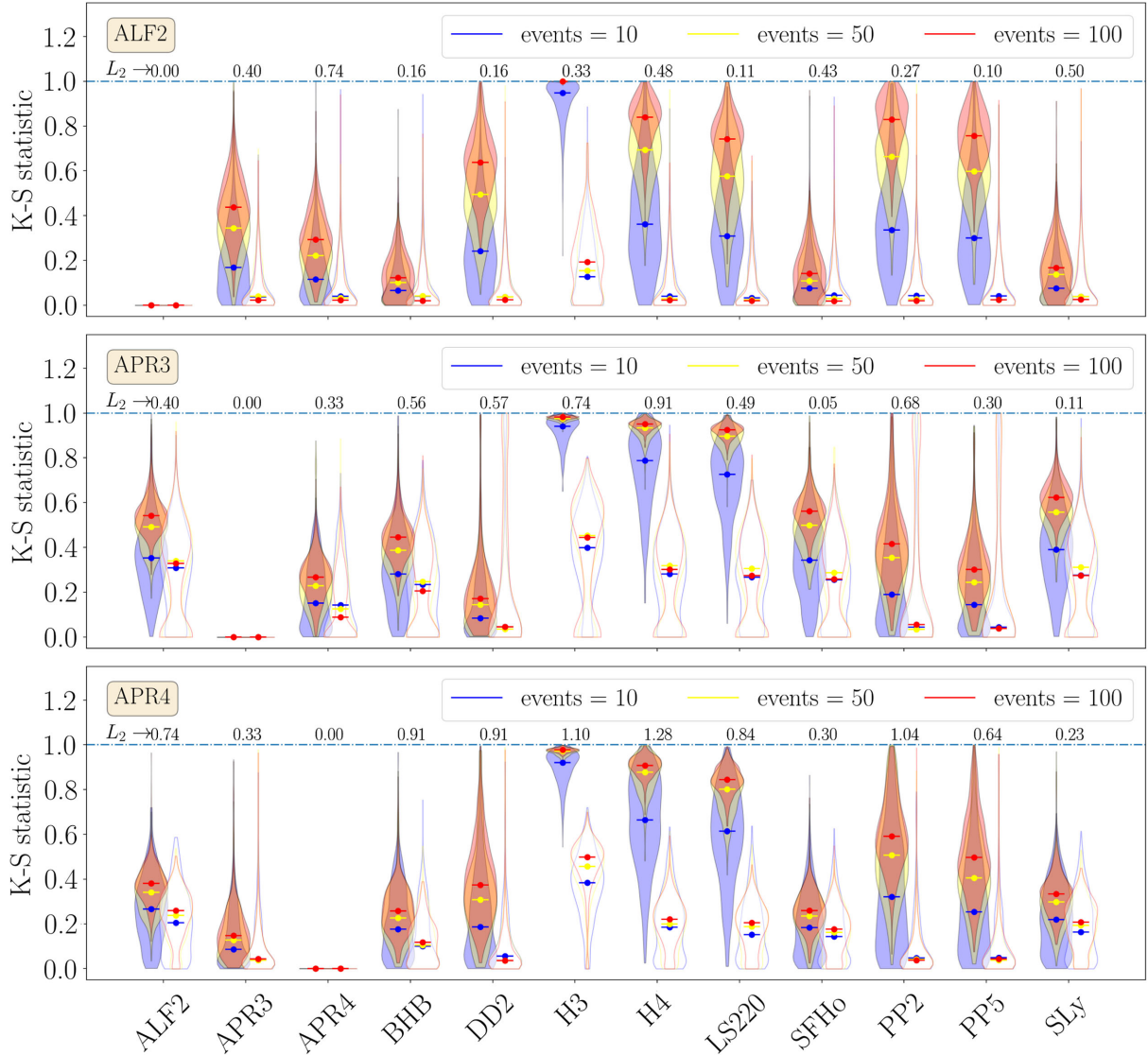


FIG. 12. Same as for Fig. 11 but for events with $\text{SNR} > 30$. The reference model is now distinguishable from many more models than when the SNR threshold was 10. A catalog of 150 events with $\text{SNR} > 30$ will be able to identify the right EOS model with good confidence.

It is evident from the figure that mass- Λ and mass-radius curves are complementary in measuring the distance between different EOS models. For example, DD2 and PP2 are most distant from APR3 in the mass-radius plane (top middle panel) while H4 happens to be most distant in the mass- Λ plane (bottom middle panel). Likewise, APR4 is most distant from ALF2 in both the mass-radius (top left) and mass- Λ (bottom left) planes. This feature highlights the importance of model selection in both mass-radius and mass-tidal deformability planes. On the contrary, the model closest to the reference EOS happens to be the same no matter the parameter space. For example, SFHo has the smallest distance from APR3 in both mass-radius and mass- Λ planes.

IV. MODEL SELECTION WITH A + AND XG NETWORKS

In this section, we will determine the ability of the ECS network to distinguish between different EOS models using the measures introduced in the previous section. We will also consider the HLVKI+ network as a fiducial. Results for mass- Λ curves are presented here; conclusions drawn from mass-radius curves are similar.

As discussed in the previous section, the preferred model is the one for which the χ^2 between the model and the inferred realization of an EOS curve is the smallest. Following the procedure described in the last section (i.e., construction of multiple realizations of the mass- Λ curve for a given population by sampling from the posterior

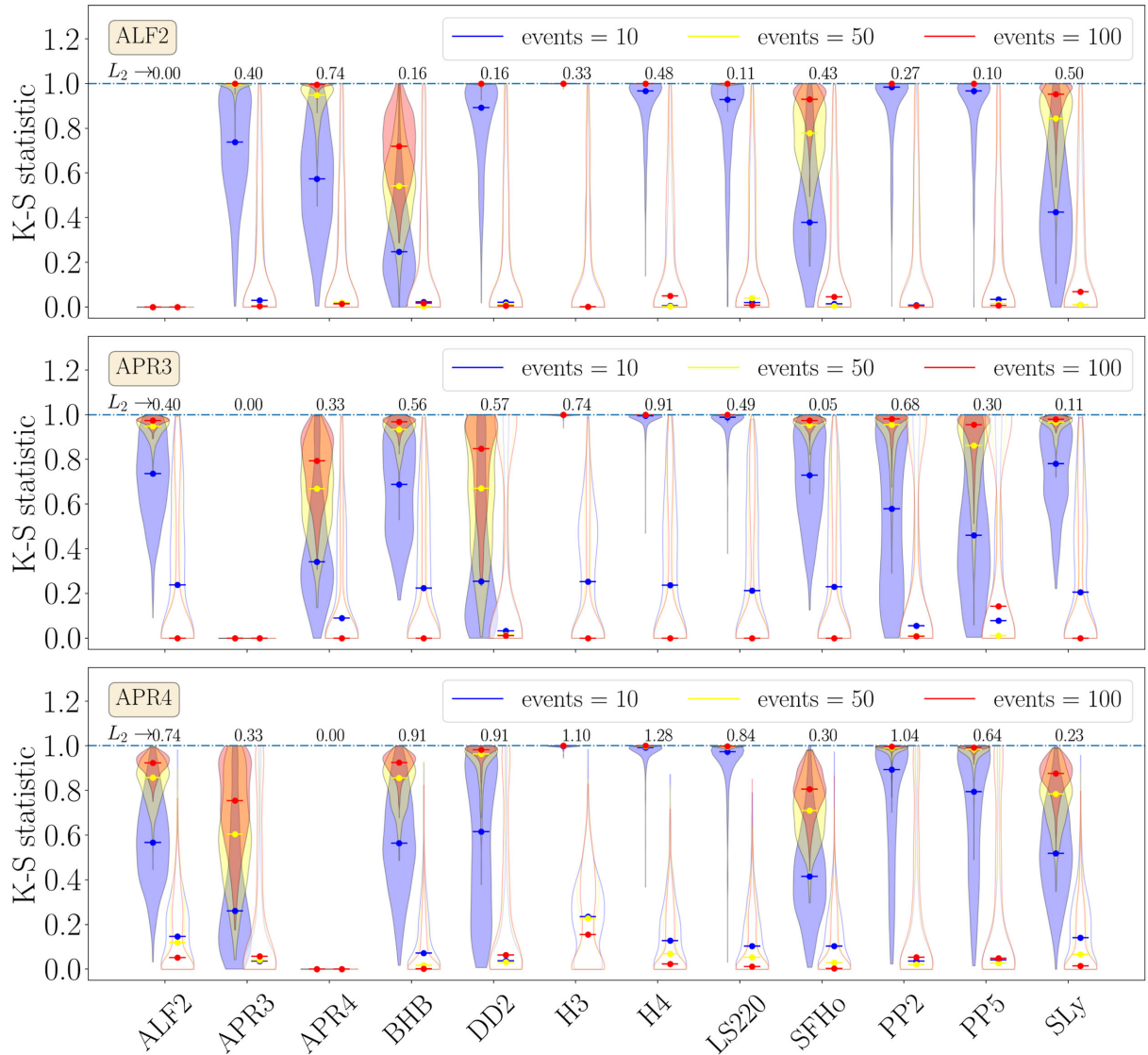


FIG. 13. Same as for Fig. 11 but for events with $\text{SNR} > 100$. Now, just ten events suffice to select the correct model.

distribution of Λ and bootstrapping over several populations of events) gives a distribution of the KS statistics, which is plotted in Figs. 10–13. In order to test the robustness of our method ALF2 (top panels), APR3 (middle panels), and APR4 (bottom panels) were in turn considered to be the true EOS model and model selection was performed over a set of 12 EOSs depicted in Fig. 9.

HLVKI+ network.—Figure 10 plots the distribution of the KS statistic for different realizations of a set of 50 events with $\text{SNR} > 10$, observed in the HLVKI+ network over a one year observing period. The distribution of the KS statistic between the reference model (ALF2 top, APR3 middle and APR4 bottom) and an alternative (shown along the x axis) is shown for both bias-corrected (filled violins) and uncorrected (empty violins) posteriors. Blue markers correspond to the KS statistic for the bias-corrected median realization. Also listed at the top of each panel are the L_2 distance between the reference EOS model and an alternative shown along the x axis.

We see that in general bias correction increases the KS statistic for the incorrect models. This implies that the population realizations that were otherwise indistinguishable or giving rise to an incorrect model selection have better distinguishability. Regardless of the true EOS, H3 is more readily distinguishable from the reference models. Of further note is that APR3 is the easiest to distinguish among the three reference EOS. But otherwise it will be difficult to converge on the true EOS using the HLVKI+ network.

It may be tempting to conclude that the reference model does not have long tails towards KS statistic > 1 but this is not correct since in a real experiment we would not know the true EOS model and hence likely to obtain violins with low medians and long posteriors for many models in the HLVKI+ detector network.

ECS network.—Next, we turn to the next-generation detector network of ECS. Instead of using the full set of the detectable population of events, we choose a subpopulation of events to distinguish between different EOS models. In Fig. 11, we randomly choose 50, 100, and 150 events from the full population of events but with network $\text{SNR} > 10$. We again find that H3 is more readily distinguishable from the reference models. Moreover, distinguishability increases, as expected, with the number of events considered for model selection. We confirm once again that bias correction leads to better distinguishability and correct model selection. This is easily seen for the EOS models H3, H4, and LS220. When the reference EOS is ALF2, correcting the bias in the tidal deformability makes the median realization of a population of 150 events distinguishable. For an injected APR3 or APR4 EOS, it makes almost every realization of the population distinguishable. We note that the events in this population are expected to be predominantly low SNR.

In Figs. 12 and 13 we take the minimum SNR to be 30 and 100, respectively, and also consider three subpopulations of 10, 50, and 100 events. We find that higher SNR events increase the distinguishability across the set of EOS models considered. We report that the median realization of a population of 100 events having a minimum SNR of 100, as is expected from a year of observation, can distinguish the correct EOS model among all the models considered in this work. We further note a divergence in the KS statistic where the statistic for the uncorrected tidal deformability estimates decreases for the SNR threshold of 100 compared to an SNR threshold of 30 for the same number of events. In contrast, the statistic increases for bias-corrected distributions. We conclude that the ECS network, with its ability to frequently detect high SNR events, will have the ability to precisely determine the EOS of dense matter.

V. CONCLUSIONS AND DISCUSSIONS

The determination of the individual tidal deformabilities and radii of a BNS system using gravitational-wave observations suffers a systematic bias due to the use of universal relations. Traditionally, this bias is accounted for by marginalizing over the residuals in the universal relations and, in the process, trading in the systematic errors for statistical ones. This procedure, nevertheless, results in a biased estimate of the relevant quantities and, therefore, to a biased model selection. In this paper, we propose a different method for an unbiased model selection that does not involve residual marginalization and, hence, is, in principle, superior to the existing procedure.

Model selection is performed by calculating the χ^2 , or alternatively, the evidence, for a model given an observation. Since a specific model is assumed in the computation of the evidence, one has knowledge of the residual associated with the model and can rectify the measured quantities by their value at that point. We note that this method does not produce model-agnostic distributions of masses and tidal deformabilities since the corrections are for a specific model.

Given an observed astrophysical population satisfying a minimum SNR condition, we create an EOS curve in the mass- Λ and mass-radius planes by drawing a representative point from the posterior distributions. We repeat this process multiple times to get a band of values in the respective planes. This bootstrapping procedure helps to not bias the inferences to a specific realization of an EOS curve. Next, we consider a set of models to be hierarchically ranked based on their χ^2 values. Here, we calculate the χ^2 for every model in the set, first with the EOS-agnostic band and then with a band obtained after correcting for systematic errors specific to chosen models. This results in two sets of χ^2 distributions which we call the *uncorrected- χ^2* and *corrected- χ^2* distribution sets.

For an unbiased inference, the model corresponding to the distribution with smaller χ^2 values is the preferred model. Unfortunately, this is not always the case since the bias-uncorrected- χ^2 distributions for the wrong model could have lower χ^2 values. Luckily, the bias-corrected- χ^2 distributions are always the lowest for the true model. However, if two EOS curves are similar (i.e., small L_2 distance) then their χ^2 distributions overlap and the models might be indistinguishable. To quantify the distinguishability of two χ^2 distributions and unbiased model selection, we use the directional KS statistic. A statistic for a pair of distributions close to 1 indicates an unbiased model selection and complete distinguishability of the pair. On the other hand, if the statistic is close to zero, the pair of distributions either overlaps or model selection is biased.² When the statistic is close to zero it is not possible to discriminate overlapping distributions from biased ones. Specifically, one can have nonoverlapping χ^2 distributions with the incorrect model having smaller values, in which case the statistic will necessarily be zero.

To ensure that our inferences are not biased by the specific realization of our population, we bootstrap over the observed population. We discussed the violin plots in the previous section, which show the distribution of the KS statistic between a reference model and an alternative

²In this study, we restrict the calculation of the KS statistic to the case where one of the χ^2 distributions is that of the reference model and so the KS statistic indicates how likely the reference model is to be chosen as the preferred one.

model. We observe that the injected model is more likely to be recovered for bias-corrected- χ^2 distributions and for a greater number of observations. The two main advantages of our method over the residual marginalization method are as follows. First, we do not need to sample over the parameters that model the residuals and hence are computationally favored. Second, the statistical errors are mostly unaffected in our method and, therefore, our model selection has greater sensitivity. Though not explored in this work, it would be interesting to do a direct comparison of the model selection prowess of our method versus the method of residual marginalization. More specifically, this would reveal the effect of the latter method on the statistical errors and its subsequent effect on the unbiased distinguishability of nearby EOS curves. We leave this to future work.

ACKNOWLEDGMENTS

We thank Philippe Landry for the internal review and his several suggestions to improve the manuscript. We thank Nathan Johnson-McDaniel, David Radice and Sukanta Bose for their comments and critical questions. We thank Monica Bapna and P. Ajith for useful discussions during a previous work showing the derivation of a more general methodology of model selection in an unpublished work. We thank Daniel Godzieba for sharing two EOSs from his piecewise polytropic collection. A.D. and B.S.S. were supported in part by NSF Grants No. PHY-2012083, No. PHY-2207638 and No. AST-2006384. This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation.

-
- [1] B. P. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), *Phys. Rev. Lett.* **119**, 161101 (2017).
 - [2] B. P. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), *Phys. Rev. X* **9**, 011001 (2019).
 - [3] B. P. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), *Phys. Rev. X* **9**, 031040 (2019).
 - [4] R. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), *Phys. Rev. X* **11**, 021053 (2021).
 - [5] R. Abbott *et al.* (LIGO Scientific, VIRGO, KAGRA Collaborations), [arXiv:2111.03606](https://arxiv.org/abs/2111.03606).
 - [6] J. Aasi *et al.* (LIGO Scientific Collaboration), *Classical Quantum Gravity* **32**, 115012 (2015).
 - [7] F. Acernese *et al.* (VIRGO Collaboration), *Classical Quantum Gravity* **32**, 024001 (2015).
 - [8] R. Abbott *et al.* (LIGO Scientific, VIRGO, KAGRA Collaborations), [arXiv:2111.03634](https://arxiv.org/abs/2111.03634).
 - [9] E. E. Flanagan and T. Hinderer, *Phys. Rev. D* **77**, 021502 (2008).
 - [10] T. Hinderer, *Astrophys. J.* **677**, 1216 (2008).
 - [11] J. Vines, E. E. Flanagan, and T. Hinderer, *Phys. Rev. D* **83**, 084051 (2011).
 - [12] G. Baym, C. Pethick, and P. Sutherland, *Astrophys. J.* **170**, 299 (1971).
 - [13] J. S. Read, B. D. Lackey, B. J. Owen, and J. L. Friedman, *Phys. Rev. D* **79**, 124032 (2009).
 - [14] J. M. Lattimer and M. Prakash, *Phys. Rep.* **621**, 127 (2016).
 - [15] L. Lindblom, *Phys. Rev. D* **82**, 103011 (2010).
 - [16] L. Lindblom, *Phys. Rev. D* **105**, 063031 (2022).
 - [17] M. F. O'Boyle, C. Markakis, N. Stergioulas, and J. S. Read, *Phys. Rev. D* **102**, 083027 (2020).
 - [18] J. R. Oppenheimer and G. M. Volkoff, *Phys. Rev.* **55**, 374 (1939).
 - [19] R. C. Tolman, *Phys. Rev.* **55**, 364 (1939).
 - [20] M. Favata, *Phys. Rev. Lett.* **112**, 101101 (2014).
 - [21] L. Wade, J. D. E. Creighton, E. Ochsner, B. D. Lackey, B. F. Farr, T. B. Littenberg, and V. Raymond, *Phys. Rev. D* **89**, 103012 (2014).
 - [22] I. Harry and A. Lundgren, *Phys. Rev. D* **104**, 043008 (2021).

- [23] T. Dietrich, T. Hinderer, and A. Samajdar, *Gen. Relativ. Gravit.* **53**, 27 (2021).
- [24] A. Buikema *et al.* (LIGO Collaboration), *Phys. Rev. D* **102**, 062003 (2020).
- [25] R. Smith *et al.*, *Phys. Rev. Lett.* **127**, 081102 (2021).
- [26] K. Yagi and N. Yunes, *Classical Quantum Gravity* **33**, 13LT01 (2016).
- [27] T. Damour and A. Nagar, *Phys. Rev. D* **80**, 084035 (2009).
- [28] J. Saes and R. F. P. Mendes, *Phys. Rev. D* **106**, 043027 (2022).
- [29] B. P. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), *Phys. Rev. Lett.* **121**, 161101 (2018).
- [30] K. Chatziioannou, C.-J. Haster, and A. Zimmerman, *Phys. Rev. D* **97**, 104036 (2018).
- [31] B. Kumar and P. Landry, *Phys. Rev. D* **99**, 123026 (2019).
- [32] Z. Carson, K. Chatziioannou, C.-J. Haster, K. Yagi, and N. Yunes, *Phys. Rev. D* **99**, 083016 (2019).
- [33] B. Biswas, *Astrophys. J.* **926**, 75 (2022).
- [34] K. Yagi and N. Yunes, *Phys. Rep.* **681**, 1 (2017).
- [35] B. P. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), *Classical Quantum Gravity* **37**, 045006 (2020).
- [36] B. D. Lackey and L. Wade, *Phys. Rev. D* **91**, 043002 (2015).
- [37] C. Pacilio, A. Maselli, M. Fasano, and P. Pani, *Phys. Rev. Lett.* **128**, 101101 (2022).
- [38] S. Ghosh, X. Liu, J. Creighton, W. Kastaun, G. Pratten, and I. M. Hernandez, *Phys. Rev. D* **104**, 083003 (2021).
- [39] J. Aasi *et al.* (LIGO Scientific Collaboration), *Classical Quantum Gravity* **32**, 115012 (2015).
- [40] F. Acernese *et al.* (VIRGO Collaboration), *Classical Quantum Gravity* **32**, 024001 (2015).
- [41] T. Akutsu *et al.* (KAGRA Collaboration), *Nat. Astron.* **3**, 35 (2019).
- [42] D. Reitze *et al.*, *Bull. Am. Astron. Soc.* **51**, 035 (2019), <https://baas.aas.org/pub/2020n7i035/release/1>.
- [43] M. Punturo *et al.*, *Classical Quantum Gravity* **27**, 084007 (2010).
- [44] S. Borhanian and B. S. Sathyaprakash, *arXiv:2202.11048*.
- [45] J. A. de Freitas Pacheco, *Astropart. Phys.* **8**, 21 (1997).
- [46] E. Vangioni, K. A. Olive, T. Prestegard, J. Silk, P. Petitjean, and V. Mandic, *Mon. Not. R. Astron. Soc.* **447**, 2575 (2015).
- [47] P. Beniamini and T. Piran, *Mon. Not. R. Astron. Soc.* **487**, 4847 (2019).
- [48] N. Aghanim *et al.*, *Astron. Astrophys. Suppl. Ser.* **641**, A6 (2020).
- [49] R. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), *Astrophys. J. Lett.* **913**, L7 (2021).
- [50] M. Alford, M. Braby, M. W. Paris, and S. Reddy, *Astrophys. J.* **629**, 969 (2005).
- [51] A. Akmal, V. R. Pandharipande, and D. G. Ravenhall, *Phys. Rev. C* **58**, 1804 (1998).
- [52] S. Banik, M. Hempel, and D. Bandyopadhyay, *Astrophys. J. Suppl. Ser.* **214**, 22 (2014).
- [53] S. Typel, G. Ropke, T. Klahn, D. Blaschke, and H. H. Wolter, *Phys. Rev. C* **81**, 015803 (2010).
- [54] B. D. Lackey, M. Nayyar, and B. J. Owen, *Phys. Rev. D* **73**, 024021 (2006).
- [55] J. M. Lattimer and F. D. Swesty, *Nucl. Phys. A* **535**, 331 (1991).
- [56] A. W. Steiner, M. Hempel, and T. Fischer, *Astrophys. J.* **774**, 17 (2013).
- [57] F. Douchin and P. Haensel, *Astron. Astrophys.* **380**, 151 (2001).
- [58] D. A. Godzieba, R. Gamba, D. Radice, and S. Bernuzzi, *Phys. Rev. D* **103**, 063036 (2021).
- [59] N. Aghanim *et al.* (Planck Collaboration), *Astron. Astrophys.* **641**, A6 (2020); **652**, C4(E) (2021).
- [60] S. Chandrasekhar, *The Mathematical Theory of Black Holes* (Oxford University Press, 1985).
- [61] S. Borhanian, *Classical Quantum Gravity* **38**, 175014 (2021).
- [62] I. Harry and T. Hinderer, *Classical Quantum Gravity* **35**, 145010 (2018).
- [63] J. M. Lattimer and M. Prakash, *Astrophys. J.* **550**, 426 (2001).
- [64] K. Yagi, *Phys. Rev. D* **89**, 043011 (2014); **96**, 129904(E) (2017); **97**, 129901(E) (2018).
- [65] S. Bernuzzi and A. Nagar, *Phys. Rev. D* **78**, 024024 (2008).
- [66] E. D. Feigelson and G. J. Babu, *arXiv:1205.2064*.