Probing Dynamical Gravity with the Polarization of Continuous Gravitational Waves

Maximiliano Isi,[1, *] Matthew Pitkin,[2, †] and Alan J. Weinstein[1, ‡]

[1]*LIGO Laboratory,*
California Institute of Technology, Pasadena, California 91125, USA
[2]*University of Glasgow, Glasgow G12 8QQ, Scotland, UK*
(Dated: March 21, 2017)

The direct detection of gravitational waves provides the opportunity to measure fundamental aspects of gravity which have never been directly probed before, including the polarization of gravitational waves. In the context of searches for continuous waves from known pulsars, we present novel methods to detect signals of any polarization content, measure the modes present and place upper-limits on the amplitude of non-tensorial components. This will allow us to obtain new model-independent, dynamical constraints on deviations from general relativity. We test this framework on multiple potential sources using simulated data from three advanced-era detectors at design sensitivity. We find that signals of any polarization will become detectable and distinguishable for characteristic strains $h \gtrsim 3 \times 10^{-27} \sqrt{1 \text{ yr}/T}$, for an observation time T. We also find that our ability to detect non-tensorial components depends only on the power present in those modes, irrespective of the strength of the tensorial strain.

PACS numbers: 04.80.Cc, 04.30.Nk, 04.50.Kd, 04.80.Nn I.

I. INTRODUCTION

The recent detection of gravitational waves (GWs) by the advanced Laser Interferometer Gravitational-wave Observatory (aLIGO) heralds the beginning of the long-awaited era of GW astronomy [1, 2]. One of the main goals of this field is to use GWs as a probe of fundamental physics in the highly-dynamical and strong-field regimes of gravity, as predicted by the general theory of relativity (GR), which are inaccessible to laboratory, Solar System or cosmological tests of gravity. The first few GW detections have already been used to place some of the most stringent constraints on deviations from GR in this domain. However, it has not been possible to use these signals to learn about the polarization content of GWs [3], a measurement highly relevant when comparing GR to many of its alternatives [4]. The reason for this is that the relative orientation of the two LIGO detectors makes it nearly impossible to unequivocally disentangle the polarizations of transient GW signals like the compact-binary coalescences (CBCs) observed so far. Since other observations usually taken to constrain the amount of allowed non-GR polarizations (e.g. [5]) are indirect and strongly model-dependent (because they are sensitive only to total radiated power) this means that there currently exist no direct measurements of GW polarizations.

Unlike CBC transients, continuous gravitational waves (CWs) are, by definition, long-lasting narrow-band signals. Although they have not yet been observed [6], CWs are expected to be sourced by stable systems, like spinning neutron stars (NS) with an asymmetric moment of inertia [7]. If detected, such signals would allow for tests of gravity complementary to those achievable with transients, including the study of GW polarizations [8].

In [8] we showed that it is possible to search for CWs in a polarization-agnostic way and to disentangle the polarization content if a signal is present. However, the data analysis methods proposed were based on a frequentist approach to statistics and suffered from the associated limitations. In this paper, we reframe the ideas of [8] in a more sophisticated Bayesian framework that allows us to achieve the following novel goals:

1. *Model-independent detection:* determine whether a set of GW detector data, prepared for any given known pulsar and from one or multiple detectors, provides evidence for the presence of an astrophysical signal of any polarization content.

2. *Model selection:* in the presence of a signal, determine whether the data favor GR or a generic non-GR model, as well as comparing specific alternative theories among themselves and to GR; combine data for multiple sources into a single statement about the validity of GR.

3. *Inference:* if the data favor the presence of a GR signal, place constraints on specific alternative theories using the tools of Bayesian parameter estimation.

Furthermore, while [8] treated only the case of a single detector, we are now able to consider the generic case of a network of detectors.

We present Bayesian methods to achieve the three goals above in the context of searches targeted to known pulsars and present sensitivity estimates for the advanced detector era, including the first generic estimates of sensitivity to non-tensorial CW polarizations ever published.

---
* misi@ligo.caltech.edu
† matthew.pitkin@glasgow.ac.uk
‡ ajw@ligo.caltech.edu
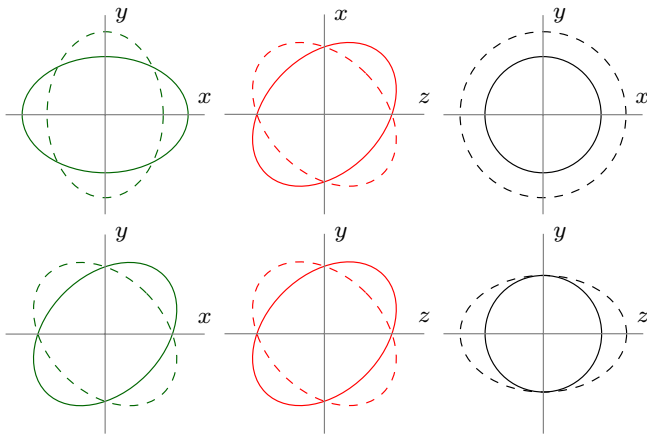
FIG. 1. *Effect of different GW polarizations on a ring of free-falling test particles.* Plus (+) and cross (×) tensor modes (green); vector-x (x) and vector-y (y) modes (red); breathing (b) and longitudinal (l) scalar modes (black). In all of these diagrams the wave propagates in the z–direction. This decomposition into polarizations was first proposed for generic metric theories in [4].

In Section II we review the basics of beyond-Einstein polarizations and the targeted pulsar CW search. In Section III, we phrase our problem in the language of model selection and explain the construction of hypotheses that will allow us to distinguish GR from non-GR signals. In section IV we specify the details of our analysis, and we explain our results in Section V. Finally, we summarize our findings and explain caveats in section VI.

II. BACKGROUND

A. Polarizations

GWs can be decomposed into different polarizations, which arise from the linearly independent components of the three-dimensional, rank-two tensor representing the spatial metric perturbation [4]. A generic metric theory of gravity may thus allow any combination of up to six independent modes: plus (+), cross (×), vector x (x), vector y (y), breathing (b) and longitudinal (l). The effect of each of these modes is represented in Fig. 1. The rank of the fields underlying any given theory determines which polarizations the theory supports: + and × correspond to tensor fields, x and y to vector fields, and b and l to scalar fields. The components of the tensor and vector pairs are not separable, in the sense that a signal model that includes one element of the group must also include the other (e.g. it is not possible to have a model that allows plus + but not ×), because the distinction between modes of the same spin is contingent on the frame of reference (relative orientation of source and detector).

As can be deduced from GR's tensorial nature, Einstein's theory only allows the existence of the + and

× polarizations. On the other hand, scalar-tensor and massive-graviton theories may also predict the presence of some b and/or l component associated to the theory's extra scalar field [9]. On top of tensor and scalar modes, bimetric theories, like Rosen or Lightman-Lee theories, also predict vector modes [10]. Furthermore, less conventional theories might, in principle, predict the existence of vector or scalar modes *only*, while still possibly being in agreement with all other non-GW tests of GR (see [11] for an example). Although all these different theoretical frameworks serve as motivation for our study, our approach to the measurement of GW polarizations is phenomenological and, thus, theory-agnostic (Sec. III). It is important to underscore that the detection of a GW signal with a non-GR polarization, no matter how small, is sufficient to falsify GR (note the converse is not true, however).

Because different polarizations have geometrically distinct effects, GW detectors will react differently to each mode. This is manifested in the detector response function F_p for each polarization p, which encodes the effect of a linearly p-polarized GW with unit amplitude, $h_p = 1$. Ground-based GW detectors, like LIGO and Virgo are quadrupolar antennas that perform low-noise measurements of the strain associated with the differential motion of two orthogonal arms. Their detector response function can thus be written as [8, 12, 13]:

$$F_+ = \frac{1}{2}\left[(\mathbf{w}_x \cdot \mathbf{d}_x)^2 - (\mathbf{w}_x \cdot \mathbf{d}_y)^2 - (\mathbf{w}_y \cdot \mathbf{d}_x)^2 + (\mathbf{w}_y \cdot \mathbf{d}_y)^2\right], \tag{1}$$

$$F_\times = (\mathbf{w}_x \cdot \mathbf{d}_x)(\mathbf{w}_y \cdot \mathbf{d}_x) - (\mathbf{w}_x \cdot \mathbf{d}_y)(\mathbf{w}_y \cdot \mathbf{d}_y), \tag{2}$$

$$F_{\mathrm{x}} = (\mathbf{w}_x \cdot \mathbf{d}_x)(\mathbf{w}_z \cdot \mathbf{d}_x) - (\mathbf{w}_x \cdot \mathbf{d}_y)(\mathbf{w}_z \cdot \mathbf{d}_y), \tag{3}$$

$$F_{\mathrm{y}} = (\mathbf{w}_y \cdot \mathbf{d}_x)(\mathbf{w}_z \cdot \mathbf{d}_x) - (\mathbf{w}_y \cdot \mathbf{d}_y)(\mathbf{w}_z \cdot \mathbf{d}_y), \tag{4}$$

$$F_{\mathrm{b}} = \frac{1}{2}\left[(\mathbf{w}_x \cdot \mathbf{d}_x)^2 - (\mathbf{w}_x \cdot \mathbf{d}_y)^2 + (\mathbf{w}_y \cdot \mathbf{d}_x)^2 - (\mathbf{w}_y \cdot \mathbf{d}_y)^2\right], \tag{5}$$

$$F_{\mathrm{l}} = \frac{1}{\sqrt{2}}\left[(\mathbf{w}_z \cdot \mathbf{d}_x)^2 - (\mathbf{w}_z \cdot \mathbf{d}_y)^2\right]. \tag{6}$$

Here, the spatial vectors $\mathbf{d}_x$, $\mathbf{d}_y$ have unit norm and point along the detector arms such that $\mathbf{d}_z = \mathbf{d}_x \times \mathbf{d}_y$ is the local zenith; the direction of propagation of the wave from a source at known sky location (specified by right ascension α, and declination δ) is given by $\mathbf{w}_z$, and $\mathbf{w}_x$, $\mathbf{w}_y$ are such that $\mathbf{w}_z = \mathbf{w}_x \times \mathbf{w}_y$. The angle between $\mathbf{w}_y$ and the celestial north is ψ, the polarization position angle, and is determined by the orientation of the source with respect to the observer. Note that, because of their symmetries, the breathing and longitudinal modes are degenerate for quadrupolar antennas, so it is enough for us to consider just one of them.

For a given detector, polarization angle and sky location, the antenna patterns (APs) of Eqs. (1–6) become simple, distinct functions of time determined by the rotation of the Earth. As we explain in Sec. III A 1, the APs of polarizations corresponding to different spins can be differentiated even in the absence of information on the source orientation; for the minority of cases in which such information exists, it can be taken into account to better distinguish among specific signal models (see Sec. IV).

Because their characteristic period (a sidereal day) is much longer than the CBC timescale (order of minutes or less), the APs are treated as constants in transient searches; however, this simplification is not valid for CW searches, since their coherent-integration time can be of the order of months to years. As we have pointed out before, this can be used to distinguish the polarization content of a signal [8]. Assuming wave frequency and speed are the same for all modes, the only differences between CWs of different polarizations arise from the sidereal-period amplitude modulations caused by each AP.

B. Continuous waves

1. Signal

A CW with constant intrinsic amplitude, phase evolution $\phi(t)$ and arbitrary polarization content will induce a strain in a quadrupolar detector which, in the time-domain, can be written as:

$$h(t) = \sum_p F_p(t) h_p(t), \tag{7}$$

where the sum is over the five independent polarizations, $p \in \{+, \times, \text{x}, \text{y}, \text{b}\}$; the F_p's are those of Eqs. (1–5), and thus implicitly depend on the relative location and orientation of source and detector by means of ψ, α & δ; the h_p term encodes the amplitude and phase of the wave before being projected onto the frame of the detector:

$$h_p(t) = a_p \cos\left(\phi(t) + \phi_p\right), \tag{8}$$

where is a_p a time-independent amplitude with a functional dependence on source parameters determined by each particular theory of gravity; $\phi(t)$ the phase evolution, a consequence of the dynamics of the source in that theory; and ϕ_p a phase offset for each polarization.

In GR, there are several mechanisms by which a NS could emit CWs, but the most likely is the presence of a non-axisymmetry in the star's moment of inertia [14]. For this type of triaxial, non-precessing source, GR predicts:

$$h_+(t) = h_0 \frac{1}{2}(1 + \cos^2 \iota) \cos \phi(t), \tag{9}$$

$$h_\times(t) = h_0 \cos \iota \sin \phi(t), \tag{10}$$

$$h_\text{x} = h_\text{y} = h_\text{b} = h_\text{l} = 0, \tag{11}$$

where ι is the inclination angle between the spin axis of the source and the observing line-of-sight, and h_0 is an overall amplitude given by:

$$h_0 = \frac{16\pi^2 G}{c^4} \frac{\epsilon I_{zz} f_{\text{rot}}^2}{r}, \tag{12}$$

where r is the source distance, f_{rot} its rotation frequency around the principal axis z, $\mathbf{I}$ the moment-of-inertia tensor and $\epsilon \equiv (I_{xx} - I_{yy})/I_{zz}$ the equatorial ellipticity. For the triaxial case, the GW frequency f is twice the rotational value f_{rot}, so that we can write:

$$\phi(t) = 2\phi_{\text{rot}}(t) + \phi_{(\text{GW}-\text{EM})}, \tag{13}$$

where ϕ_{rot} is the rotational phase as measured via electromagnetic (EM) observations and $\phi_{(\text{GW}-\text{EM})}$ is a potential, constant phase offset between the GW and EM signals that can be absorbed into the definition of the ϕ_p's in Eq. (8).

Note that other emission mechanisms may result in GW radiation at $f = f_{\text{rot}}$ [7], or even noninteger powers f_{rot} [15, 16]. Furthermore, alternative theories of gravity may (and, in general, will) support signals at any harmonic. Although in this paper we only consider the case in which only the second rotational harmonic appears in the GW phase, the analysis can be easily generalized to also include contributions from the fundamental and other multiples of f_{rot} (see Sec. VI).

2. Targeted search

We would like to search a given set of data (from one or more detectors) for CW signals coming from a specific candidate pulsar which has already been observed and timed electromagnetically. Timing solutions are obtained through the pulsar timing package TEMPO2 [17]. We want to achieve this regardless of polarization content, and to reliably distinguish between the different modes present.

If we assume all polarizations share the same phase evolution, then detector response is the only factor distinguishing CW polarizations and, thus, all the relevant information is encoded in the sidereal-day-period amplitude modulation of the signal. This allows us to focus on a narrow frequency band around the expected GW frequency by processing the data following the complex-heterodyne method developed in [18] and [19]. This procedure is summarized below.

A signal like Eqs. (7) and (8) can be rewritten in the form:

$$h(t) = \Lambda(t) e^{i\phi(t)} + \Lambda^*(t) e^{-i\phi(t)}, \tag{14}$$

$$\Lambda(t) \equiv \frac{1}{2} \sum_{p=1}^{5} a_p e^{i\phi_p} F_p(t; \psi), \tag{15}$$

with $*$ indicating complex conjugation and $\phi(t)$ given by a Taylor expansion around $f = 2f_{\mathrm{rot}}$:

$$\phi(t) = 2\pi \left(2f_{\mathrm{rot}} t_{\mathrm{SSB}} + \dot{f}_{\mathrm{rot}} t_{\mathrm{SSB}}^2 + ... \right), \qquad (16)$$

where t_{SSB} is itself a function of time given by:

$$t_{\mathrm{SSB}} = t + \Delta_{\mathrm{R}} + \Delta_{\mathrm{E}} + \Delta_{\mathrm{S}} + \Delta_{\mathrm{binary}} . \qquad (17)$$

In the above, t_{SSB} is the time measured by a clock at rest with respect to the Solar System barycenter (SSB), a point which is itself assumed to be inertial with respect to the pulsar; t is the time as measured at a given detector; Δ_{R} is the Roemer delay; Δ_{E} is the Solar-System Einstein delay; Δ_{S} is the Solar-System Shapiro delay; Δ_{binary} is the Shapiro delay originating from the motion of the pulsar in its binary (a term that vanishes for isolated sources) [19].

Because the phase evolution $\phi(t)$, including all corrections from Eq. (17), is known (with known uncertainties) from electromagnetic observations, we can digitally heterodyne the data by multiplying by $\exp\left[-i\phi(t)\right]$ so that the signal therein becomes:

$$h'(t) \equiv h(t)e^{-i\phi(t)} = \Lambda(t) + \Lambda^*(t)e^{-i2\phi(t)} \qquad (18)$$

and the frequency modulation of the first term is removed, while that of the second term is doubled. A series of low-pass filters can then be used to remove the quickly-varying term, which enables the down-sampling of the data by averaging over minute-long time bins. As a result, $\Lambda(t)$ is the only contribution from the original signal left in our data, and hence we can use Eq. (15) as the template for our search. Note that, although we started with real-valued data, after this process the data are now complex.

From Eq. (18) we see that, in the presence of a signal, the heterodyned and down-sampled noisy detector strain data B_k for the kth time bin (which can also be labeled by the Earth-frame GPS time-of-arrival at the detector, t_k) are expected to be of the form:

$$B_{\mathrm{expected}}(t_k) = \frac{1}{2}\sum_{p=1}^{5} a_p e^{i\phi_p} F_p(t_k; \psi, \alpha, \delta) + n(t_k), \quad (19)$$

where $n(t_k)$ is the heterodyned, filtered and downsampled noise in bin k, which carries no information about the GW signal. It is important to remember that, the F_p's are functions of the source orientation and sky location relative to the detector, so we have made this dependence explicit in Eqs. (15) and (19) by writing $F_p(t_k)$ as $F_p(t_k; \psi, \alpha, \delta)$. Also, recall that these functions have a characteristic period of a sidereal day ($\sim 10^{-5}$ Hz).

III. METHOD

A. Model selection

We use the tools of Bayesian model selection (also known as *second-level inference*) to determine whether the data contain a signal and, if so, whether that signal agrees with the GR prediction or not. Our procedure is hierarchical and consists of the following stages:

1. *detection*: select between the signal and noise models;

2. *test of GR*: if a signal is present, select between GR and non-GR models;

3. *upper limits*: if GR is favored, place upper limits on non-tensorial strain amplitudes, in the context of specific alternative polarization models.

This subsection covers only the first two items in this list, since the placement of upper limits belongs in the section on parameter estimation. We treat the case of a single data set in III A 1 and III A 2, while and we show how to combine results from multiple analyses in III A 3; we offer some considerations about how to approach the problem of non-Gaussian noise in III A 4.

1. Hypotheses

For any given pulsar, we would like to use *reduced* (i.e. heterodyned, filtered and downsampled) GW data to decide between the following two logically disjoint hypotheses:

1. *noise* ($\mathcal{H}_{\mathrm{N}}$): no signal, the data are drawn from a Gaussian distribution of zero mean and some standard deviation;

2. *signal* ($\mathcal{H}_{\mathrm{S}}$): the data contain noise drawn from a Gaussian distribution and a signal with the assumed phase evolution and *any* polarization content.

In order to perform model selection, we need to translate these hypotheses into the corresponding Bayesian model, which means setting a likelihood function derived from the expected noise properties and picking a multidimensional prior distribution over all parameters. It is important to underscore that a Bayesian model is *defined* by the choice of these two probability distributions.

For $\mathcal{H}_{\mathrm{N}}$, the construction of the likelihood is straightforward. First, let σ be the standard deviation of the detector noise at or near the expected GW frequency; then, for each complex-valued data point B_k, Gaussianity implies:

$$p(B_k \mid \sigma, \mathcal{H}_{\mathrm{N}}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{|B_k|^2}{2\sigma^2}\right). \qquad (20)$$

If the data are split into N_{S} segments of lengths s_j ($j = 1, \ldots, N_{\mathrm{S}}$) over which the standard deviation σ_j is assumed to remain constant, we can analytically marginalize over this parameter to obtain a likelihood for the entire data set $\mathbf{B}$ in the form of a Student's t-distribution

[19, 20]:

$$P(\mathbf{B} \mid \mathcal{H}_\mathrm{N}) = \prod_{j=1}^{N_\mathrm{S}} A_j \left(\sum_{k=\kappa_j}^{K_j} |B_k|^2 \right)^{-s_j}, \qquad (21)$$

with $A_j = (s_j - 1)!/2\pi^{s_j}$, $\kappa_j = 1 + \sum_{n=1}^{j} s_{n-1}$, $K_j = \kappa_j + s_j - 1$ and $s_0 = 0$. Data streams from N_D detectors can be analyzed coherently by generalizing this to:

$$P(\mathbf{B} \mid \mathcal{H}_\mathrm{N}) = \prod_{i=1}^{N_\mathrm{D}} \prod_{j=1}^{N_{\mathrm{S}i}} A_{i,j} \left(\sum_{k=\kappa_{i,j}}^{K_{i,j}} |B_{i,k}|^2 \right)^{-s_{i,j}}, \quad (22)$$

where i indexes detectors, $B_{i,k} \equiv B_i(t_k)$ is the datum corresponding to the ith detector at time t_k, and $A_{i,j}$, $\kappa_{i,j}$ and $K_{i,j}$ are defined analogously to s_j, κ_j above. Note that the likelihood $p(\mathbf{B} \mid \vec{\theta}, \mathcal{H})$ of some hypothesis $\mathcal{H}$, is the probability of observing the data $\mathbf{B}$ assuming $\mathcal{H}$ is true and given a specific choice of free parameters $\vec{\theta}$ from the model's parameter space Θ; in this case, however, $\Theta = \varnothing$ and $p(\mathbf{B} \mid \vec{\theta}, \mathcal{H}_\mathrm{N}) = P(\mathbf{B} \mid \mathcal{H}_\mathrm{N})$ because $\mathcal{H}_\mathrm{N}$ has no free parameters.

The case of $\mathcal{H}_\mathrm{S}$ requires more attention. One could be tempted to use Eq. (19) to define a likelihood like Eq. (22) with the substitution $|B_k| \to |B_k - \Lambda_k|$, for $\Lambda_k \equiv \Lambda(t_k)$ including all polarizations like in Eq. (15); the priors would reflect uncertainties in measured source parameters and extend over reasonable ranges for a_p and ϕ_p. However, for most realistic prior choices, that would correspond to a hypothesis that assigns most of the prior probability to regions of parameter space for which $a_p \neq 0$ for all p, thus downweighting more conservative models (including GR) that we would like to prioritize. Furthermore, a hypothesis constructed in such a way would implicitly assign equal weight to conservative models as any other unphysical model with equal number of parameters. (For instance, a model that includes the $+$ mode, but not the $\times$ mode, while allowing ψ to vary, is unphysical.)

Instead we will construct $\mathcal{H}_\mathrm{S}$ from two logically disjoint component hypotheses:

1. *GR signal* ($\mathcal{H}_\mathrm{GR}$): the data contain Gaussian noise and a GR signal with the assumed $\phi(t)$, i.e.

$$\Lambda_\mathrm{GR}(t) = \frac{h_0}{2} \left[\frac{1}{2}(1 + \cos^2 \iota) F_+(t; \psi, \alpha, \delta) \right.$$
$$\left. - i \cos \iota F_\times(t; \psi, \alpha, \delta) \right] e^{i\phi_0}, \qquad (23)$$

for some fiducial phase ϕ_0 [in the notation of Eq. (8), $\phi_+ = \phi_0$ and $\phi_\times = \phi_0 - \pi/2$];

2. *non-GR signal* ($\mathcal{H}_\mathrm{nGR}$): the data contain Gaussian noise and a signal with non-GR polarization content, but with the assumed $\phi(t)$.

In turn, $\mathcal{H}_\mathrm{nGR}$ can itself be seen as a composite hypothesis encompassing all the signal models that depart

from GR in some way. For the least degenerate case, in which the source orientation with respect to the observer (parametrized by ψ and ι) is known, we identify two kinds of non-GR signal models:

1. *enhanced GR*: signal models with the usual GR prediction with the addition of extra modes;

2. *others*: signal models that do not constrain the tensor modes to have the relative amplitude and phase predicted by GR, or that do not contain tensor contributions at all.

We denote the non-GR subhypotheses of the first kind with a subscript indicating the spin of the additional modes. For example, we use "GR+s" (meaning "GR plus scalar") for a model containing a breathing mode in addition to the usual GR template; this corresponds to a signal of the form:

$$\Lambda_\mathrm{GR+s}(t) = \frac{h_0}{2} \left[\frac{1}{2}(1 + \cos^2 \iota) F_+(t; \psi, \alpha, \delta) \right.$$
$$\left. - i \cos \iota F_\times(t; \psi, \alpha, \delta) \right] e^{i\phi_0}$$
$$+ \frac{1}{2} a_b e^{i\phi_b} F_b(t; \psi, \alpha, \delta). \qquad (24)$$

On the other hand, we denote subhypotheses of the second kind with a subscript listing the spins included in the signal. For example, "st" (meaning "scalar plus tensor") corresponds to a model with unrestricted scalar and tensor contributions:

$$\Lambda_\mathrm{st}(t) = \frac{1}{2} \left[a_+ e^{i\phi_+} F_+(t; \psi, \alpha, \delta) + a_\times e^{i\phi_\times} F_\times(t; \psi, \alpha, \delta) \right.$$
$$\left. + a_b e^{i\phi_b} F_b(t; \psi, \alpha, \delta) \right]. \qquad (25)$$

The specific examples used above (GR+s and st) were chosen to also illustrate the fact that, when the source orientation is known, a signal template that includes the $+$ and $\times$ polarizations *without* restricting their relative amplitudes and phases is not, strictly speaking, a GR template (this becomes evident by noticing that the two alternatives have a different number of non-degenerate free parameters). When the source orientation is *not* known, however, this distinction is no longer meaningful and so we consider only non-GR models that do *not* include unrestricted tensor modes (t, st, tv, stv). Note that we could have instead chosen to remove models that explicitly incorporate a GR template (GR, GR+s, GR+v, GR+sv) for equivalent results. We refer the reader to Appendix A for further details.

For sources with known orientation, there are three models that explicitly contain a GR template (GR+s, GR+v, GR+sv) and seven which do not (s, t, v, st, sv, tv, stv), for a grand total of ten non-GR combinations. $\mathcal{H}_\mathrm{nGR}$ is constructed from the logical union ("or" junc-

tion, $\vee$) of all these:

$$\mathcal{H}_{\mathrm{nGR}} \equiv \mathcal{H}_{\mathrm{GR+s}} \vee \mathcal{H}_{\mathrm{GR+v}} \vee \mathcal{H}_{\mathrm{GR+sv}} \vee \mathcal{H}_{\mathrm{s}} \vee \mathcal{H}_{\mathrm{t}} \vee$$
$$\mathcal{H}_{\mathrm{v}} \vee \mathcal{H}_{\mathrm{st}} \vee \mathcal{H}_{\mathrm{sv}} \vee \mathcal{H}_{\mathrm{tv}} \vee \mathcal{H}_{\mathrm{stv}}$$
$$= \bigvee_{m \in \tilde{M}} \mathcal{H}_m. \tag{26}$$

where, for convenience, we have defined the non-GR subscript set $\tilde{M}$:

$$\tilde{M} \equiv \{\mathrm{GR+s,\ GR+v,\ GR+sv,}$$
$$\mathrm{s,\ t,\ v,\ st,\ sv,\ tv,\ stv}\}. \tag{27}$$

By the same token, the signal hypothesis can be built from the logical union of $\mathcal{H}_{\mathrm{GR}}$ and $\mathcal{H}_{\mathrm{nGR}}$:

$$\mathcal{H}_{\mathrm{S}} \equiv \mathcal{H}_{\mathrm{GR}} \vee \mathcal{H}_{\mathrm{nGR}} = \bigvee_{m \in M} \mathcal{H}_m, \tag{28}$$

with M defined similarly to $\tilde{M}$, but also including GR:

$$M \equiv \tilde{M} \cup \{GR\}. \tag{29}$$

As mentioned above, for sources of unknown orientation, we define $\mathcal{H}_{\mathrm{nGR}}$ and $\mathcal{H}_{\mathrm{S}}$ as above, except that we exclude hypotheses containing the unrestricted tensor modes, i.e. we make the replacement $M \to M' = M \setminus \{\mathrm{t, st, tv, stv}\}$; the procedure is otherwise identical. We keep the following discussion general, by using the cardinalities $|M|$ and $|\tilde{M}|$ to denote the number of any-signal and non-GR models respectively (i.e. $|M| = 11$, $|\tilde{M}| = 10$ for known orientation, and $|M| = 7$, $|\tilde{M}| = 6$ for unknown orientation).

Note that the validity of Eqs. (26, 28) is contingent on the mutual logical independence of all the $\mathcal{H}_m$'s. This requirement is satisfied by construction, since each of the $\mathcal{H}_m$'s is defined to exclude regions of parameter space that would correspond to other hypotheses nested within it (e.g. $\mathcal{H}_{\mathrm{GR+s}}$ is defined over all values of the scalar amplitude except $h_{\mathrm{b}} = 0$, to avoid including $\mathcal{H}_{\mathrm{GR}}$). In practice, however, it is not necessary to explicitly exclude these regions of parameter space, as will explained in Sec. III A 2 below.

2. Odds

We can construct a Bayesian model for $\mathcal{H}_{\mathrm{S}}$ starting from its components: for each subhypothesis $\mathcal{H}_m$ for $m \in M$, we use a likelihood function like Eq. (21) with the substitution $|B_{i,k}| \to |B_{i,k} - \Lambda_{m,i,k}|$, i.e.

$$p(\mathbf{B} \mid \vec{\theta}, \mathcal{H}_m) = \prod_{i=1}^{N_{\mathrm{D}}} \prod_{j=1}^{N_{\mathrm{S}i}} A_{i,j} \left(\sum_{k=\kappa_{i,j}}^{K_{i,j}} |B_{i,k} - \Lambda_{m,i,k}|^2 \right)^{-s_{i,j}} \tag{30}$$

(where $\Lambda_{m,i,k}$ is the template corresponding to model m, for detector i and time-bin k), and suitable priors
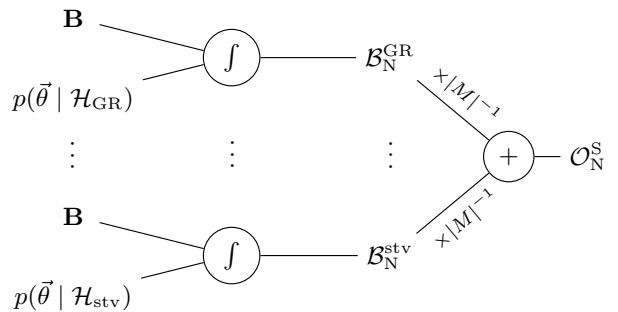

FIG. 2. *Computation of* $\mathcal{O}_{\mathrm{N}}^{\mathrm{S}}$. First, the Bayes factor $\mathcal{B}_{\mathrm{N}}^m$ is obtained from the data $\mathbf{B}$ and corresponding priors $p(\theta|\mathcal{H}_m)$ for each model $m \in M$, by evaluating the integral of Eq. (33) using a nested sampling algorithm that samples over $\vec{\theta}$ (step indicated by integral sign); these values are then added and multiplied by $P(\mathcal{H}_m)/P(\mathcal{H}_{\mathrm{N}})$ to obtain $\mathcal{O}_{\mathrm{N}}^{\mathrm{S}}$, as in Eq. (36). (Note that here we have set $P(\mathcal{H}_m)/P(\mathcal{H}_{\mathrm{N}}) = 1/|M|$, as explained Sec. IV.) The computation of $\mathcal{O}_{\mathrm{GR}}^{\mathrm{nGR}}$ is analogous.

on the model parameters $\vec{\theta}_m \in \Theta_m$; then, we combine the posteriors with priors on the models themselves to obtain the posterior for $\mathcal{H}_{\mathrm{S}}$. This last step allows us to incorporate our *a priori* beliefs about the validity of each of the components. This procedure is represented schematically in Fig. 2 and fleshed out below.

The choice of model priors can be made clearer by considering the posterior probability for the signal model. Given some set of detector data $\mathbf{B}$ and underlying assumptions I (suppressed from the following expressions), the posterior probability for $\mathcal{H}_{\mathrm{S}}$ is:

$$P(\mathcal{H}_{\mathrm{S}} \mid \mathbf{B}) = \sum_{m \in M} P(\mathcal{H}_m \mid \mathbf{B}) \tag{31}$$

by Eq. (28) and because the components are all logically independent [i.e. $\mathcal{H}_{m_1} \wedge \mathcal{H}_{m_2} = \mathrm{False}$, hence $P(\mathcal{H}_{m_1} \wedge \mathcal{H}_{m_2} \mid \mathbf{B}) = 0$ for any m_1, $m_2 \in M$ such that $m_1 \neq m_2$]. Note that this is true even for hypotheses that may contain each other as special cases. For instance, even though the GR template can be obtained from GR+s by setting the scalar amplitude to $h_{\mathrm{b}} = 0$, the points in the GR+s parameter space satisfying this condition define an infinitesimally-thin slice in hyperspace that offers no support to the prior distribution and is thus assigned no weight (see e.g. [21]).

We can expand each term on the RHS of Eq. (31) using Bayes' theorem:

$$P(\mathcal{H}_{\mathrm{S}} \mid \mathbf{B}) = \sum_{m \in M} P(\mathcal{H}_m) P(\mathbf{B} \mid \mathcal{H}_m) / P(\mathbf{B}). \tag{32}$$

Each of the terms on the RHS is made up of three factors: a marginalized likelihood $P(\mathbf{B} \mid \mathcal{H}_m)$, a prior on the model $P(\mathcal{H}_m)$, and a normalization constant $P(\mathbf{B})$.

The marginalized likelihood (also known as *evidence*)

is computed from the data:

$$P(\mathbf{B} \mid \mathcal{H}_m) = \int_{\Theta_m} p(\mathbf{B} \mid \vec{\theta}_m, \mathcal{H}_m) \, p(\vec{\theta}_m \mid \mathcal{H}_m) \, \mathrm{d}\vec{\theta}_m, \tag{33}$$

where $p(\mathbf{B} \mid \vec{\theta}_m, \mathcal{H}_m)$ is itself the likelihood. The evaluation of the multidimensional integral of Eq. (33) is the most computationally intensive part of our analysis (see Sec. IV for details).

We are free to choose the model priors (discussed in Sec. IV), as long as we satisfy the constraint:

$$P(\mathcal{H}_{\mathrm{N}}) + \sum_{m \in M} P(\mathcal{H}_m) = 1. \tag{34}$$

This is a statement about the exhaustiveness and disjointedness of the hypotheses we are considering: we assume that reality will agree with *one and only one* of the hypotheses at hand. (As we will see in Sec. VI, this assumption might not hold; for example, the noise may not be Gaussian.) The particular choice of prior for each model will encode our expectations about the corresponding theory (before seeing the data), and thus allow for some degree of subjectivity.

Note that we cannot directly compute $P(\mathbf{B})$ in a straightforward manner and without assuming that our hypothesis set is indeed exhaustive (which is not the case for non-Gaussian detector noise, see Sec. III A 4). However, the need for this computation can be avoided by looking at relative probabilities, i.e. *odds.* The odds for $\mathcal{H}_{\mathrm{S}}$ versus $\mathcal{H}_{\mathrm{N}}$ is defined as:

$$\mathcal{O}_{\mathrm{N}}^{\mathrm{S}} \equiv \frac{P(\mathcal{H}_{\mathrm{S}} \mid \mathbf{B})}{P(\mathcal{H}_{\mathrm{N}} \mid \mathbf{B})}. \tag{35}$$

Using Bayes' theorem again and canceling the $P(\mathbf{B})$ factors, this simplifies to:

$$\mathcal{O}_{\mathrm{N}}^{\mathrm{S}} = \frac{\sum P(\mathcal{H}_m) P(\mathbf{B} \mid \mathcal{H}_m)}{P(\mathcal{H}_{\mathrm{N}}) P(\mathbf{B} \mid \mathcal{H}_{\mathrm{N}})} = \sum_{m \in M} \frac{P(\mathcal{H}_m)}{P(\mathcal{H}_{\mathrm{N}})} \mathcal{B}_{\mathrm{N}}^m, \tag{36}$$

where, in the second equality, we have used the definition of the *Bayes factor*:

$$\mathcal{B}_j^i \equiv \frac{P(\mathbf{B} \mid \mathcal{H}_i)}{P(\mathbf{B} \mid \mathcal{H}_j)}, \tag{37}$$

for any two hypotheses $\mathcal{H}_i$, $\mathcal{H}_j$.

The odds in Eq. (36) can be used as a detection statistic to determine whether it is likely that the data contain a signal (of any polarization) or not. Once the presence of a signal has been established, a similar ratio can be constructed to assess agreement with GR:

$$\mathcal{O}_{\mathrm{GR}}^{\mathrm{nGR}} = \frac{P(\mathcal{H}_{\mathrm{nGR}} \mid \mathbf{B})}{P(\mathcal{H}_{\mathrm{GR}} \mid \mathbf{B})} = \sum_{m \neq GR} \frac{P(\mathcal{H}_m)}{P(\mathcal{H}_{\mathrm{GR}})} \mathcal{B}_{\mathrm{GR}}^m. \tag{38}$$

This ratio encodes the relative probability that there is a GR violation. Because it is now assumed that there is

a signal in the data, $P(\mathcal{H}_{\mathrm{N}}) = 0$ and the model priors must instead satisfy:

$$\sum_{m \in M} P(\mathcal{H}_m) = 1. \tag{39}$$

We can reduce the number of computations needed to obtain $\mathcal{O}_{\mathrm{N}}^{\mathrm{S}}$ and $\mathcal{O}_{\mathrm{GR}}^{\mathrm{nGR}}$ by using the fact that:

$$\mathcal{B}_j^i = \frac{P(\mathbf{B} \mid \mathcal{H}_i)}{P(\mathbf{B} \mid \mathcal{H}_j)} = \frac{P(\mathbf{B} \mid \mathcal{H}_i)}{P(\mathbf{B} \mid \mathcal{H}_{\mathrm{N}})} \frac{P(\mathbf{B} \mid \mathcal{H}_{\mathrm{N}})}{P(\mathbf{B} \mid \mathcal{H}_j)} = \frac{\mathcal{B}_{\mathrm{N}}^i}{\mathcal{B}_{\mathrm{N}}^j}. \tag{40}$$

This means that we need to evaluate an integral like Eq. (33) $|M|$ times per set of data, to compute $\mathcal{B}_{\mathrm{N}}^m$ for each m in M. Those $|M|$ numbers, together with the evidence for $\mathcal{H}_{\mathrm{N}}$, are enough to compute all the quantities of interest.

Instead of asking about a generic deviation from GR, we may also compare GR to a particular alternative theory. For such purpose, we will usually assign equal prior weight to GR and its alternative to compute:

$$\mathcal{O}_{\mathrm{GR}}^j = \frac{P(\mathcal{H}_j)}{P(\mathcal{H}_{\mathrm{GR}})} \mathcal{B}_{\mathrm{GR}}^j = \mathcal{B}_{\mathrm{GR}}^j, \tag{41}$$

where $\mathcal{H}_j$ may be any of the hypotheses in $\tilde{M}$ or an even more specific hypothesis. (The latter case demands an extra execution of the inference code.)

3. Multiple data sets

So far we have assumed that the data $\mathbf{B}$, corresponding to one or more GW detectors, can be analyzed coherently; however, there are cases in which we would like to combine results from sets of data analyzed incoherently. Examples are data sets corresponding to different sources or observation periods. Our Bayesian framework makes it possible to combine the respective odds in order to make an overall model selection statement (in our case, about the presence of signal or the validity of GR).

For instance, we may analyze data for N_{P} pulsars and ask about the probability that any of them contain a signal; treating each as an independent observation, the combined probability can be constructed from the odds above. Letting $\mathcal{H}_{\mathrm{S}_i}$, $\mathcal{H}_{\mathrm{N}_i}$ respectively denote signal and noise hypotheses for the ith source, while $\mathcal{H}_{\mathrm{S_{any}}}$ corresponds to a signal being present in *any* of the sources and $\mathcal{H}_{\mathrm{N_{all}}}$ corresponds to Gaussian noise in data for *all* sources:

$$\begin{aligned} {}^{(N_{\mathrm{P}})}\mathcal{O}_{\mathrm{N_{all}}}^{\mathrm{S_{any}}} &= \frac{P(\mathcal{H}_{\mathrm{S_{any}}} \mid \mathbf{B})}{P(\mathcal{H}_{\mathrm{N_{all}}} \mid \mathbf{B})} = \frac{1 - P(\mathcal{H}_{\mathrm{N_{all}}} \mid \mathbf{B})}{P(\mathcal{H}_{\mathrm{N_{all}}} \mid \mathbf{B})} \\ &= \frac{1}{P(\bigwedge_i \mathcal{H}_{\mathrm{N}_i} \mid \mathbf{B})} - 1 = \prod_{i=1}^{N_{\mathrm{P}}} \frac{1}{P(\mathcal{H}_{\mathrm{N}_i} \mid \mathbf{B}_i)} - 1 \\ &= \prod_{i=1}^{N_{\mathrm{P}}} \frac{P(\mathcal{H}_{\mathrm{S}_i} \mid \mathbf{B}_i) + P(\mathcal{H}_{\mathrm{N}_i} \mid \mathbf{B}_i)}{P(\mathcal{H}_{\mathrm{N}_i} \mid \mathbf{B}_i)} - 1 \\ &= \prod_{i=1}^{N_{\mathrm{P}}} \left(\mathcal{O}_{\mathrm{N}_i}^{\mathrm{S}_i} + 1 \right) - 1, \end{aligned} \tag{42}$$

where we have used the exclusivity and exhaustiveness of the signal and noise hypotheses, i.e.

$$P(\mathcal{H}_{S_{\text{any}}} \mid \mathbf{B}) + P(\mathcal{H}_{N_{\text{all}}} \mid \mathbf{B}) = 1, \qquad (43)$$

$$P(\mathcal{H}_{S_i} \mid \mathbf{B}_i) + P(\mathcal{H}_{N_i} \mid \mathbf{B}_i) = 1, \qquad (44)$$

with i indexing data sets. Note that the data sets for different sources ($\mathbf{B}_i$'s) are *not* conditionally independent under $\mathcal{H}_{S_{\text{any}}}$ or $\mathcal{H}_{N_{\text{all}}}$. Also, Eq. (42) does not enforce the requirement that, if signals are present in multiple sources, they all correspond to the same model from Eq. (29); such a constraint could be implemented at this stage, but is more easily enforced by examining individual values of $\mathcal{O}_N^m$ when necessary.

The construction of Eq. (42) implicitly assigns model priors to each of the meta-hypotheses $\mathcal{H}_{S_{\text{any}}}$ and $\mathcal{H}_{N_{\text{all}}}$ such that:

$$\frac{P(\mathcal{H}_{S_{\text{any}}})}{P(\mathcal{H}_{N_{\text{all}}})} = \left[\frac{P(\mathcal{H}_S)}{P(\mathcal{H}_N)} + 1\right]^{N_P} - 1, \qquad (45)$$

where we have assumed the priors for signal vs noise are equal for all sources, i.e. $P(\mathcal{H}_{S_i}) = P(\mathcal{H}_S)$ and $P(\mathcal{H}_{N_i}) = P(\mathcal{H}_N)$ for all i. When making combined statements for multiple sources, we may wish to choose $P(\mathcal{H}_S)/P(\mathcal{H}_N)$ such as to produce any desired value of $P(\mathcal{H}_{S_{\text{any}}})/P(\mathcal{H}_{N_{\text{all}}})$, say $P(\mathcal{H}_{S_{\text{any}}}) = P(\mathcal{H}_{N_{\text{all}}})$. Furthermore, one may wish to weight each pulsar differently within $\mathcal{H}_{S_{\text{any}}}$ by incorporating information about the source distance (or other parameters) into the priors via a parametrization like Eq. (12); this may improve the sensitivity of the ensemble odds to weak signals in the set, as suggested in [22]. However, using such a parametrization generally implies committing to a specific gravitational theory (or family of theories). We choose not to take such approach in this study.

Besides combining data for multiple pulsars, for a given source, we could also (incoherently) combine the results of analyses using data from different observation periods. Since the astrophysical CWs we are considering should either be present in all N_R observation runs or in none of them, the relevant odds are:

$$
\begin{aligned}
{}^{[N_R]}\mathcal{O}_N^S &= \frac{P(\mathcal{H}_S \mid \mathbf{B})}{P(\mathcal{H}_N \mid \mathbf{B})} = \sum_{m \in M} \frac{P(\mathcal{H}_m \mid \mathbf{B})}{P(\mathcal{H}_N \mid \mathbf{B})} \\
&= \sum_{m \in M} \frac{P(\mathbf{B} \mid \mathcal{H}_m)P(\mathcal{H}_m)}{P(\mathbf{B} \mid \mathcal{H}_N)P(\mathcal{H}_N)} \\
&= \sum_{m \in M} \frac{P(\mathcal{H}_m)}{P(\mathcal{H}_N)} \prod_{j=1}^{N_R} (\mathcal{B}_N^m)_j, \qquad (46)
\end{aligned}
$$

where we have again used $\mathbf{B} = \{\mathbf{B}_j\}_{j=1}^{N_R}$ to refer to the totality of data, with j indexing observation runs. The independence of the $\mathbf{B}_j$'s, conditional on $\mathcal{H}_m$ and $\mathcal{H}_N$, is applied on the last line to write the result in terms of the individual Bayes factors for each run, $(\mathcal{B}_N^m)_j$.

Similarly, we can use multiple data sets to make a single statement about deviations from GR. Once we have made N_P detections from different sources, the odds for a GR violation is:

$$
{}^{(N_P)}\mathcal{O}_{GR}^{nGR} = \sum_{m \in \tilde{M}} \frac{P(\mathcal{H}_m)}{P(\mathcal{H}_{GR})} \prod_{i=1}^{N_P} (\mathcal{B}_{GR}^m)_i, \qquad (47)
$$

where, again, i indexes sources. (See Sec. IIID of [21] for an analogous derivation.)

4. Non-Gaussian noise

Up to this point, like most other CW studies, we have assumed that the detector noise is Gaussian. However, although previous work has indicated that this is generally a very good approximation [8, 23], it is not exactly true for actual detector noise (for some frequencies more so than others). Happily, most of the model selection statements expounded so far are valid also in the presence of non-Gaussian instrumental noise, after some light reinterpretation.

If the assumption of Gaussianity does not hold, the hypotheses constructed in Sec. III A 1 are no longer exhaustive: the data may not only be explained by Gaussian noise or a signal (GR or otherwise), but also by non-Gaussian artifacts that are impossible to satisfactorily model. Nevertheless, the computation and interpretation of evidences and odds remain unchanged for all the hypotheses under consideration.

Because "noise" no longer just means "Gaussian noise", $\mathcal{O}_N^S$ (which compares the signal model vs *Gaussian* noise) has to be treated more carefully for detection purposes. Indeed, instrumental features that are clearly non-Gaussian (e.g. a loud, narrow-band artifact wandering across the frequency of interest) will generally result in a relatively large value of $\mathcal{O}_N^S$, even if there is no detectable astrophysical signal in the data. This issue affects the standard GR searches as well [23], although perhaps to a lesser degree due to the reduced signal parameter space.

It is possible to mitigate this problem by constructing a hypothesis that captures some key characteristic of instrumental features and helps discriminate those from real astrophysical signals. Perhaps the best way to do this is to take advantage of the fact that an astrophysical CW must manifest itself coherently across detectors, while the same is not true for detector artifacts [24]. We can thus define an *instrumental feature hypothesis* ($\mathcal{H}_I$) to encompass the cases in which the data are composed of Gaussian noise, or features that look like astrophysical signals but are not coherent across detectors (viz. they do not have a consistent phase evolution and they are best described by different waveform parameters).

Formally, we define $\mathcal{H}_\mathrm{I}$ by:

$$\mathcal{H}_\mathrm{I} \equiv \bigwedge_{d=1}^{N_\mathrm{D}} \left(\mathcal{H}_{\mathrm{S}_d} \vee \mathcal{H}_{\mathrm{N}_d} \right), \qquad (48)$$

where the subscript d identifies detectors. This definition does not explicitly encompass instrumental features that are coherent across some subset of the detectors. Also, note that Eq. (48) implicitly contains a term equivalent to the usual noise hypothesis $\mathcal{H}_\mathrm{N} = \bigwedge_d \mathcal{H}_{\mathrm{N}_d}$. Similarly, it also contains a term corresponding to the presence of signals in all detectors $(\bigwedge_d \mathcal{H}_{\mathrm{S}d})$. Importantly, such term is not equivalent to the *coherent* signal hypothesis $\mathcal{H}_\mathrm{S}$, as given by the multi-detector likelihood of Eq. (30):

$$\mathcal{H}_\mathrm{S} \neq \bigwedge_{d=1}^{N_\mathrm{D}} \mathcal{H}_{\mathrm{S}d}. \qquad (49)$$

While the evidence integral of Eq. (33) factorizes into single-detector terms for $\mathcal{H}_\mathrm{N}$ (due to the null parameter space), the same is not true for $\mathcal{H}_\mathrm{S}$. Furthermore, because it does not demand detector coherence, the RHS of Eq. (49) is associated with a considerably larger parameter space than the LHS. Thus, in the presence of an astrophysical signal, model selection will favor the latter due to its smaller Ockham's penalty. The same is true, of course, when comparing $\mathcal{H}_\mathrm{S}$ to $\mathcal{H}_\mathrm{I}$ as a whole.

From Eq. (48), it is straightforward to write the evidence for $\mathcal{H}_\mathrm{I}$ as

$$P(\mathbf{B} \mid \mathcal{H}_\mathrm{I}) = \prod_{d=1}^{N_\mathrm{D}} [P(\mathbf{B}_d \mid \mathcal{H}_{\mathrm{S}_d})P(\mathcal{H}_{\mathrm{S}_d} \mid \mathcal{H}_\mathrm{I})$$
$$+ P(\mathbf{B}_d \mid \mathcal{H}_{\mathrm{N}_d})P(\mathcal{H}_{\mathrm{N}_d} \mid \mathcal{H}_\mathrm{I})] \qquad (50)$$

and use this to construct the odds comparing against $\mathcal{H}_\mathrm{S}$:

$$\mathcal{O}_\mathrm{I}^\mathrm{S} = \frac{P(\mathcal{H}_\mathrm{S})}{P(\mathcal{H}_\mathrm{I})} \frac{\mathcal{B}_\mathrm{N}^\mathrm{S}}{\prod_{d=1}^{N_\mathrm{D}} \left[P(\mathcal{H}_{\mathrm{S}_d} \mid \mathcal{H}_\mathrm{I})(\mathcal{B}_{\mathrm{N}_d}^{\mathrm{S}_d} - 1) + 1 \right]}. \qquad (51)$$

Here we have used Eq. (50), together with the fact that $P(\mathcal{H}_{\mathrm{S}_d} \mid \mathcal{H}_\mathrm{I}) + P(\mathcal{H}_{\mathrm{N}_d} \mid \mathcal{H}_\mathrm{I}) = 1$ and $P(\mathbf{B} \mid \mathcal{H}_\mathrm{N}) = \prod_d P(\mathbf{B} \mid \mathcal{H}_{\mathrm{N}d})$, to write $\mathcal{O}_\mathrm{I}^\mathrm{S}$ as a function of the detector-coherent signal vs noise Bayes factor $\mathcal{B}_\mathrm{N}^\mathrm{S}$, the single-detector signal vs noise Bayes factors $\mathcal{B}_\mathrm{N}^{\mathrm{S}_d}$, and model priors $P(\mathcal{H}_\mathrm{S})$, $P(\mathcal{H}_\mathrm{I})$ and $P(\mathcal{H}_{\mathrm{S}_d} \mid \mathcal{H}_\mathrm{I})$.

As usual, we are free to choose the model priors to give more or less weight to different hypotheses. For example, we recover choice of [23] (Appendix A3) by setting $P(\mathcal{H}_{\mathrm{S}_d} \mid \mathcal{H}_\mathrm{I}) = 0.5$ for all d and $P(\mathcal{H}_\mathrm{S}) = P(\mathcal{H}_\mathrm{I}) \times 0.5^{N_\mathrm{D}}$ such that:

$$\ln \mathcal{O}_\mathrm{I}^\mathrm{S} = \ln \mathcal{B}_\mathrm{N}^\mathrm{S} - \sum_{d=1}^{N_\mathrm{D}} \ln \left(\mathcal{B}_{\mathrm{N}_d}^{\mathrm{S}_d} + 1 \right). \qquad (52)$$

(When comparing to Appendix A3 of [23], however, note that in that work "I" is used to denote both the background information and the "incoherent-signal-or-noise" hypothesis, which can be identified with our $\mathcal{H}_\mathrm{I}$.)

There is reason to believe that $\ln \mathcal{O}_\mathrm{I}^\mathrm{S}$, with model priors as in Eq. (52), is quite good at picking out instrumental features, even for data from just two instruments [23]. (Note that we would expect the discriminatory power of $\ln \mathcal{O}_\mathrm{I}^\mathrm{S}$ to grow with the number of detectors available.) However, at the end of the day, we can never be fully confident that $\mathcal{H}_\mathrm{I}$ will indeed capture all non-astrophysical disturbances. To address this, we may always treat $\ln \mathcal{O}_\mathrm{N}^\mathrm{S}$ and $\ln \mathcal{O}_\mathrm{I}^\mathrm{S}$ as any generic detection statistic and use estimates of the background distribution to establish significance.

B. Parameter estimation

Besides choosing between different models, we can use Bayesian statistics to obtain posterior probability density functions (PDFs) on the parameters of a given template (*first–level inference*). In the absence of a loud signal, this can be used to obtain credible intervals that yield upper–limits on the amplitudes of GR deviations.

For a model $\mathcal{H}$ with N parameters, an N-dimensional posterior PDF covering the parameter space Θ can be obtained from Bayes' theorem:

$$p(\vec{\theta} \mid \mathbf{B}, \mathcal{H}) = \frac{p(\mathbf{B} \mid \vec{\theta}, \mathcal{H}) \, p(\vec{\theta} \mid \mathcal{H})}{P(\mathbf{B} \mid \mathcal{H})}, \qquad (53)$$

for $\vec{\theta}$ in Θ, and with $p(\vec{\theta} \mid \mathcal{H})$ the prior over Θ. To obtain a one-dimensional PDF for a single parameter (call it θ_i), the N-dimensional distribution must be marginalized over all nuisance parameters (viz. all parameters except θ_i):

$$p(\theta_i \mid \mathbf{B}, \mathcal{H}) = \int_{\Theta'} p(\vec{\theta} \mid \mathbf{B}, \mathcal{H}) \, \mathrm{d}^{N-1}\theta_j$$
$$\propto \int_{\Theta'} p(\mathbf{B} \mid \vec{\theta}, \mathcal{H}) \, p(\vec{\theta} \mid \mathcal{H}) \, \mathrm{d}^{N-1}\theta_j, \quad (54)$$

where $0 < j \leq N$, such that $j \neq i$, and Θ' denotes the parameter space Θ with the ith dimension removed. Note that the equality has been replaced by a relation of proportionality because we have excluded the evidence $P(\mathbf{B} \mid \mathcal{H})$ from the expression. (Although of great importance for model selection, this quantity is uninteresting for the purposes of parameter estimation and can be treated as a simple normalization constant.) As discussed in Sec. IV, we evaluate Eq. (54) using the same nested-sampling algorithm used to compute the evidence.

Eq. (54) can be used to place upper limits on model parameters; in particular, we will use it to place limits on the amplitude of GR deviations. Consider, for instance, the case of a scalar-tensor theory that can be encapsulated by our GR+s model as described in the previous section; the 95%–credible upper limit on the strength of the breathing mode is $h_\mathrm{b}^{95\%}$, defined by:

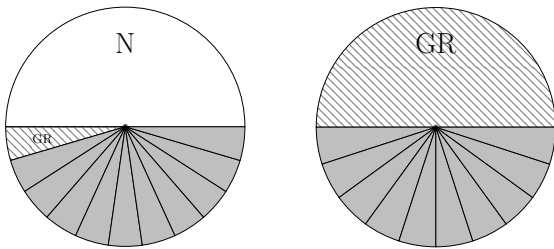$$0.95 = \int_{\min(h_\mathrm{b})}^{h_\mathrm{b}^{95\%}} p(h_\mathrm{b} \mid \mathbf{B}, \mathcal{H}_\mathrm{GR+s}) \, \mathrm{d}h_\mathrm{b}, \qquad (55)$$

FIG. 3. *Model priors.* Distribution of prior probability over subhypotheses for the construction of $\mathcal{O}_{\mathrm{N}}^{\mathrm{S}}$ (left) and $\mathcal{O}_{\mathrm{GR}}^{\mathrm{nGR}}$ (right), according to Eqs. (36) and (38) respectively. For $\mathcal{O}_{\mathrm{N}}^{\mathrm{S}}$, we assign equal weight to the $\mathcal{H}_{\mathrm{N}}$ (white) and $\mathcal{H}_{\mathrm{S}}$ (gray); as in Eq. (57), we make no *a priori* distinction between non-GR models (solid) and GR (hatched). For $\mathcal{O}_{\mathrm{GR}}^{\mathrm{nGR}}$, we set equal prior probability for $\mathcal{H}_{\mathrm{GR}}$ and $\mathcal{H}_{\mathrm{nGR}}$, distributing the prior equally among non-GR models, as in Eq. (59).

where $\min(h_{\mathrm{b}})$ is the minimum value of h_{b} allowed by the prior.

Note that there may be reasons to compute posteriors under different priors than when computing Bayes factors. In particular, it is conventional to present upper limits obtained using a uniform prior over some broad range of the amplitude parameters. With a uniform prior, the posterior is trivially related to the likelihood. This approach produces a more conservative upper limit than other choices, e.g. a Jeffreys prior (see Appendix B).

IV. ANALYSIS

We quantify our ability to use Bayesian model selection to detect CW signals and determine their polarization content as described above. To do this, we use one year of simulated data from three advanced interferometric detectors at design sensitivity: LIGO Hanford (H1), LIGO Livingston (L1) and Virgo (V1). Detector noise is simulated by drawing from a Gaussian distribution with zero mean and standard deviation corresponding to the power spectral density (PSD) of each detector at the frequency of the pulsar. (Previous work has shown that these are good assumptions for actual reduced detector data [8, 23].)

As mentioned in the previous section, the key step in our analysis is the computation of the evidence integral of Eq. (33) for the 12 hypotheses under consideration (one noise model plus 11 signal submodels). We carry this out using a version of the Bayesian inference code used for targeted pulsar search in [25, 26], which we modified to handle signal models other than GR. This inference code is itself built on the implementation of Skilling's nested-sampling algorithm [27] in the `LALInference` package [28], part of the LIGO Algorithm Library Suite [29]. This is the same inference software used for CBC analyses, including GW150914 [30].

In computing likelihoods, we take source location, frequency and frequency derivatives as known quantities (relevant uncertainties are negligible for this analysis). This is also true for orientation parameters $\cos\iota$ and ψ, when such measurements exist (Table 3 in [6]), although we may instead choose to marginalize over measurement errors; in particular, all plots for the Crab pulsar (PSR J0534+2200) in Sec. V are produced using known values of orientation parameters. For most pulsars, however, we lack any orientation information and so $\cos\iota$ and ψ are treated as unknowns with flat priors over appropriate ranges [31]. Unless otherwise indicated, priors uniform in the logarithm are used for amplitude parameters (h_0 or h_p's), since these are the least informative priors for scaling coefficients (also known as "Jeffreys priors") [32]; we make the somewhat arbitrary choice of restricting the strain amplitudes to the 10^{-28}–10^{-24} range. Flat priors are placed over all phase offsets (ϕ_0 and all the ϕ_p's).

We follow common practice by, according to the principle of indifference (see e.g. Ch. 5 of [33]), assigning equal prior probability to the signal and noise models, i.e. we let

$$P(\mathcal{H}_{\mathrm{S}}) = P(\mathcal{H}_{\mathrm{N}}) = 1/2 \ . \tag{56}$$

We must also decide how to split the prior among the different $\mathcal{H}_m$'s when computing $\mathcal{O}_{\mathrm{N}}^{\mathrm{S}}$ and $\mathcal{O}_{\mathrm{GR}}^{\mathrm{nGR}}$. In the former case we choose to distribute the prior weight uniformly among all signal models, so that:

$$P(\mathcal{H}_m) = |M|^{-1}/2, \tag{57}$$

with $|M|$ the cardinality of M (i.e. the number of signal models). In the latter, however, we prioritize GR by setting:

$$P(\mathcal{H}_{\mathrm{GR}} \mid \mathcal{H}_{\mathrm{S}}) = 1/2, \tag{58}$$

$$P(\mathcal{H}_m \mid \mathcal{H}_{\mathrm{S}}) = |\tilde{M}|^{-1}/2. \tag{59}$$

For example, in the case of fixed source orientation, where we consider all 10 non-GR models, we have $|M| = 11$ and $|\tilde{M}| = 10$, respectively implying $P(\mathcal{H}_m) = 1/22$ and $P(\mathcal{H}_m|\mathcal{H}_{\mathrm{S}}) = 1/20$ for the two cases above. This distribution is illustrated schematically in Fig. 3.

Note that these might not be the only justifiable options; for example, we might want to prioritize $\mathcal{H}_{\mathrm{GR}}$ when constructing $\mathcal{H}_{\mathrm{S}}$ in order to better deal with a noise background that does not conform to our assumption of Gaussianity. (Other strategies to tackle non-Gaussian noise are discussed in Sec. III A 4.) In any case, the code is sufficiently flexible to make different choices for the model priors if desired.

V. RESULTS

A. Model selection

1. Signal versus noise

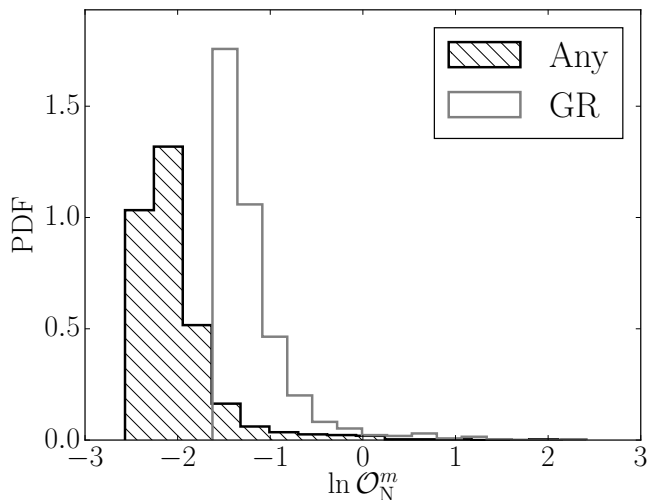We first show that $\mathcal{O}_{\mathrm{N}}^{\mathrm{S}}$, as defined in Eq. (36), can be used to discriminate signals of any polarization from

FIG. 4. *Signal vs noise log-odds background distributions for any-signal and GR hypotheses.* Histograms of $\ln \mathcal{O}_\mathrm{N}^\mathrm{S}$ (black line, hatched) and $\ln \mathcal{O}_\mathrm{N}^\mathrm{GR}$ (gray line) over an ensemble of 1000 simulated noise instantiations corresponding to the Crab pulsar. For each instantiation, three time series of Gaussian noise were produced using the design noise spectra of H1, L1 and V1, as outlined in section IV; the data are analyzed coherently across detectors. (Note that here $\ln \mathcal{O}_\mathrm{N}^\mathrm{GR} = \ln \mathcal{B}_\mathrm{N}^\mathrm{GR}$, since we assign equal weight to both models.)

Gaussian noise, without significant loss of sensitivity to GR signals. The black histogram in Fig. 4 shows the distribution of the natural logarithm of this quantity (henceforth, "log-odds"), obtained from the analysis of an ensemble of noise instantiations corresponding to a single source—in this case, the Crab pulsar. For comparison, the gray histogram in Fig. 4 is the analogous distribution for $\ln \mathcal{B}_\mathrm{N}^\mathrm{GR}$ (note that $\mathcal{B}_\mathrm{N}^\mathrm{GR} = \mathcal{O}_\mathrm{N}^\mathrm{GR}$ if we assign equal priors to the GR and Gaussian noise models, cf. Eq. (36) with $m = \mathrm{GR}$); this is the value computed in regular, GR-only targeted pulsar searches, although with different signal amplitude priors [23].

For both quantities shown in Fig. 4, a negative value marks a preference for the noise model ($\mathcal{H}_\mathrm{N}$, as defined at the beginning of section III A 1). However, note that a conservative (as determined by the priors) analysis should not be expected to strongly favor $\mathcal{H}_\mathrm{N}$, since the presence of a weak signal below the noise threshold cannot be discarded; this explains why the ranges in the plots of Fig. 4 do not extend to more negative values. Generally speaking, the magnitude of the signal prior volume (viz. the volume of parameter space allowed by the signal model, weighted by the prior function) will determine the mean of background distributions like Fig. 4, which will be more negative the greater the signal volume. This is a manifestation of an implicit Ockham's penalty.

The relationship between the odds for different signal hypotheses vs noise is illustrated in Fig. 5, which shows violin plots representing the noise-ensemble distributions of log-odds for GR (left panel), all non-GR models (center

panel) and the signal hypotheses constructed from them (right panel), all of them versus noise. (The construction of $\mathcal{H}_\mathrm{S}$ is outlined in section III A 1, Eq. (36) in particular.) Interestingly, Fig. 5 shows the relationship between $\ln \mathcal{B}_\mathrm{N}^m$ (equivalently, $\ln \mathcal{O}_\mathrm{N}^m$) and the number of degrees of freedom (a proxy for the prior volume) of model m: models with more degrees of freedom have a greater prior volume and get correspondingly downweighted, resulting in more negative values of $\ln \mathcal{B}_\mathrm{N}^m$.

If the data contain a sufficiently loud signal of any polarization, the evidence for $\mathcal{H}_\mathrm{S}$ will surpass that for $\mathcal{H}_\mathrm{N}$, and this can be used to establish a detection. Fig. 6 shows the response of $\ln \mathcal{O}_\mathrm{N}^\mathrm{S}$ and $\ln \mathcal{B}_\mathrm{N}^\mathrm{GR}$ to the presence of GR and non-GR signals. The general features of these plots confirm our expectations that for weak, subthreshold signals, the analysis should not be able to distinguish between the signal and noise models, yielding a Bayes factor close to unity (more precisely, a value of $\ln \mathcal{O}_\mathrm{N}^\mathrm{S}$ consistent with the background distributions of Fig. 4). Note that, in agreement with Fig. 4, the noise baseline for $\ln \mathcal{O}_\mathrm{N}^\mathrm{S}$ lies below that of $\ln \mathcal{B}_\mathrm{N}^\mathrm{GR}$, due to its greater prior volume.

For stronger (detectable) signals, the basic form of our likelihood functions, Eq. (20), leads us to expect $\ln \mathcal{O}_\mathrm{N}^\mathrm{S}$ to scale linearly with the square of the signal-to-noise ratio (SNR):

$$\ln \mathcal{B}_\mathrm{N}^m \sim \left(\mathbf{B} \cdot \boldsymbol{\Lambda}_\mathrm{MP} - |\boldsymbol{\Lambda}_\mathrm{MP}|^2/2 \right) /\sigma^2 \propto (h_\mathrm{inj}/\sigma)^2 , \quad (60)$$

where the standard deviation σ proxies the PSD and we let $\boldsymbol{\Lambda}_\mathrm{MP}$ be the time-series vector corresponding to the maximum probability template; for a stationary PSD, this implies $\ln \mathcal{B}_\mathrm{N}^m \propto h_\mathrm{inj}^2$, as observed in Fig. 6. The spread around the trendline is due to the individual features of each noise instantiation and (much less so) to numerical errors in the computation of the evidence, Eq. (33). For details on numerical uncertainty, see Appendix C.

From the left panel of Fig. 6, we conclude that $\ln \mathcal{O}_\mathrm{N}^\mathrm{S}$ can be as good an indicator of the presence of GR signals as $\ln \mathcal{B}_\mathrm{N}^\mathrm{GR}$ itself; this implies that we may include non-GR polarizations in our search without significantly sacrificing sensitivity to GR signals. However, the power of $\ln \mathcal{O}_\mathrm{N}^\mathrm{S}$ lies in responding also to non-GR signals. As an example of this, the right panel in Fig. 6 shows $\ln \mathcal{O}_\mathrm{N}^\mathrm{S}$ and $\ln \mathcal{B}_\mathrm{N}^\mathrm{GR}$ as a function of the amplitude of a fully non-GR injection. Here, we have chosen to inject a particular model of vector signal developed in [11], but the results are generic.

Note that, for sufficiently loud signals, $\mathcal{H}_\mathrm{GR}$ becomes preferable over $\mathcal{H}_\mathrm{N}$ (hence $\ln \mathcal{B}_\mathrm{N}^\mathrm{GR} > 0$), even when the injection model does not match the search; this is because the noise evidence drops faster than GR's and becomes very small (i.e. the data do not look at all like Gaussian noise, although they do not match the expected GR signal well either). The particular SNR at which this occurs will depend on the overlap between the antenna patterns of the injection and those of GR, and will consequently vary among sources.
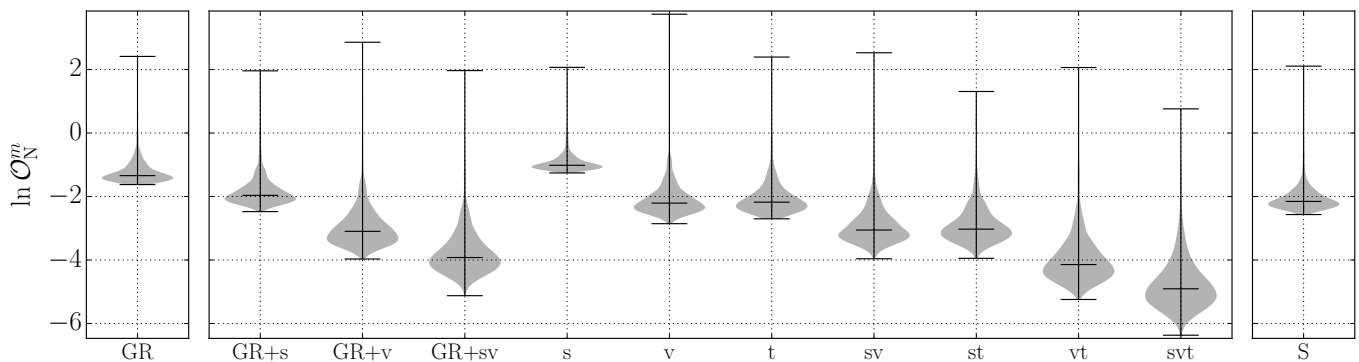
FIG. 5. *Signal vs noise log-odds background distributions for all subhypotheses.* Violin plots representing histograms of the log-odds of several models versus noise, computed over an ensemble of 1000 simulated noise instantiations each corresponding to H1, L1 and V1 design data prepared for the Crab pulsar; the data are analyzed coherently across detectors. As indicated by the labels on the x-axis, the left panel results from an analysis using the GR model, the center panel from all non-GR models and the right panel from the all-signal model; the latter is produced from the combination of the GR and non-GR results (section III A 1). Except for the any-signal case, the quantity histogrammed is just $\ln \mathcal{B}_{\mathrm{N}}^m$, which is the same as $\ln \mathcal{O}_{\mathrm{N}}^m$ if neither $\mathcal{H}_m$ nor $\mathcal{H}_{\mathrm{N}}$ are favored *a priori* (hence the label on the y-axis). Black lines mark the range and median of each distribution. Note that the latter is correlated to the number of degrees of freedom (prior volume) of the model, *viz.* models with more variables yield more negative $\ln \mathcal{B}_{\mathrm{N}}^m$'s due to their larger Ockham's penalty. (The black and gray histograms in Fig. 4 respectively show the same distributions as the right- and leftmost panels here.)
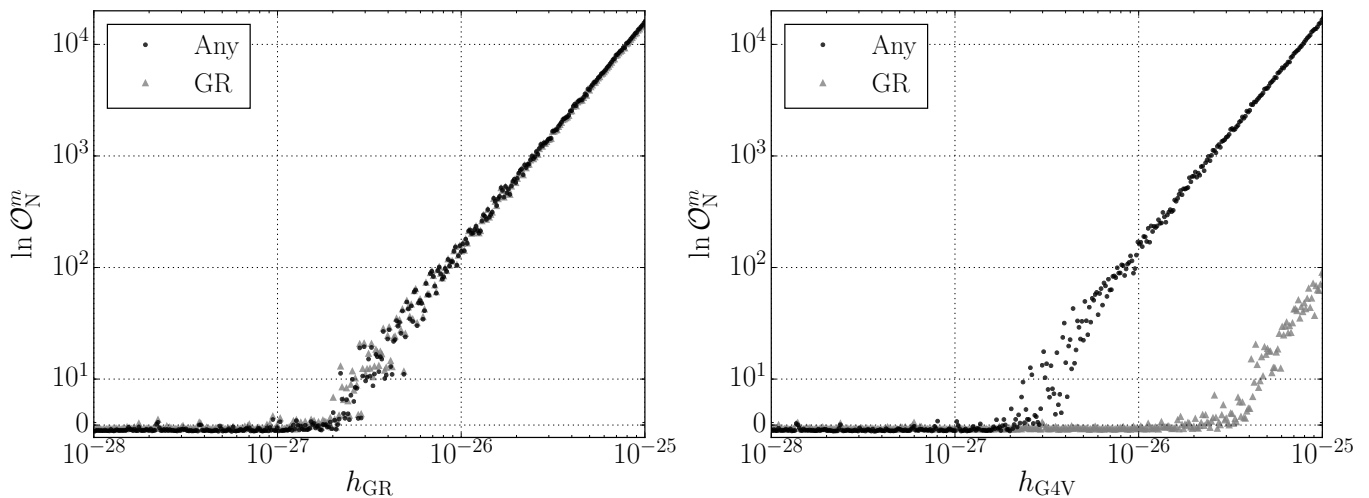
FIG. 6. *Expected sensitivity to GR and vector injections.* Log-odds of any-signal ($\mathcal{H}_{\mathrm{S}}$, black circles) and GR ($\mathcal{H}_{\mathrm{GR}}$, gray triangles) versus noise ($\mathcal{H}_{\mathrm{N}}$) hypotheses, as a function of injection amplitude, for signals corresponding to both GR (left) and a specific vector-only model, denoted G4V (right) [11]. The any-signal odds is defined in Eq. (36). The amplitudes are defined by $h_{\mathrm{GR}}^2 = a_+^2 + a_\times^2$ and $h_{\mathrm{G4V}}^2 = a_\times^2 + a_\mathrm{y}^2$. Each of the 500 points corresponds to a data instantiation (one time series for each detector: H1, L1 and V1) made up of Gaussian noise plus a simulated Crab-pulsar signal of the indicated strength. The injections were performed with random values of the nuisance phase parameters, and the data were analyzed coherently across detectors. A logarithmic scale is used for the y-axis, except for a linear stretch corresponding to the first decade. While the any-signal odds is equally sensitive to tensor and vector signals, the GR analysis is only able to distinguish vector signals from noise for extremely loud injections.

For the interesting case of scalar-tensor theories (for us, templates composed of GR plus an extra breathing component, and denoted "GR+s"), the behavior is slightly different. This is both because GR+s has an extra amplitude degree of freedom (a_b) and, as discussed in section III A 1, because $\mathcal{H}_{\mathrm{GR}}$ can be recovered as a special case of $\mathcal{H}_{\mathrm{GR+s}}$ (namely, when $a_\mathrm{b} \to 0$). In Fig. 7, we present the log-odds of signal versus noise hypotheses as a function of injected GR ($h_{\mathrm{GR}} \equiv \sqrt{a_+^2 + a_\times^2}$, x-axis) and scalar ($h_\mathrm{b} \equiv a_\mathrm{b}$, y-axis) strengths. These plots divide the h_b–h_{GR} plane in roughly two regions where the corresponding signal model ($\mathcal{H}_{\mathrm{S}}$, $\mathcal{H}_{\mathrm{GR}}$ or $\mathcal{H}_{\mathrm{GR+s}}$) is preferred (black) and where it is not (red). The latter cor-
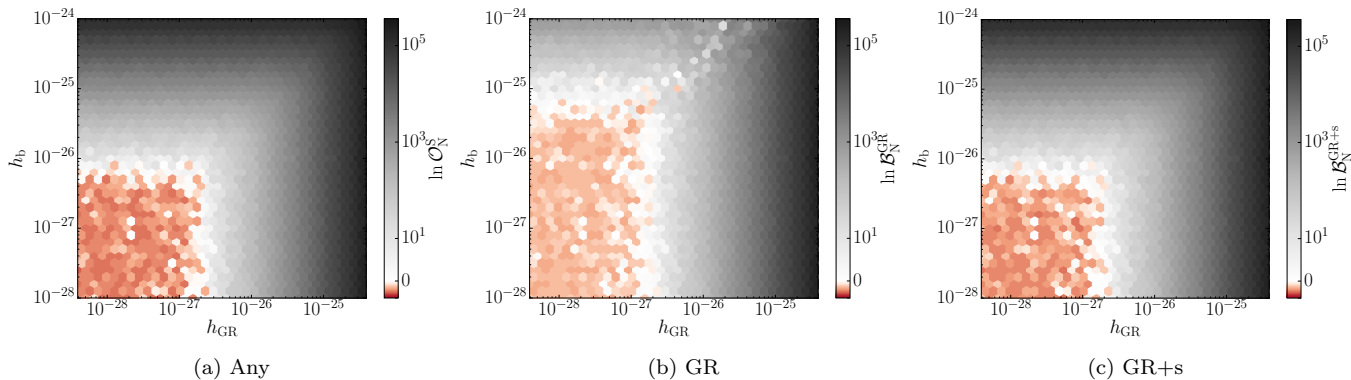
(a) Any   (b) GR   (c) GR+s

FIG. 7. *Expected sensitivity to scalar-tensor injections.* Log-odds of any-signal ($\mathcal{H}_\mathrm{S}$, left), GR ($\mathcal{H}_\mathrm{GR}$, center) and GR+s ($\mathcal{H}_\mathrm{GR+s}$, right) hypotheses versus noise. The any-signal odds is defined in Eq. (36). Each plot was produced by analyzing 2500 instantiations of data (one time series for each detector: H1, L1 and V1) made up of Gaussian noise plus a simulated Crab-pulsar GR+s signal of the indicated scalar ($h_\mathrm{b} \equiv a_\mathrm{b}$) and tensor ($h_\mathrm{GR}$) amplitudes. The color of each hexagon represents the average value of the log-odds in that region of parameter space; color is normalized logarithmically, except for a linear stretch in the $(-1, 1)$ range. As evidenced by the extended red region in the central plot, the GR analysis is only sensitive to the tensor component of the injections and misidentifies strong scalar signals as noise; nevertheless, if the scalar component is larger than $\sim 5 \times 10^{-26}$, the GR analysis will disfavor the noise hypothesis, even for a small tensor component, as in the right panel of Fig. 6. In contrast, the any-signal analysis is sensitive to the total power of the injected signal, regardless of polarization.
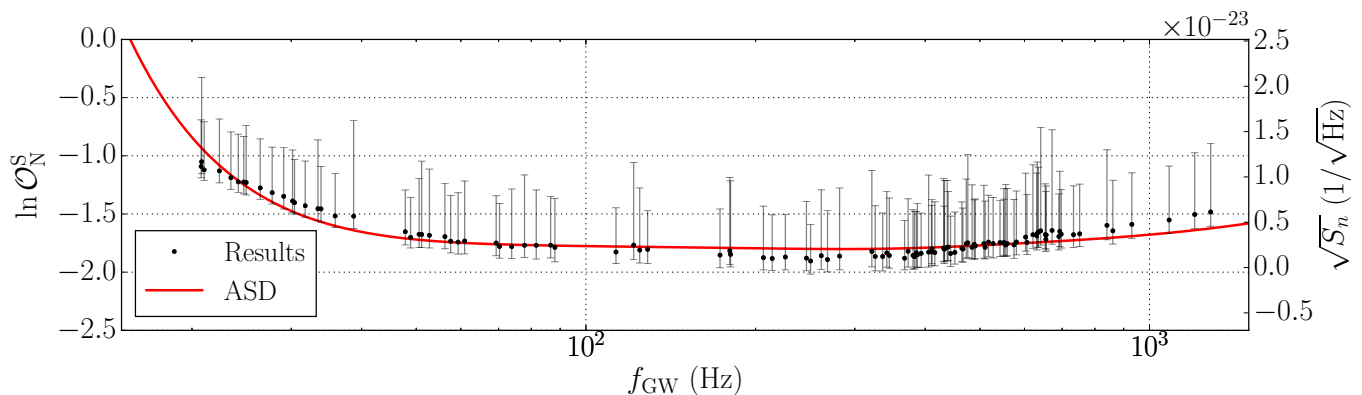
FIG. 8. *Signal log-odds vs GW frequency for noise-only data.* Circles mark the mean of the distribution of $\ln \mathcal{O}_\mathrm{N}^\mathrm{S}$, as a function of the expected GW frequency for each pulsar in our set; vertical lines indicate one-sided standard deviations for each source. Each datapoint and corresponding bars summarize the shape of a distribution like Fig. 4 for each of the pulsars, but produced from only 100 runs per source. Lower noise amplitude helps discard a greater region of signal parameter space, resulting in a lower $\ln \mathcal{O}_\mathrm{N}^\mathrm{S}$ value and explaining the observed dependence of the results on the effective noise amplitude spectral density $\sqrt{S_n}$ (ASD, red curve), corresponding to the harmonic mean of each detector PSD.

responds to the area of parameter space associated with subthreshold signals that cannot be detected.

As expected, the best coverage is obtained when analyzing the data using the model matching the injection, GR+s, (rightmost plot) or the all-signal model (leftmost plot). In both these cases, the results improve with either scalar or tensor SNR. In contrast, the GR analysis (center plot) is most sensitive to tensor strain, although the signal model is still favored for strong-enough scalar signals; this is the same behavior observed in Fig. 6.

We have produced distributions of background $\ln \mathcal{O}_\mathrm{N}^\mathrm{S}$, like those of Fig. 4, for a selection of pulsars in the sensi-

tive band of the three detectors under consideration. In Fig. 8, these are represented by their respective means and one-sided standard deviations as a function of the pulsar's GW frequency. The frequency dependence is explained by variations in the instrumental noise spectra. This is explained by the fact that, for a particular prior choice, more information is gained from the data if the noise floor is lower: with less noise it is possible to discard the presence of weaker signals, so the value of $\ln \mathcal{O}_\mathrm{N}^\mathrm{S}$ decreases.
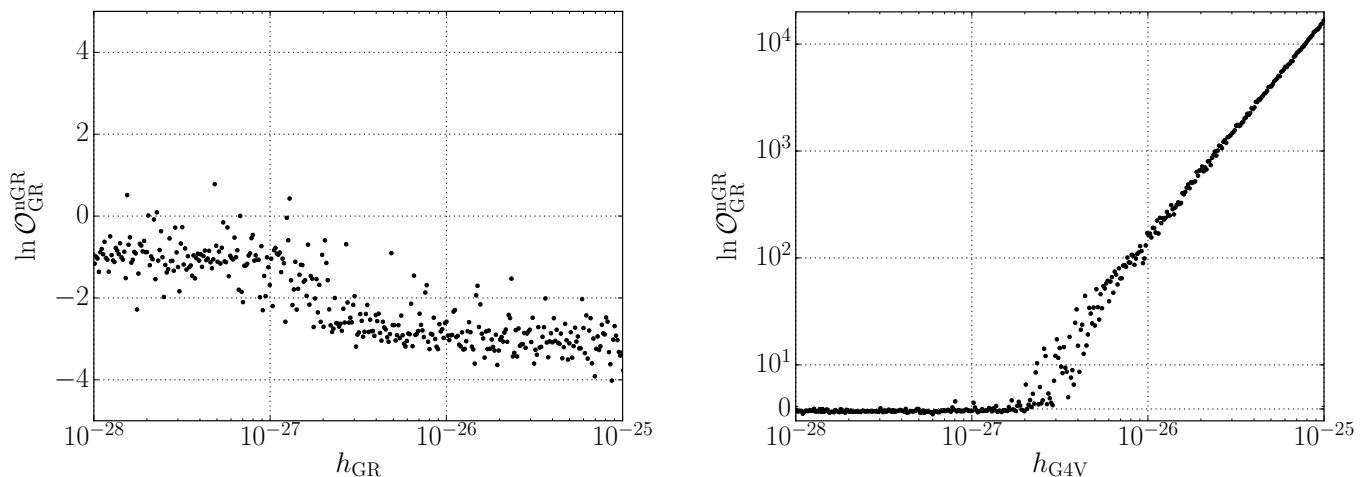
FIG. 9. *Categorizing tensor and vector injections* ($\mathcal{H}_{\mathrm{nGR}}$ vs $\mathcal{H}_{\mathrm{GR}}$). Non-GR vs GR log-odds, as a function of effective injection amplitude, for both GR (left) and a specific vector-only model, denoted G4V (right). The amplitudes are defined by $h_{\mathrm{GR}}^2 = a_+^2 + a_\times^2$ and $h_{\mathrm{G4V}}^2 = a_\mathrm{x}^2 + a_\mathrm{y}^2$. Each of the 500 points corresponds to a data instantiation (one time series for each detector: H1, L1 and V1) made up of Gaussian noise plus a simulated Crab-pulsar signal of the indicated strength. The injections were performed with random values of the nuisance phase parameters, and the data were analyzed coherently across detectors. Note that, on the right, a logarithmic scale is used for the y-axis, except for a linear stretch corresponding to the first decade. While the G4V signals are clearly identified as not conforming to GR, the conclusion is less strong for GR signals because the analysis cannot discard the presence of subthreshold non-GR components hidden by the noise.
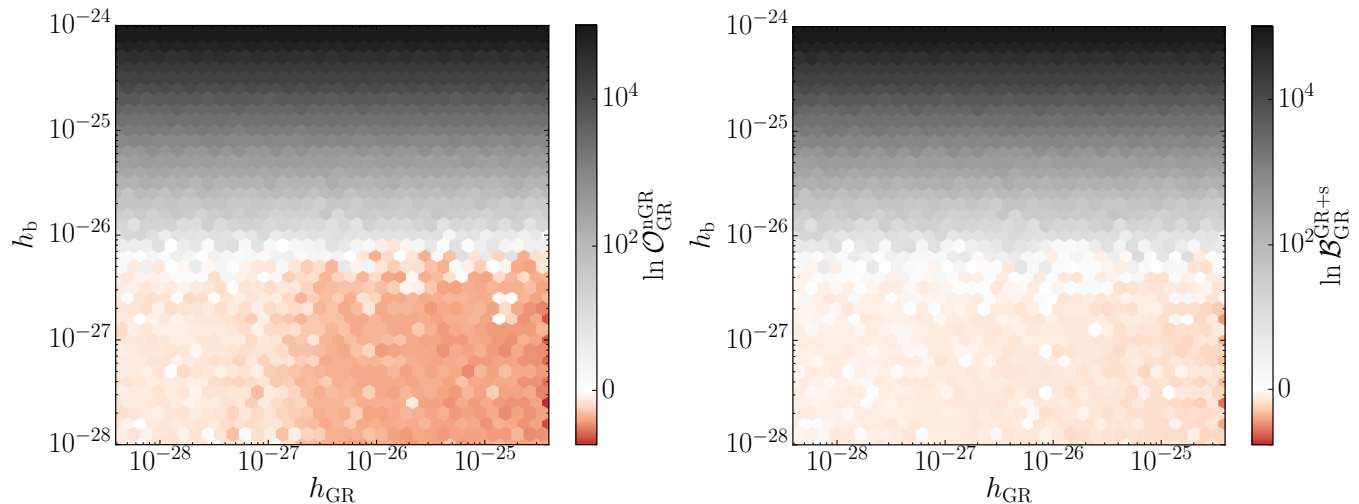
FIG. 10. *Categorizing scalar-tensor injections* ($\mathcal{H}_{\mathrm{nGR}}$ & $\mathcal{H}_{\mathrm{GR+s}}$ vs $\mathcal{H}_{\mathrm{GR}}$). Log-odds comparing the non-GR and GR+s hypotheses to GR. The non-GR odds is defined in Eq. (38). Each plot was produced by analyzing 2500 instantiations of data (one time series for each detector: H1, L1 and V1) made up of Gaussian noise plus a simulated Crab-pulsar GR+s signal of the indicated scalar ($h_\mathrm{b} \equiv a_\mathrm{b}$) and tensor ($h_{\mathrm{GR}}$) amplitudes. The color of each hexagon represents the average value of the log-odds in that region of parameter space; color is normalized logarithmically, except for a linear stretch in the $(-1, 1)$ range. Note that a horizontal slice taken over the red region of the left plot produces a series of points like those in the left panel of Fig. 9.

2. GR vs non-GR

In the presence of a signal, $\mathcal{O}_{\mathrm{GR}}^{\mathrm{nGR}}$, as defined by Eq. (38), indicates whether there is reason to believe there is a GR violation or not. Because there could always be an unresolvably small departure from GR, we do not expect

our analysis (with priors as chosen) to ever strongly favor the GR hypothesis; rather, in the presence of a GR signal we will find that $\ln \mathcal{O}_{\mathrm{GR}}^{\mathrm{nGR}}$ remains relatively close to zero, simply meaning that there is no strong evidence for or against non-GR features. This is indeed the behavior observed in the left panel of Fig. 9, where $\ln \mathcal{O}_{\mathrm{GR}}^{\mathrm{nGR}}$
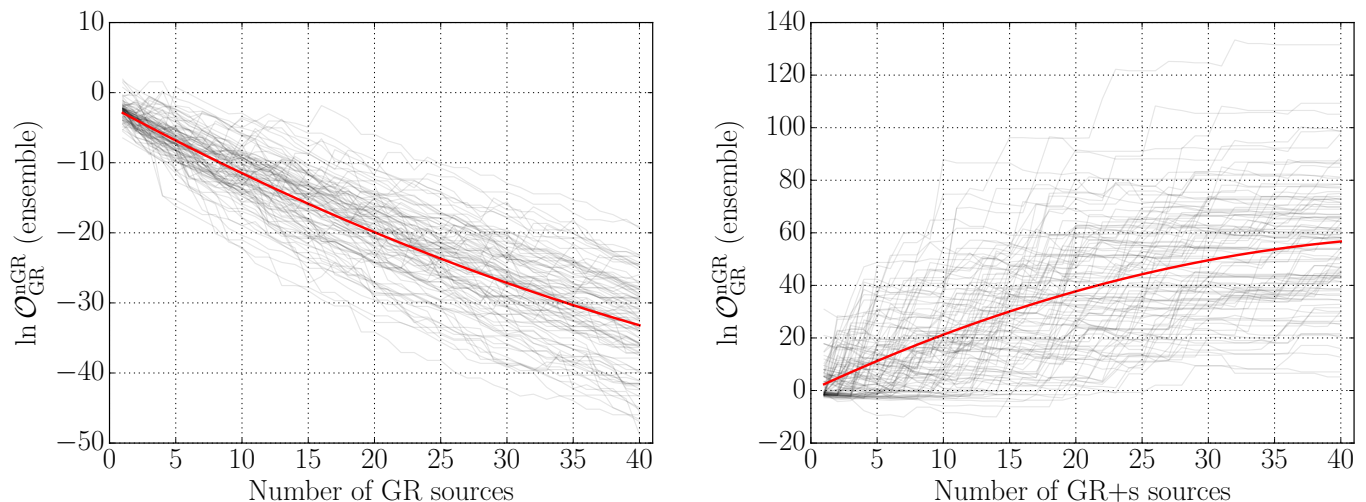
FIG. 11. *Ensemble non-GR vs GR log-odds.* Non-GR vs GR log-odds computed from data for multiple sources, vs the number of sources in the set. Each light-gray trace marks a possible progression of the ensemble log-odds as new sources are added; the red line corresponds to the best quadratic fit. For each pulsar, we chose an arbitrary data instantiation containing a GR (left) or GR+s (right); GR signals are restricted to $10^{-27} < h_{\rm GR} < 10^{-26}$, while GR+s signals also satisfy $0.3 < h_{\rm b}/h_{\rm GR} < 1$. We compute the value of $\ln \mathcal{O}_{\rm GR}^{\rm nGR}$ for each signal in the set and combine them according to Eq. (47) to obtain the ensemble value plotted in the y-axis.

is shown to be roughly insensitive to GR injection amplitude. For values of $h_{\rm GR}$ below certain threshold (which, in this case, is around 3×10^{-27}), the search does not detect a signal and, consequently, no information is gained for or against $\mathcal{H}_{\rm GR}$, i.e. $\ln \mathcal{O}_{\rm GR}^{\rm nGR} \sim 0$. The difference between the two populations (below and above threshold) is determined mainly by the choice of amplitude priors.

The behavior of $\mathcal{O}_{\rm GR}^{\rm nGR}$ is less ambiguous in the presence of a non-GR signal. For instance, if the data contain a detectable signal that completely lacks tensor components, then $\mathcal{O}_{\rm GR}^{\rm nGR}$ will unequivocally reflect this. This is evidenced by the growth of $\ln \mathcal{O}_{\rm GR}^{\rm nGR}$ with injected non-GR SNR in the right panel of Fig. 9.

As might be expected, $\mathcal{O}_{\rm GR}^{\rm nGR}$ responds to non-GR signals that include a tensor component with a combinations of features from both panels of Fig. 9. As an example, the left plot of Fig. 10 shows $\mathcal{O}_{\rm GR}^{\rm nGR}$ in the presence of GR+s injections, as a function of injected GR and scalar strengths. This plot can be split into three clearly demarcated regions: one in which the signal is not detected (light red, bottom left), one in which the signal is detected and the non–GR model is preferred (black, top), and one which the signal is detected but where the evidence for a deviation from GR is not clear due to the predominance of the tensorial component (darker red, bottom right). The first corresponds to the subthreshold population on either side of Fig. 9, while the second and third correspond to the above-threshold populations on the right and left sides of Fig. 9 respectively. For reference, Fig. 10 also includes the direct comparison of GR+s and GR on the right.

We can make a stronger statement about the agreement of the data with GR by making use of signals from multiple sources, as discussed in Sec. III A 3. The power of combining multiple signals is illustrated in Fig. 11, where $\ln \mathcal{O}_{\rm GR}^{\rm nGR}$, as defined in Eq. (47), is plotted vs number of GR (left) and GR+s (right) signals detected. Note that this presumes that, for each source, the presence of a signal has already been established from the value of $\ln \mathcal{O}_{\rm N}^{\rm S}$. Computing the ensemble $\ln \mathcal{O}_{\rm GR}^{\rm nGR}$, as done here, is a good way of summarizing the information contained in the data about the relative likelihoods between the two models, but it provides no information not already present in the set of single-source odds.

B. Parameter estimation

When no conclusive evidence for a CW is found in the data, we are still interested in placing upper limits on the strength of possible signals (up to some credibility), and this is done as explained in Sec. III B. By the same token, if a signal consistent with GR is detected, we can always place an upper limit on the amplitude of non-GR modes, even if the odds indicate there is no clear sign of a GR violation.

For instance, we can get a quantitative estimate of our sensitivity to scalar modes from a given source by looking at the distribution of $h_{\rm b}^{95\%}$, defined in Eq. (55), computed for a set of noise-only data instantiations. Such distribution for the Crab pulsar is presented in the left panel of Fig. 12. Similarly, the right panel presents estimates for the sensitivity to vector modes coming from the Crab pulsar, assuming a vector-tensor model. In this case, however, the quantity plotted is the total, effective
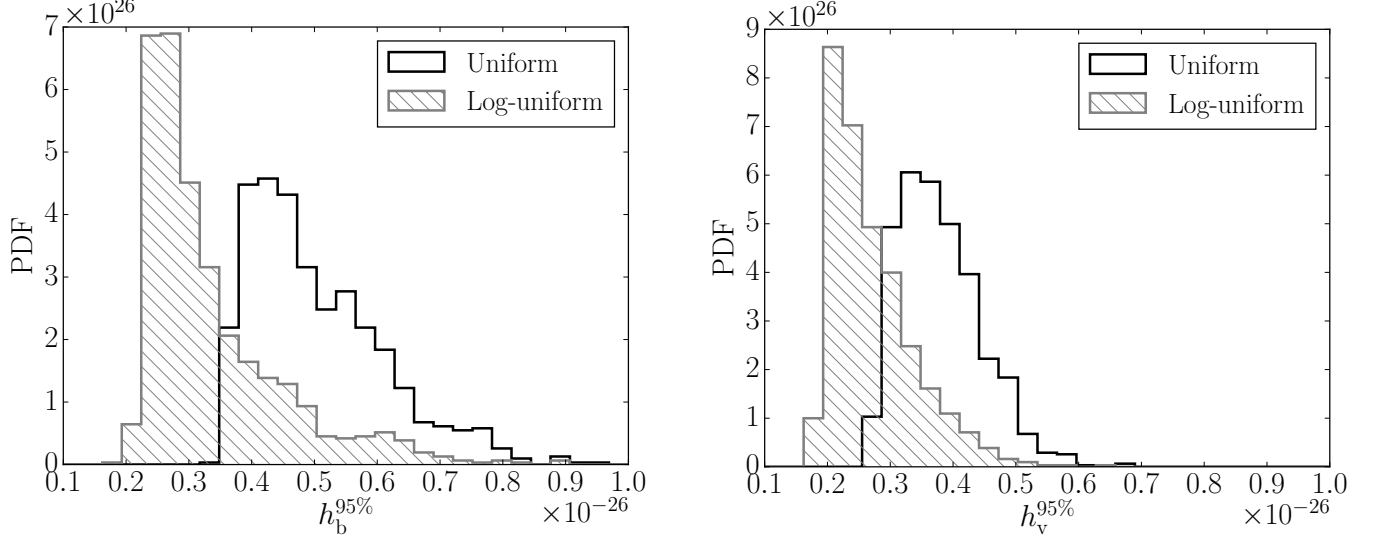
FIG. 12. *Expected Crab non-GR upper-limits in absence of signal.* Histogram of 95%-credible upper limits for the scalar (left) and vector (right) amplitudes, for a set of 1000 noise-only data sets, computed using priors uniform in the amplitude (black) or uniform in the logarithm of the amplitude (hatched gray); the differences between these two priors are discussed in detail in Appendix B. The effective vector amplitude h_v is defined by $h_\mathrm{v}^2 = h_\mathrm{x}^2 + h_\mathrm{y}^2$. Each instantiation (one time series for each detector: H1, L1 and V1) is made up of simulated Gaussian noise with standard deviation given by the advanced design PSDs. The mean of these PDFs scales with PSD, as $\sqrt{S_\mathrm{n}(f)/T}$, where T is the observation time (assuming stationary noise). Scalar and vector upper-limits are produced using the GR+s and GR+v models respectively.
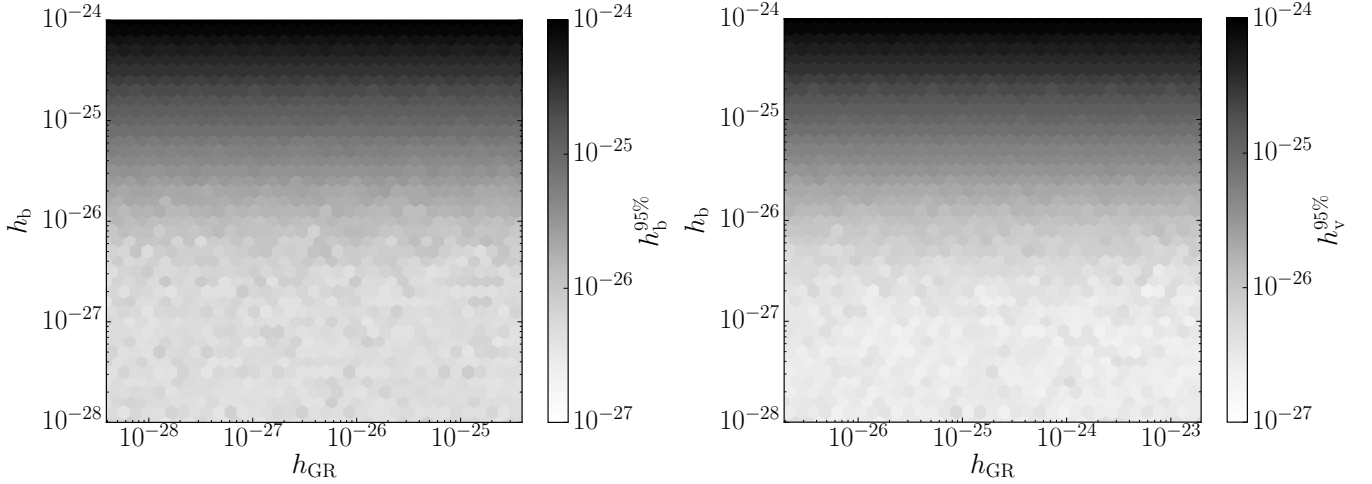
FIG. 13. *Expected Crab scalar and vector upper-limits in presence of GR+s and GR+v signals.* Shading represents the 95%-credible upper limit for the scalar ($h_\mathrm{b}^{95\%}$, left) and vector ($h_\mathrm{v}^{95\%}$, right) amplitudes, vs the amplitude of injected GR (h_GR, x-axis) and corresponding non-GR (h_b on the left, and h_v on the right; y-axis) components. Each plot was produced by analyzing 2500 instantiations of data (one time series for each detector: H1, L1 and V1) made up of Gaussian noise plus a simulated Crab-pulsar GR+s (left) or GR+v (right) signal with indicated strains. Note that the injection amplitudes are defined by $h_\mathrm{GR}^2 \equiv a_+^2 + a_\times^2$, $h_\mathrm{b} \equiv a_\mathrm{b}$ and $h_\mathrm{v}^2 \equiv a_\mathrm{x}^2 + a_\mathrm{y}^2$. The color of each hexagon represents the average value of the upper-limit in that region of parameter space. We underscore that the upper limits, $h_\mathrm{b}^{95\%}$ and $h_\mathrm{v}^{95\%}$, are well-defined even when the non-GR component is strong enough to be detected, as is the case for the darker-colored regions. Note that $h_\mathrm{b}^{95\%}$ and $h_\mathrm{v}^{95\%}$ are insensitive to h_GR.
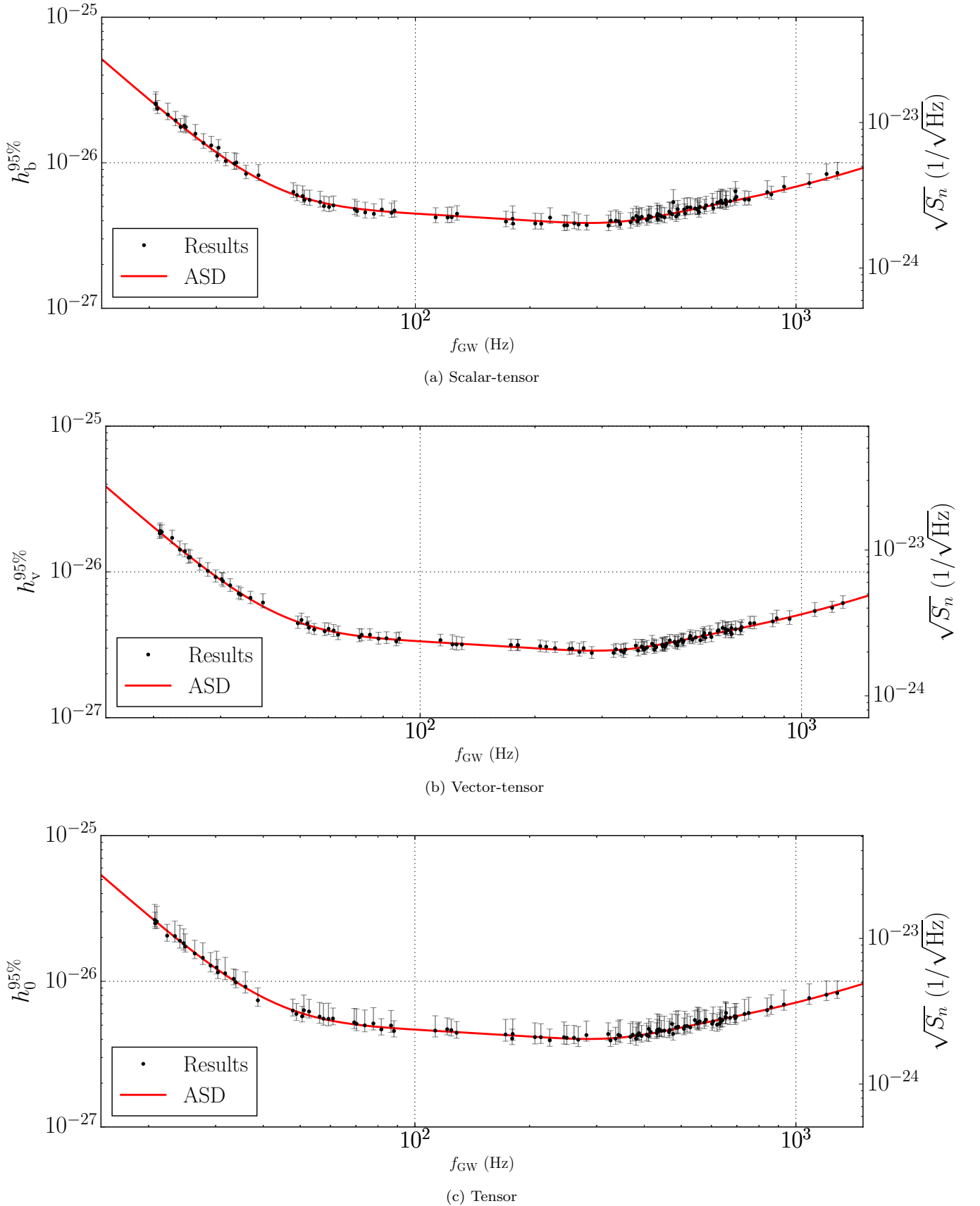
(a) Scalar-tensor

(b) Vector-tensor

(c) Tensor

FIG. 14. *Expected upper-limits in absence of signal vs GW frequency.* Circles mark the mean of the distribution of $h_{\rm b}^{95\%}$ (top), $h_{\rm v}^{95\%}$ (middle) and $h_{0}^{95\%}$ (bottom), as a function of expected GW frequency for each pulsar in our set; vertical lines mark one-sided standard deviations for each source. Each datapoint and corresponding bars summarize the shape of a distribution like those of Fig. 12, but produced from 100 noise instantiations each. The scalar and vector upper limits were produced assuming GR+s and GR+v models respectively, while h_0 is the GR-only amplitude. We use unifom priors in all amplitude parameters (see Fig. 12 and Appendix B).

vector strain amplitude $h_{\mathrm{v}} = \sqrt{a_{\mathrm{x}}^2 + a_{\mathrm{y}}^2}$. These plots include distributions produced using the same log-uniform prior used to obtain Bayes factors, as well as more conservative ones obtained using uniform amplitude priors (see Appendix B). In either case, the magnitude of $h_{\mathrm{v}}^{95\%}$ is comparable to that of $h_{\mathrm{b}}^{95\%}$.

Interestingly, our ability to measure scalar and vector amplitudes is unaffected by the presence of other modes. We illustrate this for the Crab pulsar in Fig. 13, which results from analyzing data with GR+s (left) and GR+v (right) injections. There we plot $h_{\mathrm{b}}^{95\%}$ as a function of scalar and tensor injection amplitudes on the left, and $h_{\mathrm{v}}^{95\%}$ as a function of vector and tensor injection amplitudes on the right. From these plots, one can conclude that $h_{\mathrm{b}}^{95\%}$ and $h_{\mathrm{v}}^{95\%}$ are sensitive only to the corresponding scalar and vector components, and not by h_{GR}.

As shown previously in the literature, the mean of distributions like those of Fig. 12 will scale with $\sqrt{S_{\mathrm{n}}(f)/T}$, where $S_{\mathrm{n}}(f)$ is the effective PSD of the detector noise at the expected GW frequency f and T is the integration time (cf. Eq. (26) and Fig. 1 of [19]). Because of this, the mean of this distribution will vary with the source's expected GW frequency, as shown in Fig. 14. Note that, following convention, these upper limits are computed using uniform amplitude priors, which means they are a factor of a few less stringent than those obtained with a log-uniform prior (see Fig. 12 and Appendix B). Also, for completeness, Fig. 14 also includes the expected GR upper limits, $h_0^{95\%}$; these values are the same as would be obtained by the standard GR-only search.

VI. CONCLUSION

We have developed a Bayesian framework to detect CW signals from known sources regardless of polarization content, to disentangle the modes present in a given signal, and to constrain the amplitudes of extra polarizations that may be hiding under the noise. We have implemented this as an extension of LIGO's Bayesian targeted CW search pipeline [26], and thus benefit from the power of the nested sampling algorithm on which it is based.

We have tested our methods on one year of simulated noise for three advanced-era detectors at design sensitivity (H1, L1, V1), and prepared for a set of multiple known sources in their frequency band. This allows us to estimate our future sensitivity to CW polarizations, in this most optimistic case. Under these conditions and for the Crab pulsar in particular, we expect signals of any polarization to become detectable for characteristic strain amplitudes $h \gtrsim 3 \times 10^{-27}$ (Figs. 6 and 7); this threshold will vary among sources, due to differences in position (sky location and orientation) and detector PSD at the expected GW frequency (cf. e.g. Fig. 8). Furthermore, the value of this threshold will decrease linearly with the square-root of the observation time [19].

A signal louder than the detection threshold will allow us to determine whether its polarization content is consistent with GR or not, and the strength of this statement will depend almost exclusively on the power of the non-GR component (Figs. 9 and 10). In other words, from a model-selection standpoint, the non-GR hypothesis will only be unequivocally favored if the total power in non-GR modes is greater than the threshold value, regardless of the strength of the GR modes. However, for signals that do not satisfy this, we may always place upper limits on non-tensorial amplitudes and thus constrain deviations from GR; for instance, Fig. 14 presents the most optimistic expectations for 95%-credible upper limits for scalar and vector amplitudes of CW signals from all pulsars in our set ($h_{\mathrm{b}}^{95\%} \sim h_{\mathrm{v}}^{95\%} \sim 3 \times 10^{-27}$ in the best case). As far as we are aware, these are the first generic estimates of sensitivity to scalar and vector polarizations ever published [34].

Although the results presented here made use of simulated Gaussian noise, the procedure is identical for actual detector data. Furthermore, the assumption of Gaussianity has been shown to hold relatively well for real CW data [8], so the actual sensitivity limits should not be far from those presented here. If the data are strongly non-Gaussian, however, one must be careful in using $\ln \mathcal{O}_{\mathrm{N}}^{\mathrm{S}}$ for detection purposes and may instead wish to adopt one of the strategies suggested in Sec. III A 4.

Another important limitation of our results is that here we only consider CW signals emitted at $f = 2f_{\mathrm{rot}}$, while it is to be expected that other mechanisms (within GR or not) allow emission at other harmonics, $f = f_{\mathrm{rot}}$ in particular. Yet, the only change required to account for this is to modify the template in Eq. (7) to include terms at different harmonics; the ability to do this already exists within our current infrastructure. We also assume that other aspects of the waves, like their speed, remain in agreement with the GR prediction, an assumption that will be relaxed in a future study.
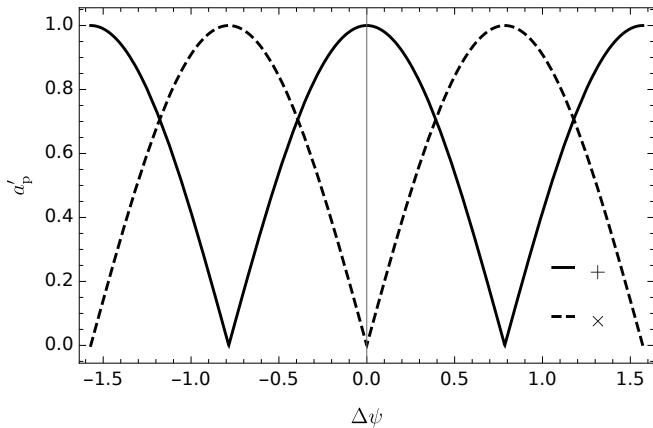
FIG. 15. *Effect of changing polarization angle.* Norm of the complex plus (a'_+, solid line) and cross ($a'_\times$, dashed line) weights after rotating the source by $\Delta\psi$ in the plane of the sky, i.e. letting $\psi \rightarrow \psi' = \psi + \Delta\psi$; this transformation is expressed in Eqs. (A5) and (A6). In this case, we start from $a_+ = 1$, $a_\times = 0$ and $\psi = 0$.

Appendix A: Tensor modes

A distinction can be made between GR and a free-tensor model that includes $+$ and $\times$ but does not restrict their relative amplitudes (denoted "t"). The former has four free parameters (overall amplitude, h_0; overall phase, ϕ_0; inclination, ι; polarization, ψ) and corresponds to a signal template of the form:

$$\Lambda_{\rm GR}(t) = \frac{h_0}{2}\left[\frac{1}{2}(1+\cos^2\iota)F_+(t;\psi) - i\cos\iota F_\times(t;\psi)\right]e^{i\phi_0}.$$
(A1)

This is a special case of the free-tensor model, which has five parameters (plus amplitude, a_+; cross amplitude, $a_\times$; plus phase, ϕ_+; cross phase, $\phi_\times$; polarization, ψ) and whose template is:

$$\Lambda_{\rm t}(t) = \frac{1}{2}\left[a_+ e^{i\phi_+}F_+(t;\psi) + a_\times e^{i\phi_\times}F_\times(t;\psi)\right].\quad\text{(A2)}$$

If ψ and ι are known, it is clear that the two models are different, since $\mathcal{H}_{\rm GR}$ has two free parameters (h_0, ϕ_0) and $\mathcal{H}_{\rm t}$ has four (a_+, $a_\times$, ϕ_+, $\phi_\times$). If the orientation is *not* fixed, however, the two models span the same signal space. This is because there is a degeneracy between ψ and a_+, $a_\times$ due to the way the antenna patterns depend on the source orientation:

$$F_+(t;\psi') = F_+(t;\psi)\cos 2\Delta\psi + F_\times(t;\psi)\sin 2\Delta\psi,\quad\text{(A3)}$$

$$F_\times(t;\psi') = F_\times(t;\psi)\cos 2\Delta\psi - F_+(t;\psi)\sin 2\Delta\psi,\quad\text{(A4)}$$

with $\psi' = \psi + \Delta\psi$. Eqs. (A3) and (A4) can be derived from Eqs. (1) and (2) respectively, as in [13]. Consequently, changing $\psi \rightarrow \psi'$ in Eq. (A2) is equivalent to

leaving ψ fixed while replacing the plus and cross complex amplitudes by:

$$a'_+ e^{i\phi'_+} = a_+ e^{i\phi_+}\cos 2\Delta\psi - a_\times e^{i\phi_\times}\sin 2\Delta\psi,\quad\text{(A5)}$$

$$a'_\times e^{i\phi'_\times} = a_\times e^{i\phi_\times}\cos 2\Delta\psi + a_+ e^{i\phi_+}\sin 2\Delta\psi.\quad\text{(A6)}$$

This is illustrated in Fig. 15.

The relationship between the different tensor model parameters can be appreciated in the posterior probability plots of Fig. 16. The degeneracy between ψ and a_+, $a_\times$ is particularly evident in Fig. 16d, where the one-dimensional PDF for ψ shows that this parameter cannot be constrained, even for a loud signal. Furthermore, joint posteriors between ψ and a_+ & $a_\times$ confirm that this is due to the degeneracy from Eqs. (A3) and (A4), as seen by comparing these two-dimensional PDFs to Fig. 15. Less importantly, these relations also cause ψ and $\cos\iota$ to become correlated (see Fig. 17).

Because their signal templates are degenerate when ψ is allowed to vary, the distinction between $\mathcal{H}_{\rm GR}$ and $\mathcal{H}_{\rm t}$ becomes less meaningful. This can be seen explicitly by comparing the values of $\ln\mathcal{B}_{\rm GR}^{\rm t}$ in the cases of known and unknown orientations, as in Fig. 18. As shown in Fig. 19, the ambiguity between $\mathcal{H}_{\rm GR}$ and $\mathcal{H}_{\rm t}$ propagates into $\mathcal{H}_{\rm nGR}$, making it no longer appropriate to treat $\mathcal{H}_{\rm t}$ as a non-GR subhypothesis. Therefore, for sources of unknown orientation, we redefine $\mathcal{H}_{\rm nGR}$ to be composed of the models in $\tilde{M}' = \tilde{M}\setminus\{{\rm t,st,tv,stv}\}$, as explained in the main text, i.e.

$$\tilde{M}' = \{{\rm GR+s, GR+v, GR+sv, s, v, sv}\},\quad\text{(A7)}$$

$$M' = \{{\rm GR, GR+s, GR+v, GR+sv, s, v, sv}\}.\quad\text{(A8)}$$

Appendix B: Amplitude priors

Previous CW Bayesian searches targeted to known pulsars have always applied a flat prior on the signal amplitude parameter [23]. This is because flat priors, if wide enough, cause the posterior to be only determined by the likelihood (up to normalization), yielding more conservative upper limits on the signal strength. Furthermore, unlike with priors uniform in the logarithm of the quantity, upper limits derived with flat priors will generally not depend on the limits set by the prior (again, assuming the range allowed extends from zero amplitude to some large value that does not truncate the likelihood).

Upper limits obtained using log-uniform priors (uniform in the logarithm of the quantity) will, generally, be dependent on the range of the prior, although not strongly. For example, consider a one-dimensional problem on some positive parameter x. For simplicity, further assume we have a flat likelihood between $x = 0$ and an upper cutoff at $x = x_{\rm max}$; then, $x_{\rm max}$ will necessarily also be an upper bound for the posterior. Because the likelihood is uniform, below the cutoff the posterior will be
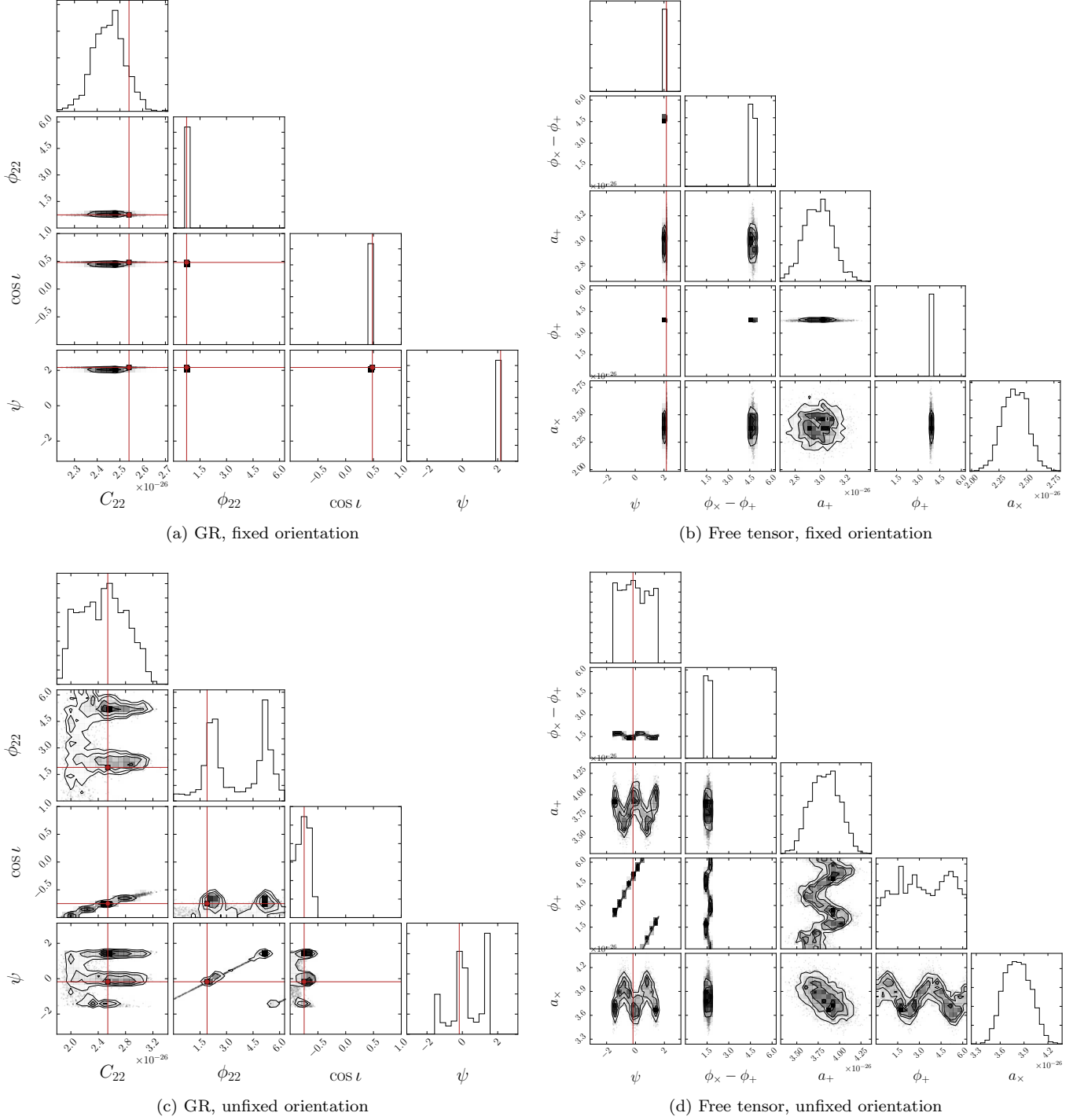
(a) GR, fixed orientation

(b) Free tensor, fixed orientation

(c) GR, unfixed orientation

(d) Free tensor, unfixed orientation

FIG. 16. *Tensor posteriors in presence of signal.* Posterior PDFs for parameters of $\mathcal{H}_{\mathrm{GR}}$ (left) and $\mathcal{H}_{\mathrm{t}}$ (right) with fixed (top) and unfixed (bottom) source orientation (ψ, ι). The data analyzed contain signals with parameters indicated by the red lines; note that $C_{22} = h_0/2$ is the quantity actually used to parametrize GR amplitudes in our nested sampling code [20]. For fixed orientation, both the GR (a) and free-tensor (b) analyses accurately determine the amplitude and phase of the injected signal; for the latter, a_+ and $a_\times$ are constrained to lie within a region consistent with $h_{\mathrm{inj}}^2 = a_+^2 + a_\times^2$ and $a_+/a_\times = (1 + \cos^2 \iota)/(2 \cos \iota)$, for an effective injection amplitude given by $h_{\mathrm{inj}}^2 = h_0^2 (1 + \cos^2 \iota)^2/4 + h_0^2 \cos^2 \iota$. [In both (a) and (b), $\cos \iota$ and ψ are known, and their resolution in these plots is limited by binning only.] When the orientation is allowed to vary, we observe correlations between the recovered GR amplitude and the orientation parameters in panel (c), explained by Eqs. (9–11) and (A3, A4); the measurements of $\cos \iota$ and ψ will also be correlated, as can be seen by noting that ψ is not defined for a source with spin axis along the line of sight (see Fig. 17). As discussed in the text, a free-tensor analysis with unfixed orientation is unable to constrain ψ due to the correlations with the amplitude parameters shown in the first column of panel (d), in agreement with Fig. 15; this degeneracy makes the distinction between $\mathcal{H}_{\mathrm{GR}}$ and $\mathcal{H}_{\mathrm{t}}$ less meaningful (see Fig. 18).
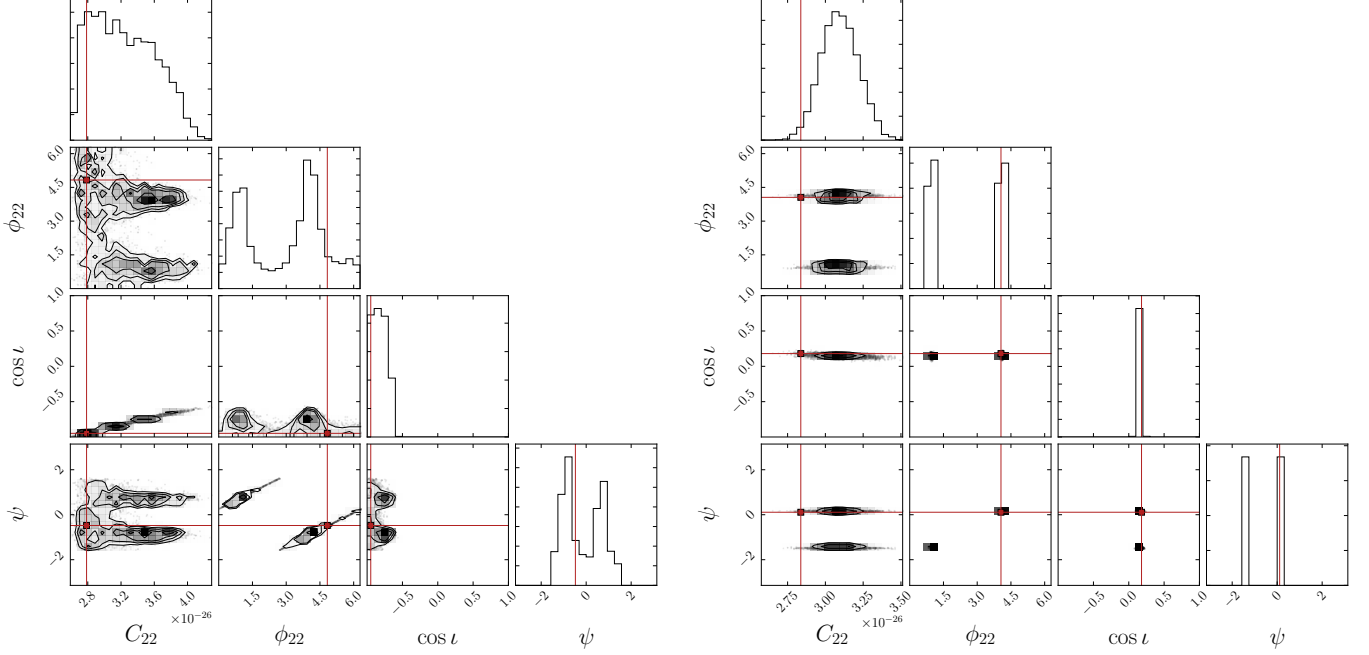
FIG. 17. *Effect of inclination.* Posterior PDFs for parameters of $\mathcal{H}_{\mathrm{GR}}$ with unfixed source orientation (ψ, ι). The data sets analyzed contain signals with parameters indicated by the red lines; note that $C_{22} = h_0/2$ is the quantity actually used to parametrize GR amplitudes in our nested sampling code [20]. On the left, the injected signal corresponds to a face-off source $(\cos \iota \approx -1)$, making it difficult to constrain the polarization angle ψ; on the right, the injection has similar amplitude but corresponds to an edge-on source $(\cos \iota \approx 0)$, making it easy to constrain ψ [modulo $\pi/2$ due to the $2\Delta\psi$ dependence of Eqs. (A3) and (A4)].

(a) Fixed orientation

(b) Unfixed orientation
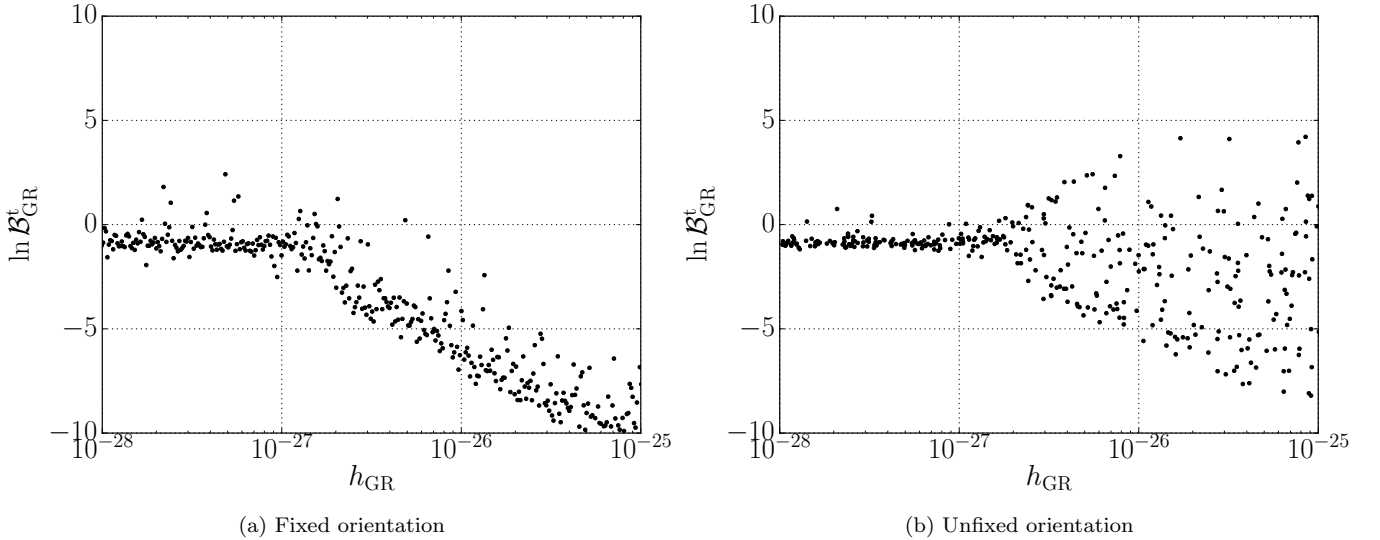
FIG. 18. *Free-tensor vs GR.* Natural logarithm of the Bayes factor comparing $\mathcal{H}_{\mathrm{t}}$ to $\mathcal{H}_{\mathrm{GR}}$, as a function of GR injection amplitude for fixed (left) and unfixed (right) source orientation. On the left, the analysis correctly gives preference to $\mathcal{H}_{\mathrm{GR}}$ for signals above the detection threshold; on the right, however, the analysis is unable to satisfactorily distinguish between $\mathcal{H}_{\mathrm{t}}$ and $\mathcal{H}_{\mathrm{GR}}$, due to the orientation degeneracies discussed in appendix A.
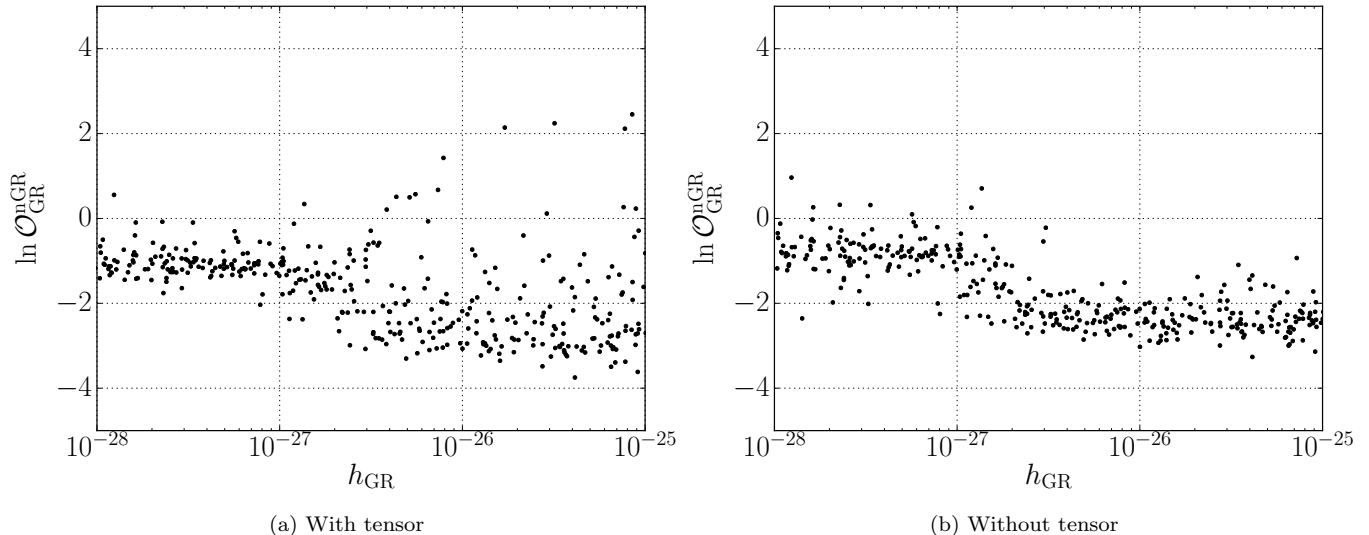
(a) With tensor

(b) Without tensor

FIG. 19. *Non-GR vs GR for unfixed orientation.* Natural logarithm of the Bayes factor comparing $\mathcal{H}_{\mathrm{nGR}}$ to $\mathcal{H}_{\mathrm{GR}}$ vs GR injection amplitude for unfixed source orientation. On the left, $\mathcal{H}_{\mathrm{nGR}}$ is constructed from the models in $\tilde{M}$, as defined by Eq. (27); on the right, $\mathcal{H}_{\mathrm{nGR}}$ is constructed from the models in $\tilde{M}' = \tilde{M} \setminus \{\mathrm{t}, \mathrm{st}, \mathrm{tv}, \mathrm{stv}\}$, to account for the ambiguity between $\mathcal{H}_{\mathrm{GR}}$ and $\mathcal{H}_{\mathrm{t}}$ for sources of unknown ψ. Note the improvement after removing the free-tensor models from the non-GR hypothesis.

determined, up to normalization, by the prior only, i.e. for $x < x_{\max}$,

$$p(x \mid \mathbf{B}, \mathcal{H}) \propto p(x \mid \mathcal{H}). \tag{B1}$$

Now consider a log-uniform prior $p(x \mid \mathcal{H}) \propto \mathrm{d}(\log x) \propto 1/x$, with a lower bound $x_{\min}$, such that $0 < x_{\min} < x_{\max}$. Because such prior is uniform in the $\log x$, this implies that the 95%-credible upper limit on x will be given by:

$$\log x^{95\%} = \log x_{\min} + 0.95(\log x_{\max} - \log x_{\min})$$
$$= \log\left(x_{\max}^{0.95}/x_{\min}^{0.95-1}\right). \tag{B2}$$

Since $x_{\max}$ is set by the likelihood, if the prior is changed by rescaling $x_{\min}$ by a factor α,

$$x_{\min} \to x'_{\min} = \alpha x_{\min}, \tag{B3}$$

then, for a given set of data, the upper limit becomes $x_\alpha^{95\%}$, satisfying:

$$x_\alpha^{95\%}/x^{95\%} = \alpha^{0.05}. \tag{B4}$$

Thus, the dependence of the upper limit on the range defined by the log-uniform prior is quite weak, as illustrated in Fig. 20. This explains why upper limits obtained with a log-uniform prior differ only by a factor of a few from those obtained with a flat one, as seen in Fig. 12.

However, the flat priors do not properly represent our ignorance of the scale of the signal amplitude. This problem manifests itself in negative Bayes factors that too quickly favor the noise hypothesis if no loud signal is clearly present, rather than reflecting our expectation

that a signal might be hiding under the noise. This can be seen in Fig. 21, where we show the distributions of $\ln \mathcal{B}_{\mathrm{N}}^{\mathrm{GR}}$, obtained for several noise-only data instantiations for the Crab pulsar, corresponding to flat and log-uniform priors in the GR amplitude parameter, h_0; a uniform prior results in lower values of $\ln \mathcal{B}_{\mathrm{N}}^{\mathrm{GR}}$ that strongly favor $\mathcal{H}_{\mathrm{N}}$. This behavior is not specific to the GR model.

For most of our analysis, we choose to apply priors uniform in the logarithm of all amplitude quantities. However, for the sake of consistency with previous searches and in order to make our limits more conservative, we also present upper limits produced using flat amplitude priors, as shown in Fig. 12.

Appendix C: Numerical error

The fractional numerical error in the computation of the natural logarithm of the evidence by nested sampling is usually estimated by:

$$\delta\left[\ln P\left(\mathbf{B} \mid \mathcal{H}\right)\right] \sim \sqrt{H/N_{\mathrm{live}}}, \tag{C1}$$

where N_{live} is the number of of live points and H is the information gained in the analysis:

$$H \equiv \int_\Theta p(\vec{\theta} \mid \mathbf{B}, \mathcal{H}) \, \ln \frac{p(\vec{\theta} \mid \mathbf{B}, \mathcal{H})}{p(\vec{\theta} \mid \mathcal{H})} \mathrm{d}\vec{\theta}, \tag{C2}$$

a quantity that is easy to estimate from the output of the nested sampling code [27, 35].

An example of the actual statistical error as function of SNR is presented in Figs. 22 & 23, where the injected GR signal amplitude serves as proxy for ρ (for fixed PSD).
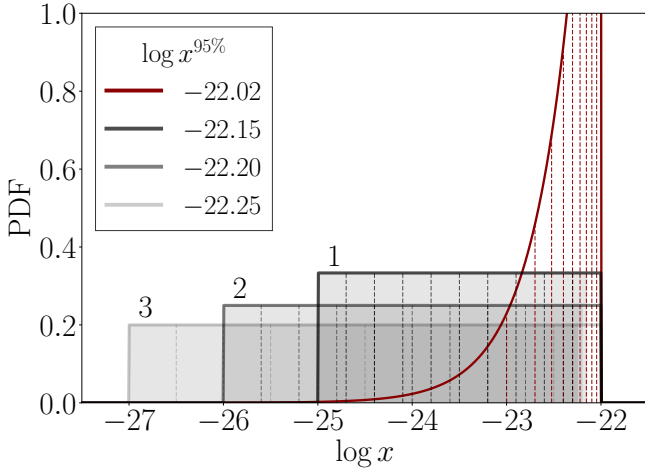
FIG. 20. *Log-uniform prior and upper limits.* For a 1D random variable x, we show the probability densities corresponding to a uniform likelihood with upper cutoff $\log x_{max} = -22$ (red) and log-uniform priors with different lower cutoffs ($\log x_{min} = -25$ for box 1, $\log x_{min} = -26$ for box 2 and $\log x_{min} = -27$ for box 3). Vertical dashed lines mark areas of equal probability mass for each distribution. The combined effect of the likelihood and each of the prior distributions is to produce 95%-credible upper limits on x with values shown in the legend. The value obtained using only the likelihood corresponds to that obtained with a uniform prior with a broad enough range. As expected from Eq. (B4), the upper limit is not very sensitive to the lower bound set by the prior.
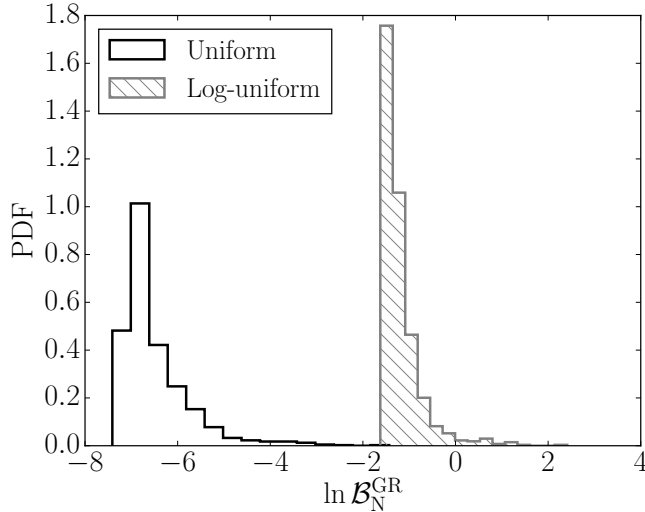

FIG. 21. *Log-uniform vs flat amplitude priors.* The logarithm of the GR vs noise Bayes factor is computed for 1000 instantiations of Crab pulsar noise. For the GR amplitude h_0, we apply priors uniform in the quantity (black) and uniform in the logarithm of the quantity (hatched gray). The flat prior causes one to more strongly favor the noise model, due to a larger implicit Ockham's penalty.
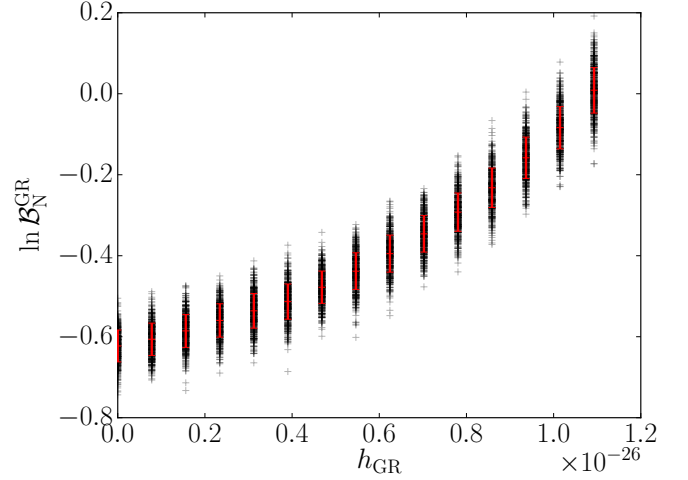

FIG. 22. *Numerical error in Bayes factor computation.* The logarithm of the GR vs noise Bayes factor is computed 500 times for different values of injected GR signal amplitude. The noise realization is not varied between computations with the same injection strength, only the seed for the random number generator used by the nested sampling algorithm. The red bars mark one standard deviation around the mean.
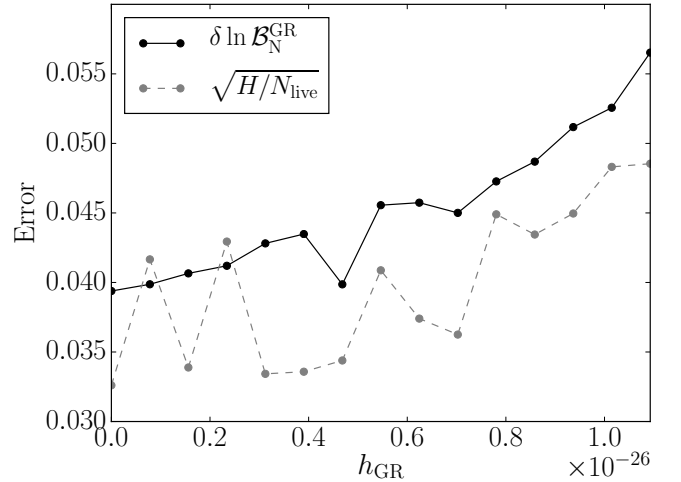

FIG. 23. *Observed error and prediction.* Error in the computation of the logarithm of the GR vs noise Bayes factor as a function of injected GR signal amplitude. The solid, black curve corresponds to measured standard deviations from the computation of $\ln \mathcal{B}_N^{GR}$ 500 times per injection strength (red bars in Fig. 22). The dashed, gray curve shows the theoretical prediction for the error in the logarithm of the evidence, Eq. (C1).

From these plots it becomes apparent that, although the actual error might exceed the estimator of Eq. (C1), its absolute magnitude is quite small and should not affect our results. In any case, Eq. (C1) indicates that any level of accuracy may be achieved by increasing the number of live points (at the cost of increased computational

burden). For more details on the numerical error of the nested sampling algorithm in `LALInference`, we refer the reader to Sec. IVB of [36].

[1] B. P. Abbott et al., (The LIGO Scientific Collaboration, and The Virgo Collaboration), Phys. Rev. Lett. **116**, 061102 (2016).

[2] B. P. Abbott et al., (The LIGO Scientific Collaboration, and The Virgo Collaboration), Phys. Rev. Lett. **116**, 241103 (2016).

[3] B. P. Abbott et al., (The LIGO Scientific Collaboration, and The Virgo Collaboration), Phys. Rev. Lett. **116**, 221101 (2016).

[4] D. M. Eardley, D. L. Lee, A. P. Lightman, R. V. Wagoner, and C. M. Will, Phys. Rev. Lett. **30**, 884 (1973).

[5] P. C. C. Freire, N. Wex, G. Esposito-Farèse, J. P. W. Verbiest, M. Bailes, B. A. Jacoby, M. Kramer, I. H. Stairs, J. Antoniadis, and G. H. Janssen, Mon. Not. R. Astron. Soc. **423**, 3328 (2012).

[6] J. Aasi et al., (The LIGO Scientific Collaboration, and The Virgo Collaboration), Astrophys. J. **785**, 119 (2014), arXiv:1309.4027.

[7] M. Zimmermann and E. Szedenits, Phys. Rev. D **20**, 351 (1979).

[8] M. Isi, A. J. Weinstein, C. Mead, and M. Pitkin, Phys. Rev. D **91**, 082002 (2015).

[9] C. M. Will, Living Rev. Relativ. **17** (2014), 10.12942/lrr-2014-4.

[10] K. Chatziioannou, N. Yunes, and N. Cornish, Phys. Rev. D **86**, 022004 (2012).

[11] C. Mead, "Gravitatational Waves in G4v," (2015), arXiv:1503.04866.

[12] A. Nishizawa, A. Taruya, K. Hayama, S. Kawamura, and M.-a. Sakagami, Phys. Rev. D **79**, 082002 (2009).

[13] A. Błaut, Phys. Rev. D **85**, 043005 (2012).

[14] D. I. Jones and N. Andersson, Mon. Not. R. Astron. Soc. **331**, 203 (2002).

[15] B. J. Owen, L. Lindblom, C. Cutler, B. F. Schutz, A. Vecchio, and N. Andersson, Phys. Rev. D **58**, 084020 (1998), arXiv:9804044 [gr-qc].

[16] R. Bondarescu, S. A. Teukolsky, and I. Wasserman, Phys. Rev. D **79**, 104003 (2009), arXiv:0809.3448.

[17] R. T. Edwards, G. B. Hobbs, and R. N. Manchester, Mon. Not. R. Astron. Soc. **372**, 1549 (2006), arXiv:0607664 [astro-ph].

[18] T. M. Niebauer, A. Rüdiger, R. Schilling, L. Schnupp, W. Winkler, and K. Danzmann, Phys. Rev. D **47**, 3106 (1993).

[19] R. Dupuis and G. Woan, Phys. Rev. D **72**, 102002 (2005), arXiv:0508096 [gr-qc].

[20] M. Pitkin, J. Veitch, and G. Woan, "A nested sampling code for targeted searches for continuous gravitational waves from pulsars," (in prep.).

[21] T. G. F. 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).

[22] X. Fan, Y. Chen, and C. Messenger, Phys. Rev. D **94**, 084029 (2016).

[23] B. P. Abbott et al., (The LIGO Scientific Collaboration, and The Virgo Collaboration), (2017), arXiv:1701.07709.

[24] D. Keitel, R. Prix, M. A. Papa, P. Leaci, and M. Siddiqi,

Phys. Rev. D **89**, 064023 (2014), arXiv:1311.5738.

[25] M. Pitkin, C. Gill, J. Veitch, E. Macdonald, and G. Woan, J. Phys. Conf. Ser. **363**, 012041 (2012), arXiv:1203.2856.

[26] M. Pitkin, J. Veitch, and G. Woan, "A nested sampling code for targeted searches for continuous gravitational waves from pulsars," (in preparation).

[27] J. Skilling, Bayesian Anal. **1**, 833 (2006).

[28] 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), arXiv:1409.7215.

[29] https://wiki.ligo.org/DASWG/LALSuite.

[30] B. P. Abbott, (The LIGO Scientific Collaboration, and The Virgo Collaboration), Phys. Rev. Lett. **116**, 241102 (2016), arXiv:1602.03840 [gr-qc].

[31] M. Pitkin, C. Gill, D. I. Jones, G. Woan, and G. S. Davies, Mon. Not. R. Astron. Soc. **453**, 4400 (2015), arXiv:1508.00416.

[32] E. Jaynes, IEEE Trans. Syst. Sci. Cybern. **4**, 227 (1968).

[33] D. Sivia and J. Skilling, *Data Analysis: A Bayesian Tutorial*, 2nd ed. (Oxford University Press, New York, 2006) p. 264.

[34] Note that sensitivity estimates presented in [8] were restricted to the specific vector-only model of [11].

[35] C. R. Keeton, Mon. Not. R. Astron. Soc. **414**, 1418 (2011), arXiv:1102.0996.

[36] J. Veitch and A. Vecchio, Phys. Rev. D **81** (2010).