Stealth bias in gravitational-wave parameter estimation

Michele Vallisneri and Nicolás Yunes

Stealth Bias in Gravitational-Wave Parameter Estimation

Michele Vallisneri[1] and Nicolás Yunes[2]

[1] *Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA.*
[2] *Department of Physics, Montana State University, Bozeman, MT 59717, USA.*

Inspiraling binaries of compact objects are primary targets for current and future gravitational-wave observatories. Waveforms computed in General Relativity are used to search for these sources, and will probably be used to extract source parameters from detected signals. However, if a different theory of gravity happens to be correct in the strong-field regime, source-parameter estimation may be affected by a *fundamental bias*: that is, by systematic errors due to the use of waveforms derived in the incorrect theory. If the deviations from General Relativity are not large enough to be detectable on their own and yet these systematic errors remain significant (i.e., larger than the statistical uncertainties in parameter estimation), fundamental bias cannot be corrected in a single observation, and becomes *stealth bias*. In this article we develop a scheme to determine in which cases stealth bias could be present in gravitational-wave astronomy. For a given observation, the answer depends on the detection signal-to-noise ratio and on the strength of the modified-gravity correction. As an example, we study three representative stellar-mass binary systems that will be detectable with second-generation ground-based observatories. We find that significant systematic bias can occur whether or not modified gravity can be positively detected, for correction strengths that are not currently excluded by any other experiment. Thus, stealth bias may be a generic feature of gravitational-wave detections, and it should be considered and characterized, using expanded models such as the parametrized post-Einstein framework, when interpreting the results of parameter-estimation analyses.

PACS numbers: 04.80.Cc, 04.80.Nn, 04.30.-w, 04.50.Kd

I. INTRODUCTION

Although General Relativity (GR) has been shown to be an excellent description of gravitation in all astronomical, astrophysical, and laboratory observations carried out so far [1], this theory remains largely untested in the nonlinear and dynamical *strong-field* regime, where gravitational fields are large and rapidly evolving, and velocities are not small. Future gravitational-wave (GW) observations of the late inspiral and coalescence of compact binaries will probe this regime and allow for new tests of GR. To do so, however, the assumption that GR is correct must be relaxed in GW science, and data analysis must be carried out with more general waveform families than those predicted by GR.

The parameterized post-Einsteinian (ppE) framework [2] was recently proposed as a generic solution for this need. In this framework, one deforms the GR waveforms through amplitude (A) and phase (Ψ) corrections expanded in a polynomial basis: $\delta\Psi = \beta_i u^{b_i}$ and $\delta A = \alpha_i u^{a_i}$, where u is the (dimensionless) reduced GW frequency and $(\alpha_i, \beta_i, a_i, b_i)$ are ppE parameters. Of these, a_i and b_i control the type of modified-gravity (MG) corrections, while α_i and β_i control their magnitude. Recent data-analysis investigations of this framework, both in its full form [3] and in a reduced version restricting the allowed b_i to those that appear in post-Newtonian (PN) expansions [4], suggest that ppE could be used in realistic settings.

The simplest ppE model includes a single amplitude and phase correction to the waveform, and it is still sufficiently general to extend the PN-coefficient tests[1] of Refs. [5–7] to all currently studied MG models, including Brans–Dicke theory [8–16], dynamical Chern–Simons gravity [17–21], phenomenological massive-graviton propagation [9, 10, 14, 22–24], gravitational Lorentz violation [25], gravitational parity violation [19, 26], violations of Local Position Invariance [27], and the existence of extra-dimensions [28]. More complicated ppE models posit a piecewise-specified hybrid waveform model that includes the merger and ringdown phases, as well as multi-exponent inspiral corrections [2].

Neglecting the prospect that GR might be an inaccurate description of Nature in the strong-field regime may lead to the mischaracterization of the GW Universe once detections are made. This pitfall has been labeled *fundamental bias* [2] in analogy to other types of bias in observational astronomy. Imagine, for example, that the behavior of gravitation is correctly described by a MG theory that allows for scalar radiation only in the very late stages of inspiral. If so, the final inspiral will proceed faster than would be expected in GR, and less power may be emitted in GWs during ringdown. Such events might still be detected with GR templates, albeit with suboptimal efficiency, but they will appear closer than they really are, introducing a systematic error in the estimated luminosity distance.

This paper discusses how we may quantify the degree to which fundamental bias could lead us to incorrect in-

---

[1] These tests treat the coefficients in the PN expansion of the inspiral-waveform phasing as independent parameters rather than fixed functions of the binary parameters, and verify the consistency of their measured values.

ferences about the parameters of detected GW sources. The magnitude of the errors introduced will of course depend on the magnitude of the MG correction. In this first study, we will model the correction with the simplest ppE inspiral waveform, which includes a single-exponent modification to the phasing. We ask the following question: As a function of correction exponent b and magnitude β, what is the detection signal-to-noise ratio such that the systematic error due to using GR templates is equal to the statistical, noise-induced error in the determination of the source parameters, such as the chirp mass and mass ratio? We then ask: At that SNR, is there sufficient statistical evidence, as quantified by Bayesian model comparison [29], to claim that the MG model is preferred? If the systematic error is relatively large, but model comparison still prefers GR, we would find ourselves in a situation of *stealth bias* [3] that makes it logically impossible to correct the systematic effects of MG theories, at least in individual observations.

Fundamental bias in GW observations is similar to the *mismodeling bias* considered by Cutler and Vallisneri (who call it *theoretical* error [30]): the idea is that the approximations used to solve the Einstein equations and construct waveform templates may introduce parameter-estimation errors that are larger than statistical uncertainties. Cutler and Vallisneri developed simple formulas to estimate likely mismodeling errors; in this paper we use them in the ppE context. We also use Vallisneri's recent formula [31] to compute the Bayesian odds ratio, as measured in GW observations, for GR versus a MG theory linked to GR by one or more continuous parameters. These formulas are valid for sufficiently large detection SNRs, but for weaker signals a full-scale Monte Carlo approach is needed [3]. In this paper, we limit ourselves to exemplifying the logic of our approach with simpler analytic tools of more limited applicability, and we leave a detailed Monte Carlo analysis for future work.

As an example, we concentrate on the signals from the circular–adiabatic inspiral of nonspinning stellar-mass binaries, as detected by a network of second-generation ground-based observatories. The signals include ppE phasing corrections with a range of exponents and strengths. We find that, for likely detection SNRs, three cases are indeed possible, depending on the strength of the correction:

   i. Fundamental bias is significant, but statistical evidence is sufficient to positively detect MG effects, so the bias can be corrected;

   ii. Fundamental bias is significant, but statistical evidence is *insufficient* to detect MG effects, so stealth bias is present;

   iii. Fundamental bias is not significant and MG effects cannot be detected.

For exponents b corresponding to quadrupole- or higher-order PN terms (i.e., $b \geq -5/3$), stealth bias can occur

for correction strengths β that are *not* currently excluded by other constraints, such as those from binary pulsars.

This paper is organized as follows. Section II describes our formulation; Sec. III discusses our analysis and results; Sec. IV presents our conclusions. Throughout this paper we use the conventions of Misner, Thorne, and Wheeler [32], such as geometric units with $G = c = 1$.

II. FORMULATION

In this section we describe our analytical formulation. Subsection II A introduces ppE waveforms for inspiraling binaries; Subsecs. II B and II C provide details about the estimation of fundamental bias and the model-comparison detection of MG; Subsec. II D ties all of these together to characterize the impact of fundamental bias in typical detections and the occurrence of stealth bias.

A. Systematic bias and the ppE framework

GW searches performed with matched filtering can be affected by systematic bias if the theoretical signal templates used in the search do not match exactly the true signals that appear in the data. Symbolically, $h_{\text{true}} = h_{\text{theory}} + \delta h$, where the unmodeled correction δh can have different origins [33]:

- **Mismodeling**: caused by the approximations used in solving the equations that describe the motion of the GW source, and the generation and propagation of GWs [30].

- **Instrumental**: caused by the approximations used in modeling the response of the detector to GWs [34, 35].[2]

- **Astrophysical**: caused by unmodeled astrophysics (e.g., neutron-star hydrodynamics in NS–NS inspirals) or objects (e.g., a third star in the close vicinity of a binary) [33, 36, 37].

- **Fundamental**: caused by MG corrections to Einstein's equations [2, 3, 31].

Mismodeling bias is remedied by deriving ever more accurate solutions to the field equations: for binary inspirals, this goal is currently pursued by pushing the PN approximation to higher orders, and by integrating together

———

[2] Of course, the first thing that comes to mind when considering instrumental effects is the parameter-estimation error due to detector noise, which has a stochastic rather than systematic nature. These errors *may* be difficult to characterize statistically if the noise is significantly non-Gaussian or nonstationary, but they are distinct from the systematic effects that we address in this paper, since they can be alleviated only by improving the instrument itself, rather than its theoretical description or the modeling of sources.

PN and numerical-relativity results with analytical resummation and fitting techniques, such as the effective-one-body scheme. Instrumental bias is reduced by careful detector modeling and characterization. Astrophysical bias is expected to be irrelevant for most ground-based binary observations; but even if this were not the case, astrophysical effects should present themselves differently (or not at all) in observations of different systems, depending on the contingent characteristics of the perturbation. By contrast, fundamental bias, if present, should appear consistently in all observed systems, although it would of course depend on the specific system parameters.

In this paper, we concentrate on the inspiral signals from compact-binary coalescences. We consider inspirals that are circular and adiabatic, with negligible spin effects, and neglect mismodeling bias by assuming the GW emission is well-described by the restricted PN waveform in the frequency-domain, stationary-phase approximation [38–41]. In GR, the resulting signals can be written as

$$h_{\mathrm{GR}}(f) = A_{\mathrm{GR}}(f) e^{i\Psi_{\mathrm{GR}}(f)}, \quad (1)$$

where $A_{\mathrm{GR}}(f) = \mathcal{A} u^{-7/6} [1 + \cdots]$ (we neglect PN amplitude corrections, symbolized with ellipses in the above equation), $u = \pi \mathcal{M} f$ is the reduced frequency, $\mathcal{M} = \eta^{3/5} M$ the chirp mass, $\eta = m_1 m_2/M^2$ the symmetric mass ratio, $M = m_1 + m_2$ the total mass, and f the GW frequency. The constant amplitude $\mathcal{A}$ depends on the chirp mass, the luminosity distance, and the detector's antenna patterns [38–41]. The quantity $\Psi_{\mathrm{GR}}(f)$ in Eq. (1) is the GW phase, given in the PN approximation by

$$\Psi_{\mathrm{GR}} = 2\pi f t_c - \phi_c - \frac{\pi}{4}$$
$$+ \frac{3u^{-5/3}}{128} \left\{ 1 + \sum_{k=2}^{7} \left[\psi_k + \frac{1}{3}\bar{\psi}_k \log(u) \right] \eta^{-k/5} u^{k/3} \right\}, \quad (2)$$

where the constant coefficients $(\psi_k, \bar{\psi}_k)$ can be obtained (for instance) from those in Ref. [30].

Under these assumptions, the unmodeled corrections enumerated above can be represented by a continuous and (in principle) predictable deformation of the GW phase $\Psi(f)$ and amplitude $A(f)$. The particular deformation depends on the systematic effect. For mismodeling, we expect corrections within the structure of the PN series ($\delta A \propto f^{-7/6+k_A/3}$ for the amplitude and $\delta\Psi \propto f^{-5/3+k_\phi/3}$ for the phase, with integers $k_A, k_\phi > 0$), beyond the highest known perturbative order (i.e., $k_A > 5$ for the amplitude and $k_\phi > 7$ for the phase).

For astrophysical effects, we expect corrections to arise almost always with "negative" PN exponents [33, 36, 37, 42]. For example, an accretion disk [42], the presence of a third body [36], and orbital eccentricity [42] all introduce GW phase corrections $\delta\Psi \propto f^{-5/3-k'_\phi}$, with integer $k'_\phi > 0$. Physically, this frequency dependence corresponds to

astrophysical effects becoming less important for tighter binaries, where strong-field effects become dominant.

Moving on to unmodeled corrections originating from fundamental physics, the amplitude and phase deformations δA and $\delta\Psi$ can always be expressed as sums of frequency powers, provided that δA and $\delta\Psi$ remain *analytic* at all frequencies sampled during the inspiral:

$$\delta A = A_{\mathrm{GR}}(f) \sum_{k=1}^{K} \alpha_k u^{a_k}, \qquad \delta\Psi(f) = \sum_{k=1}^{K} \beta_k u^{b_k}, \quad (3)$$

where $(\alpha_k, \beta_k, a_k, b_k) \in \mathbb{R}$ for all k, and where we have included the A_{GR} prefactor in δA. We have neglected possible logarithmic terms for simplicity, but they can be included easily in the same fashion. This ppE model introduces $4K$ new parameters in the waveform, but the simplest version of this model would allow only a single exponent:

$$\delta A = A_{\mathrm{GR}}(f) \alpha u^a, \qquad \delta\Psi(f) = \beta u^b. \quad (4)$$

Indeed, it can be shown that such a simple parametric deformation is sufficiently general to model all known MG corrections to the waveform to leading PN order [2, 3], provided that the two tensor polarizations are dominant, as in GR. Otherwise, a second term would be needed in the phase and amplitude [43, 44].

Furthermore, a convincing argument can be made that a and b should be restricted to a few discrete values [44]. Suppose that the adiabatic-inspiral waveform is derived from an energy-balance equation with modified binding energy and flux of the form

$$E = E_0 v^2 \left[1 + (\cdots)_{\mathrm{PN}} + \delta E v^k \right], \quad (5)$$
$$\dot{E} = \dot{E}_0 v^{10} \left[1 + (\cdots)_{\mathrm{PN}} + \delta\dot{E} v^m \right], \quad (6)$$

where v is the relative velocity of the binary components, ellipses stand for higher-order PN terms, and $E_0, \dot{E}_0, \delta E$ and $\delta\dot{E}$ are all constants that may depend on the source parameters and on the MG coupling constants. The exponents k and m must be integers, otherwise E or $\dot{E}$ would not be analytic, and we would lose the guarantee that the equations have a unique solution of hyperbolic character by the Picard–Lindelöf theorem[3]. Furthermore, we must have $k \geq -2$ and $m \geq -10$, otherwise E and $\dot{E}$ would diverge in the $v \to 0$, weak-field limit. These constraints lead to the deformations

$$\delta A = A_{\mathrm{GR}}(f) \alpha u^{\bar{a}/3}, \qquad \delta\Psi(f) = \beta u^{\bar{b}/3}, \quad (7)$$

---

[3] Given the differential equation $dy/dt = f(t, y(t))$, with initial value $y(t_0) = y_0$, a unique solution exists for all $t \in (t_0 - \epsilon, t_0 + \epsilon)$ provided that f is Lipschitz continuous in y and continuous in t. A noninteger value of k and m would lead to a differential equation with a non-Lipschitz continuous source term, with possible loss of uniqueness.

where $(\bar{a}, \bar{b}) \in \mathbb{Z}$, with $(\bar{a}, \bar{b}) > (-10, -15)$ [44].

In this paper we concentrate on phasing corrections by setting $\alpha = 0$ and choosing $\bar{b} \in \{-7, -6, -5, -4, -3, -2, -1, 1, 2\}$. Different values of $\bar{b}$ represent different types of MG effects: $\bar{b} = -7$ corresponds to the leading–PN-order correction in Brans–Dicke theory [8, 9, 14–16, 23, 45] or in Einstein–dilaton–Gauss–Bonnet gravity [19, 46]; $\bar{b} = -3$ to the leading-order term in a phenomenological massive graviton theory [9, 11, 12, 14, 22, 25, 45], $\bar{b} \geq -5$ (but $\neq -4$) to the modified-PN scheme of Refs. [4–7], and $\bar{b} = -1$ to dynamical Chern–Simons gravity [19, 21, 47]. Notice, in particular, that the modified-PN scheme is clearly a sub-case of the ppE scheme. We omit $\bar{b} = 0$ because the resulting correction would be degenerate with an arbitrary constant in the phase. We do not consider $\bar{b} < -7$ because the values that we study provide enough information to observe a consistent trend as $\bar{b}$ becomes more negative. Moreover, for $\bar{b} \leq -7$, binary pulsar observations can do a better job at constraining modified gravity theories than GWs observations [48]. We also do not consider $\bar{b} > 2$, as this would correspond to terms of higher than 3.5 PN order, which we do not account for in the Ψ_{GR}.

B. Quantifying the bias

Let us assume that a GW detection is reported for a dataset s that contains the waveform

$$h_{\mathrm{full}}(\vec{\theta}_{\mathrm{tr}}) = h(\vec{\theta}_{\mathrm{tr}}) + \delta h(\vec{\theta}_{\mathrm{tr}}), \qquad (8)$$

where $\vec{\theta}_{\mathrm{tr}}$ is the vector of parameters that describes the GW source and source–detector geometry, h is the approximated waveform family used to filter the data, and δh is the unmodeled correction to h. Following Cutler and Vallisneri [30], we compute the *theoretical* error $\delta\vec{\theta}_{\mathrm{th}}$ induced by matched-filtering with h instead of h_{full}.

The theoretical error $\delta\vec{\theta}_{\mathrm{th}}$ is defined as the displacement $\vec{\theta}_{\mathrm{bf}} - \vec{\theta}_{\mathrm{tr}}$ between the true parameters $\vec{\theta}_{\mathrm{tr}}$ and the *best-fit* parameters $\vec{\theta}_{\mathrm{bf}}$ that would maximize the likelihood in the absence of noise. When $\delta h(\vec{\theta}_{\mathrm{tr}})$ is negligibly small, $\vec{\theta}_{\mathrm{bf}} = \vec{\theta}_{\mathrm{tr}}$; as $\delta h(\vec{\theta}_{\mathrm{tr}})$ grows in magnitude, $\vec{\theta}_{\mathrm{bf}}$ is displaced further and further away along the parameter-space direction in which $h(\vec{\theta}_{\mathrm{bf}})$ can reproduce $h(\vec{\theta}_{\mathrm{tr}}) + \delta h(\vec{\theta}_{\mathrm{tr}})$ most closely.

To leading order in δh, $\delta\vec{\theta}_{\mathrm{th}}$ is given by [30]

$$\delta\vec{\theta}_{\mathrm{th}} = (F_{\mathrm{bf}}^{-1})^{\alpha\beta} \big(h_{,\beta}(\vec{\theta}_{\mathrm{bf}}) \big| \delta h(\vec{\theta}_{\mathrm{bf}})\big), \qquad (9)$$

where $h_{,\beta} = \partial h / \partial \theta^\beta$ are partial derivatives of the waveform with respect to source parameters, $F_{\alpha\beta} = (h_{,\alpha}|h_{,\beta})$ is the *Fisher matrix*, here evaluated at θ_{bf}, and

$$(g_1|g_2) = 4 \operatorname{Re} \int_0^\infty \frac{\tilde{g}_1^*(f)\tilde{g}_2(f)}{S_n(f)} \, \mathrm{d}f, \qquad (10)$$

is the noise-weighted signal inner product, with $S_n(f)$ the one-sided power spectral density of detector noise (see, e.g., [38]). The inner product defines a signal *norm* $|h|$ by way of $|h|^2 = (h|h)$.

In Eq. (9), the waveform correction δh is projected onto the waveform derivatives, and the projection cosines are mapped into parameter errors by the inverse Fisher matrix F^{-1}, thus taking into account possible parameter covariances. Note that the resulting $\delta\vec{\theta}_{\mathrm{th}}$ is independent of the detection SNR $|h_{\mathrm{full}}(\vec{\theta}_{\mathrm{tr}})|$, since both F and $(h_{,\beta}(\theta_{\mathrm{bf}})|\delta h(\theta_{\mathrm{bf}}))$ are quadratic in the waveform amplitude.

Equation (9) is accurate only for small $\delta\vec{\theta}_{\mathrm{th}}$—more precisely, for perturbations small enough that $h(\vec{\theta}_{\mathrm{bf}} - \delta\vec{\theta}_{\mathrm{th}}) \simeq h(\vec{\theta}_{\mathrm{bf}}) - h_{,\alpha}\delta\theta_{\mathrm{th}}^\alpha$. Cutler and Vallisneri [30] discuss more sophisticated versions of Eq. (9) that can be applied to larger perturbations, but in this paper we adopt the simpler Eq. (9), not least because the other ingredients in our formulation depend on $\delta h(\vec{\theta}_{\mathrm{tr}})$ being small.

C. Detecting modified gravity

Following Vallisneri [31] (see also [3]), we define a MG correction δh to the signal h to be detectable when the *odds ratio* of the Bayesian evidences for the MG and pure-GR scenarios, used as a detection statistic, is large enough that the *false-alarm probability* of favoring the MG hypothesis when GR is in fact correct is suitably small. More precisely, we compute the odds ratio

$$\mathcal{O} = \frac{P(\mathrm{MG}|s)}{P(\mathrm{GR}|s)} = \frac{P(\mathrm{MG}) \int p(s|\vec{\theta}, \vec{\lambda}) \, p(\vec{\theta}, \vec{\lambda}) \, \mathrm{d}^k\theta \, \mathrm{d}^m\lambda}{P(\mathrm{GR}) \int p(s|\vec{\theta}) \, p(\vec{\theta}) \, \mathrm{d}^k\theta}, \qquad (11)$$

where $P(\mathrm{MG})$ and $P(\mathrm{GR})$ are the prior probabilities that MG and GR are correct, $p(s|\vec{\theta}, \vec{\lambda})$ is the likelihood that the detector data s contains the MG waveform $h(\vec{\theta}) + \delta h(\vec{\theta}, \vec{\lambda})$, $p(s|\vec{\theta})$ is the likelihood that s contains the pure-GR waveform $h(\vec{\theta})$, and $p(\vec{\theta}, \vec{\lambda}) = p(\vec{\theta})p(\vec{\lambda})$ and $p(\vec{\theta})$ are the prior probability densities for the source parameters $\vec{\theta}$ and the MG parameters $\vec{\lambda}$.

If the true signal is MG, using MG templates would improve the fit to the data and increase the maximum value attained by the MG likelihood relative to the GR likelihood. On the other hand, the evidence for the more complicated, higher-dimensional MG model is reduced by the smaller *prior mass* within the support of the likelihood—the mechanism by which Bayesian inference embodies Occam's principle of parsimony. As signals get stronger, the improvement in the likelihood grows exponentially with the (squared) detection SNR, and eventually it overcomes the effect of the priors.

Even if we fix the true signal, the odds ratio $\mathcal{O}$ remains a random variable, because it depends on the realization of detector noise, by way of the likelihoods. If the GW signal itself (rather than the MG correction) has been detected with a sufficiently large detection SNR

$|h_{\text{full}}(\vec{\theta}_{\text{tr}}, \vec{\lambda})|$, it can be shown [31] that Eq. (11) becomes remarkably simple: for the cases when the underlying signal is pure-GR or MG respectively, we find that

$$\begin{aligned}
\mathcal{O}_{\text{GR}} &= \mathcal{N} e^{x^2/2}, \\
\mathcal{O}_{\text{MG}} &= \mathcal{N} e^{x^2/2 + \sqrt{2}\, x\, \text{SNR}_{\text{MG}} + \text{SNR}_{\text{MG}}^2}.
\end{aligned} \qquad (12)$$

In this equation:

- The dependence on the noise realization enters exclusively through x, a normal random variable with zero mean and unit variance. Technically, x is given by the inner product of detector noise with the "MG-unique" component of the MG correction δh (the component that is orthogonal to the derivatives of the signal with respect to the GR parameters).

- The constant of proportionality $\mathcal{N}$ is the same in both rows: it is a function of $P(\text{MG})$, $P(\text{GR})$, and of the estimation errors and prior-density widths for the MG parameters $\vec{\lambda}$. As we shall see below, the fact that this constant is the same under both hypotheses allows us to disregard it when we analyze the statistics of our detection scheme.

- Last, $\text{SNR}_{\text{MG}} \equiv \text{SNR}\sqrt{1 - \text{FF}}$ is (by definition) the *effective MG-detection SNR*, with FF the *fitting factor* between the pure-GR and MG waveforms,

$$\text{FF}(\vec{\theta}, \vec{\lambda}) = \max_{\vec{\theta}'} \frac{\big(h(\vec{\theta}')\big| h(\vec{\theta}) + \delta h(\vec{\theta}, \vec{\lambda})\big)}{|h(\vec{\theta}')| \times |h(\vec{\theta}) + \delta h(\vec{\theta}, \vec{\lambda})|}. \qquad (13)$$

This FF describes the maximum fraction of detection SNR that can be recovered when searching for MG signals using GR templates. The FF is itself independent of SNR, and in our application it is given by

$$\frac{\big(h(\vec{\theta}_{\text{tr}} + \delta\vec{\theta}_{\text{th}})\big| h(\vec{\theta}_{\text{tr}}) + \delta h(\vec{\theta}_{\text{tr}}, \vec{\lambda})\big)}{|h(\vec{\theta}_{\text{tr}} + \delta\vec{\theta}_{\text{th}})| \times |h(\vec{\theta}_{\text{tr}}) + \delta h(\vec{\theta}_{\text{tr}}, \vec{\lambda})|}, \qquad (14)$$

or also by $1 - \frac{1}{2}|\delta h(\vec{\theta}_{\text{tr}}, \vec{\lambda})|^2 / |h(\vec{\theta}_{\text{tr}})|^2$ [31]. These formulas are equivalent, but only to leading order in $\vec{\lambda}$.

By solving Eqs. (12) for $\mathcal{O}_{\text{GR}}(x)$ and $\mathcal{O}_{\text{MG}}(x)$, we can derive the cumulative probability distributions of $\mathcal{O}_{\text{GR}}$ and $\mathcal{O}_{\text{MG}}$ from the normal probability distribution $p(x) = e^{-x^2/2}/\sqrt{2\pi}$. For instance,

$$P(\mathcal{O}_{\text{GR}} > \mathcal{O}^*) = \int_{x(\mathcal{O}^*)}^{\infty} p(x)\, \mathrm{d}x : \qquad (15)$$

this is exactly the false-alarm probability P_F in a *decision scheme* that declares an MG detection when the odds ratio $\mathcal{O}$ computed from the data is larger than the *threshold* $\mathcal{O}^*$. By contrast, the true-detection probability (also known as the detection *efficiency*) is given

by $P_E \equiv P(\mathcal{O}_{\text{MG}} > \mathcal{O}^*)$, which we compute along the lines of Eq. (15), and which is a function of SNR_{MG}. In practice, we set $\mathcal{O}^*$ by requiring a sufficiently small P_F. In this paper, we use $P_F = 10^{-4}$, which seems appropriate for the tens of detections expected from second-generation ground-based detectors (but see also the discussion in Sec. III of [31]). Because all that matters in the decision scheme is how often $\mathcal{O}_{\text{GR}}$ and $\mathcal{O}_{\text{MG}}$ are larger than the same threshold $\mathcal{O}^*$, we can work with renormalized odds ratios divided by $\mathcal{N}$, which dispenses with the need of computing $\mathcal{N}$, and shows that the priors $P(\text{GR})$ and $P(\text{MG})$ do not enter the statistics of the decision scheme (although they enter its interpretation [31]).

We can now replace the vague statement that began this section with a precise quantitative condition: we will say that a MG correction becomes detectable once SNR_{MG} is large enough to yield an efficiency of $1/2$. This is a conventional value, but once P_E reaches $1/2$ it grows rapidly to 1 with increasing SNR_{MG}. For $P_F = 10^{-4}$, we get $P_E = 1/2$ when $\text{SNR}_{\text{MG}} = 2.75$. Because $\text{SNR}_{\text{MG}} \equiv \text{SNR}\sqrt{1 - \text{FF}}$, the SNR at which MG effects are detectable is then simply $\text{SNR}^{\text{detect}} = 2.75(1 - \text{FF})^{-1/2}$. We thus confirm our expectation that MG detectability improves for larger detection SNR or for larger MG corrections, which lead to smaller FF.

D. Statistical error and stealth bias

We have by now collected the tools to estimate the fundamental error $\delta\vec{\theta}_{\text{th}}$ due to MG corrections, and to determine the detection SNR necessary to positively detect the presence of MG corrections of a certain magnitude. The last ingredient in our formulation are the statistical errors in the determination of the source parameters $\vec{\theta}$, which we estimate as the (square-root) diagonal elements of the inverse Fisher matrix, computed at $\vec{\theta}_{\text{tr}}$ for GR waveforms:

$$\delta\theta^\alpha_{\text{stat}} = \sqrt{(F_{\text{tr}}^{-1})^{\alpha\alpha}} \qquad (16)$$

(no summation is implied by the repeated index). This expression implies that $\delta\theta^\alpha_{\text{stat}}$ scales as $1/\text{SNR}$. Indeed, to leading order in $1/\text{SNR}$ the inverse Fisher matrix describes the variance of the maximum-likelihood parameter estimator across noise realizations, as well as the shape of the Bayesian posterior parameter distribution around its mode when priors can be neglected [49].

For a fixed MG-correction magnitude β, the statistical error decreases with increasing detection SNR, but the systematic bias remains constant and eventually becomes the limiting factor for parameter-estimation accuracy. If this happens before MG can be positively detected, the observation may suffer from stealth bias. Thus, we will look for the detection SNR at which $\max_\alpha \delta\theta^\alpha_{\text{th}}/\delta\theta^\alpha_{\text{stat}} = 1$, and compare it to the detection SNR at which $\text{SNR}_{\text{MG}} \equiv \text{SNR}\sqrt{1 - \text{FF}} = 2.75$. We

will denote these two limiting SNRs as $\mathrm{SNR}^{\mathrm{bias}}$ and $\mathrm{SNR}^{\mathrm{detect}}$, respectively.

We should not forget that our formulation, which relies on leading-order expansions of the waveforms as functions of the source parameters, is only valid for sufficiently large detection SNRs, for which the expected errors in the source parameters are small enough to confine these to a small neighborhood of the true parameters, where the leading-order expansion is accurate. Unfortunately, it is not possible to give a generally valid estimate of the required SNR, which will depend on the specific system and waveform family under consideration (for instance, Ref. [50] finds that $\mathrm{SNR} > 20$ would be sufficient for Initial-LIGO detections of nonspinning stellar-mass binaries). We reckon however that leading-order results will *not* be predictive of parameter-estimation performance for most near-threshold detections, and therefore for most initial detections with ground-based detectors.

Nevertheless, for a given system and waveform model it is possible to check whether first-order results are at least self consistent. To do this, we adopt Vallisneri's 2008 criterion [49] for the statistical errors predicted by the Fisher-matrix formalism, and we require that the error made in approximating $h(\vec{\theta}_{\mathrm{tr}} + \delta\vec{\theta}_{\mathrm{stat}}) - h(\vec{\theta}_{\mathrm{tr}})$ as $h_{,\alpha}\delta\theta^\alpha_{\mathrm{stat}}$ be sufficiently small (0.1 in norm) on most (95%) of the $1-\sigma$ error surface described by $(F^{-1})^{\alpha\beta}$. For the theoretical errors, we require that the FFs computed in the two ways described below Eq. (13) be consistent to 1%. The resulting self-consistency thresholds lie somewhere in the middle of the curves in Figs. 1. This circumstance does not invalidate our analysis, since the most interesting detections for GR-testing purposes will be at the high-SNR end, while the low-SNR results should be at least indicative of general trends.

III.   ANALYSIS AND RESULTS

We examine three representative binary GW sources for second-generation interferometric detectors such as Advanced LIGO [51]: neutron-star–neutron-star binaries with $(1.4+1.4)M_\odot$ component masses; neutron-star–black-hole binaries with $(1.4+5)M_\odot$ masses; and black-hole–black-hole binaries with $(5+10)M_\odot$ masses.

We concentrate on the inspiral phase of coalescence, which we model as quadrupolar and adiabatically quasi-circular with 3.5-PN-accurate phasing. We truncate the waves at the innermost stable circular orbit of a point-particle in a Schwarzschild background (assuming GR), and we neglect spin effects and PN amplitude corrections. The resulting waveforms are described by nine parameters: the two masses (or the chirp and reduced masses), the time and inspiral phase at coalescence, two sky-position angles, two angles that describe the binary inclination and GW polarization, and the luminosity distance (see [30] and [3] for a similar waveform prescription). We assume a simultaneous detection by three second-generation detectors with the LIGO Hanford, LIGO Liv-

ingston, and Virgo geometries and relative delays [52], and with identical broadband-configuration power spectral densities, as given by Eq. (10) of [3]. Furthermore, we assume that GW-detector noise is Gaussian and stationary, as required by the Cutler–Vallisneri [30] and Vallisneri [31] formalisms.

For these systems, we consider ppE phasing corrections $\delta\Psi$ as described in Sec. II A, and we compare $\mathrm{SNR}^{\mathrm{bias}}$ and $\mathrm{SNR}^{\mathrm{detect}}$ as a function of the MG-correction magnitude β for a range of exponents $\bar{b}$. For each mass combination, each $\bar{b}$, and each β, we randomly select 1,000 configurations of the phase and angle parameters from the appropriate uniform distributions (e.g., sky positions are chosen randomly on the celestial sphere). The luminosity distance is reabsorbed in the SNR scaling, while the time of coalescence has no effect on our computation. For each configuration we compute $\mathrm{SNR}^{\mathrm{detect}}$ and $\mathrm{SNR}^{\mathrm{bias}}$ and we report their median values. The condition $\max_\alpha \delta\theta^\alpha_{\mathrm{th}}/\delta\theta^\alpha_{\mathrm{stat}} = 1$ that yields the latter is usually satisfied first for the chirp-mass parameter. Statistical fluctuations around the median turn out to be rather small (a few percent).

The $\delta\vec{\theta}_{\mathrm{stat}}$ consistency check is satisfied for detection SNRs ranging from 10 to 100, typically ~ 50, but our results for lower SNRs should be at least representative of trends. The $\delta\vec{\theta}_{\mathrm{th}}$ check is satisfied for a maximum β that depends on $\bar{b}$, and which sets the largest β that we investigate. By contrast, the smallest β that we study corresponds to $\min(\mathrm{SNR}^{\mathrm{detect}}, \mathrm{SNR}^{\mathrm{bias}}) = 100$, a relatively large detection SNR that would be achieved very rarely in volume-limited searches [53].

Figure 1 presents the main results of this paper, with the solid curves plotting $\mathrm{SNR}^{\mathrm{detect}}$ (once again, the detection SNR above which MG can be detected positively), and the dashed curves plotting $\mathrm{SNR}^{\mathrm{bias}}$ (the SNR above which the largest ratio of fundamental error to statistical error reaches one). Both sets of curves are plotted as a function of the MG-correction magnitude β; the curves in each set correspond to $\bar{b} = -7, -5, -4, -3, -2, -1, 1$, and 2, from left to right as labeled. The top, mid, and bottom panels show results for our $(1.4+1.4)M_\odot$, $(1.4+5)M_\odot$, and $(5+10)M_\odot$ systems, respectively.

This figure reveals a few interesting features. First, for the same detection SNR, more massive systems require larger β before MG can be detected. This must happen because the larger the mass, the fewer the number of useful GW cycles in the detector's band, so the signals become relatively featureless, and higher FFs can be obtained for the same β. Of course, this behavior would probably be affected if one included the merger and ring-down phases of coalescence. Second, all the curves are relatively steep, with typical detection SNRs mapping out a small range of β for each $\bar{b}$. Indeed, our analytical formulation predicts $\mathrm{SNR}^{\mathrm{detect}}$ and $\mathrm{SNR}^{\mathrm{bias}}$ proportional to β^{-1} at leading order (this is because $\delta h \propto \beta$, and both $\delta\vec{\theta}_{\mathrm{th}}$ and $\mathrm{SNR}_{\mathrm{MG}} \propto \delta h$). The curves are less steep (and much closer) at the lower SNRs (~ 10) expected for the the first, volume-limited detections of ground-based de-
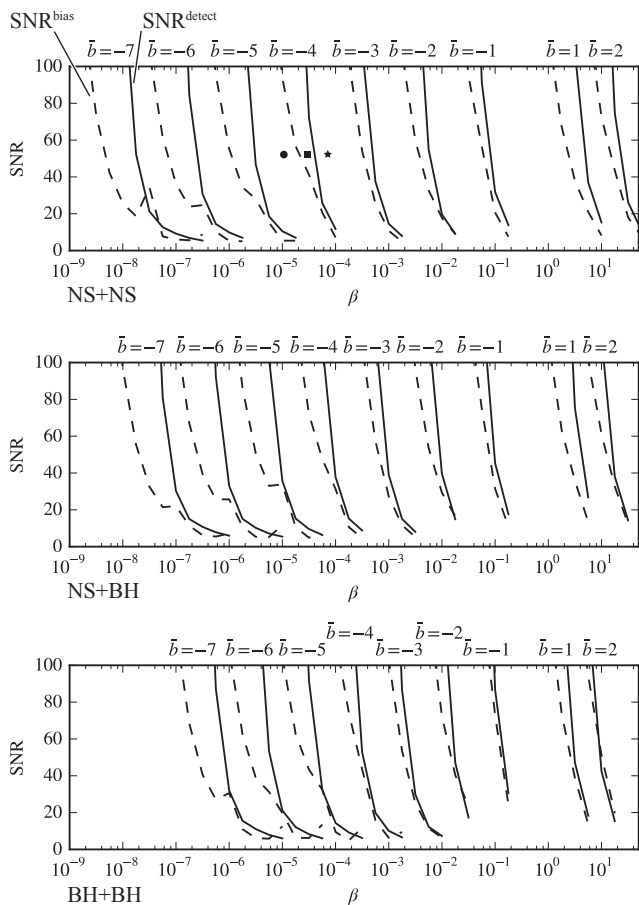
FIG. 1. $\text{SNR}^{\text{detect}}$ (solid curves) and SNR^{bias} (dashed curves) as a function of β for $\bar{b} = -7, -6, -5, -4, -3, -2, -1, 1$ and 2 (left to right), for a NS–NS system with $(m_1, m_2) = (1.4, 1.4)\, M_\odot$ (top), a NS–BH system with $(m_1, m_2) = (1.4, 5)\, M_\odot$ (mid) and a BH–BH system (bottom) with $(m_1, m_2) = (5, 10)\, M_\odot$. The symbols in the top plot are described in the main text.

tectors; remember however that the linearized-waveform approximation is probably inappropriate for those detections. Third, while we computed all our results for positive β, our formulation is invariant with respect to a change in its sign: with everything else held fixed, the change yields $\delta h \to -\delta h$ and $\delta\vec{\theta}_{\text{th}} \to -\delta\vec{\theta}_{\text{th}}$, which leaves $\text{SNR}^{\text{detect}}$ and SNR^{bias} unchanged. Fourth, the dashed curves present non-monotonic behavior at small SNR, which occurs when the best-determined parameter switches from the chirp mass to another parameter (usually, but not for all realizations, the total mass). Fifth, and most important, the SNR^{bias} curves lie almost always to the *left* of the $\text{SNR}^{\text{detect}}$ curves, so *stealth bias* (which corresponds to the area between the solid and dashed line) *is a generic feature for these waveforms*, and it becomes more prevalent for more negative $\bar{b}$.

Thus, if ground-based GW detectors observe a signal from these binaries with a certain detection SNR, and

Nature happens to deviate from GR in the strong-field regime, three cases are possible (with somewhat blurry boundaries) depending on the value of β. They are illustrated by the small symbols in the top panel of Fig. 1 for the $\bar{b} = -4$ case, and listed below:

- **Overt bias: significant fundamental bias, detectable MG deviation**: $\text{SNR}^{\text{bias}}(\beta) < \text{SNR}^{\text{detect}}(\beta) < \text{SNR}$ (rightmost symbol, star). The systematic error induced by using GR templates is larger than the statistical uncertainty, but we are able to infer that the signal is described by a non-GR theory. Thus, we can correct the fundamental bias, as long as our MG model is sufficiently close to the true theory of gravity.

- **Stealth bias: significant fundamental bias, undetectable MG deviation**: $\text{SNR}^{\text{bias}}(\beta) < \text{SNR} < \text{SNR}^{\text{detect}}(\beta)$ (middle symbol, square). The systematic error is larger than the statistical uncertainty, but we do not have sufficient statistical evidence to determine that the signal is described by a MG theory. Thus, we cannot correct the fundamental bias. Parameter-estimation results could be prefaced with the caveat that they may include a systematic error as large as the statistical uncertainty *if* GR is incorrect in the strong-field with β in this range.

  For instance, for the systems in Fig. 1, at detection SNR ~ 50 the median statistical uncertainty in the chirp mass (the best-determined parameter) is 0.02–0.04%, and the stealth systematic errors range from $1\times$ (at the small-β end corresponding to $\text{SNR}^{\text{bias}} = 50$) to several $\times$ those values (at the large-β end corresponding to $\text{SNR}^{\text{detect}} = 50$).

- **Negligible bias: insignificant fundamental bias, undetectable MG deviation**: $\text{SNR} < \text{SNR}^{\text{bias}}(\beta) < \text{SNR}^{\text{detect}}(\beta)$ (leftmost symbol, circle). Although we cannot determine that the signal is described by a MG theory (i.e., the signal is indistinguishable from the GR prediction), the systematic error is smaller than the statistical uncertainty. The accuracy of parameter estimation is dominated by statistical errors and not by modified gravity effects. This is also trivially the case when GR is correct.

We stress that the generic appearance of stealth bias is an empirical finding that we have established only for these waveforms. Stealth bias is not inevitable: if we had found that in general $\text{SNR}^{\text{bias}}(\beta) > \text{SNR}^{\text{detect}}(\beta)$, systematic error would only appear in observations where the presence of MG corrections is obvious. Although our analysis is robust for the sources we examined, we cannot extend our conclusions to different sources, different waveform families, or different detectors. We leave this to future work.

$\bar{b}$	-7	-6	-5	-4	-3	-2	-1	1	2
upper limit on β	4×10^{-10}	4×10^{-7}	3×10^{-4}	3×10^{-1}	3×10^{2}	3×10^{5}	5×10^{8}	4×10^{14}	2×10^{17}

TABLE I. Binary pulsar constraints on β for different values of $\bar{b}$ [48].

Of course, overt and (particularly) stealth bias are only a concern if they are not already excluded *a priori* by previously obtained constraints on β for a given $\bar{b}$. For instance, binary-pulsar observations [48] provide very strong constraints for $\bar{b} \leq -6$ (see Table I), but deteriorate rapidly for corrections corresponding to higher-order PN terms [3]. For this reason, it would be very interesting to map all current binary-pulsar and Solar-System constraints to the ppE framework.

IV. CONCLUSIONS

In this paper we introduced a scheme to determine if the systematic errors induced by the *a priori* assumption that GR is always correct could become as significant as noise-induced statistical uncertainties *before* a MG effect could be detected confidently using Bayesian model comparison. We find that indeed such stealth biases appear generically in analysis of inspiral GWs from stellar-mass compact binaries, as observed by second-generation ground-based detectors, at least for certain ranges of detection SNRs and MG correction strengths.

The possibility of stealth bias should *not* be regarded as casting a veil of uncertainty on the interpretation of GW detections. Rather, it is a warning that inferences from GW observations must be considered in the context of all the other evidence for GR as the true theory of gravitation, as well as all other constraints on MG theories.

For simplicity, we studied non-eccentric, non-precessing systems with no PN amplitude corrections to the waveforms, concentrating only on second-generation ground-based detectors. We employed a simple analytical formulation that is valid for sufficiently large detection SNRs. Obvious extensions of this work could encompass more complete waveform models (perhaps including also the merger and ringdown phases of binary coalescence, using the hybrid ppE models); different source populations; low-SNR detections, which can be tackled with Cutler and Vallisneri's "ODE" approach to evaluate fundamental bias [30] and with Monte Carlo integration to evaluate the Bayesian odds ratio [3]; space-borne detectors or third-generation ground-based detectors. Indeed, we have no doubt that the full power of these techniques will be unleashed once actual detections are made, to look for hints of MG theories and put GR on an ever-firmer footing.

V. ACKNOWLEDGMENTS

[1] C. M. Will, Living Reviews in Relativity **9**, 3 (2006), arXiv:gr-qc/0510072.

[2] N. Yunes and F. Pretorius, Phys. Rev. D **80**, 122003 (2009), arXiv:0909.3328 [gr-qc].

[3] N. Cornish, L. Sampson, N. Yunes, and F. Pretorius, Phys. Rev. D **84**, 062003 (2011), arXiv:1105.2088 [gr-qc].

[4] 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), arXiv:1110.0530 [gr-qc].

[5] K. G. Arun, B. R. Iyer, M. S. S. Qusailah, and B. S. Sathyaprakash, Phys. Rev. **D74**, 024006 (2006), arXiv:gr-qc/0604067.

[6] K. G. Arun, B. R. Iyer, M. S. S. Qusailah, and B. S. Sathyaprakash, Class. Quantum Grav. **23**, L37 (2006), arXiv:gr-qc/0604018.

[7] C. K. Mishra, K. G. Arun, B. R. Iyer, and B. S. Sathyaprakash, Phys. Rev. D **82**, 064010 (2010), arXiv:1005.0304 [gr-qc].

[8] C. M. Will, Phys. Rev. D **50**, 6058 (1994), arXiv:gr-qc/9406022.

[9] C. M. Will and N. Yunes, Class. Quantum Grav. **21**, 4367 (2004), arXiv:gr-qc/0403100.

[10] E. Berti, A. Buonanno, and C. M. Will, Class. Quantum Grav. **22**, S943 (2005), arXiv:gr-qc/0504017.

[11] A. Stavridis and C. M. Will, Phys. Rev. D **80**, 044002 (2009), arXiv:0906.3602 [gr-qc].

[12] K. G. Arun and C. M. Will, Class. Quantum Grav. **26**, 155002 (2009), arXiv:0904.1190 [gr-qc].

[13] D. Keppel and P. Ajith, Phys. Rev. D **82**, 122001 (2010), arXiv:1004.0284 [gr-qc].

[14] K. Yagi and T. Tanaka, Phys. Rev. D **81**, 064008 (2010), arXiv:0906.4269 [gr-qc].

[15] N. Yunes, P. Pani, and V. Cardoso, Phys. Rev. D **85**, 102003 (2012), arXiv:1112.3351 [gr-qc].

[16] J. Healy, T. Bode, R. Haas, E. Pazos, P. Laguna, *et al.*, (2011), arXiv:1112.3928 [gr-qc].

[17] C. F. Sopuerta and N. Yunes, Phys. Rev. D **80**, 064006 (2009), arXiv:0904.4501 [gr-qc].

[18] N. Yunes and F. Pretorius, Phys. Rev. D **79**, 084043 (2009), arXiv:0902.4669 [gr-qc].

[19] K. Yagi, L. C. Stein, N. Yunes, and T. Tanaka, Phys. Rev. D **85**, 064022 (2012), arXiv:1110.5950 [gr-qc].

[20] K. Yagi, N. Yunes, and T. Tanaka, Phys. Rev. D **86**, 044037 (2012), arXiv:1206.6130 [gr-qc].

[21] K. Yagi, N. Yunes, and T. Tanaka, Physical Review Letters **109**, 251105 (2012), arXiv:1208.5102 [gr-qc].

[22] C. M. Will, Phys. Rev. D **57**, 2061 (1998), arXiv:gr-qc/9709011.

[23] P. D. Scharre and C. M. Will, Phys. Rev. D **65**, 042002 (2002), arXiv:gr-qc/0109044.

[24] E. Berti, J. Gair, and A. Sesana, Phys. Rev. D **84**, 101501 (2011), arXiv:1107.3528 [gr-qc].

[25] S. Mirshekari, N. Yunes, and C. M. Will, Phys. Rev. D **85**, 024041 (2012), arXiv:1110.2720 [gr-qc].

[26] N. Yunes, R. O'Shaughnessy, B. J. Owen, and S. Alexander, Phys. Rev. D **82**, 064017 (2010), arXiv:1005.3310 [gr-qc].

[27] N. Yunes, F. Pretorius, and D. Spergel, Phys. Rev. D **81**, 064018 (2010), arXiv:0912.2724 [gr-qc].

[28] K. Yagi, N. Tanahashi, and T. Tanaka, Phys. Rev. D **83**, 084036 (2011), arXiv:1101.4997 [gr-qc].

[29] P. C. Gregory, *Bayesian Logical Data Analysis for the Physical Sciences* (Cambridge University Press, 2005).

[30] C. Cutler and M. Vallisneri, Phys. Rev. **D76**, 104018 (2007), arXiv:0707.2982 [gr-qc].

[31] M. Vallisneri, Phys. Rev. D **86**, 082001 (2012), arXiv:1207.4759 [gr-qc].

[32] C. W. Misner, K. Thorne, and J. A. Wheeler, *Gravitation* (W. H. Freeman & Co., San Francisco, 1973).

[33] N. Yunes, B. Kocsis, A. Loeb, and Z. Haiman, Phys. Rev. Lett. **107**, 171103 (2011), arXiv:1103.4609 [astro-ph.CO].

[34] S. Vitale, W. Del Pozzo, T. G. F. Li, C. Van Den Broeck, I. Mandel, B. Aylott, and J. Veitch, Phys. Rev. D **85**,

[35] J. T. Whelan, E. L. Robinson, J. D. Romano, and E. H. Thrane, (2012), arXiv:1205.3112 [gr-qc].

[36] N. Yunes, M. Coleman Miller, and J. Thornburg, Phys. Rev. D **83**, 044030 (2011), arXiv:1010.1721 [astro-ph.GA].

[37] B. Kocsis, N. Yunes, and A. Loeb, Phys. Rev. D **84**, 024032 (2011), arXiv:1104.2322 [astro-ph.GA].

[38] C. Cutler and É. E. Flanagan, Phys. Rev. D **49**, 2658 (1994), arXiv:gr-qc/9402014.

[39] E. E. Flanagan and S. A. Hughes, Phys. Rev. D **57**, 4535 (1998), arXiv:gr-qc/9701039 [gr-qc].

[40] E. E. Flanagan and S. A. Hughes, Phys. Rev. D **57**, 4566 (1998), arXiv:gr-qc/9710129 [gr-qc].

[41] S. Droz, D. J. Knapp, E. Poisson, and B. J. Owen, Phys. Rev. D **59**, 124016 (1999), arXiv:gr-qc/9901076.

[42] N. Yunes, K. G. Arun, E. Berti, and C. M. Will, Phys. Rev. D **80**, 084001 (2009), arXiv:0906.0313 [gr-qc].

[43] K. Arun, Class. Quantum Grav. **29**, 075011 (2012), arXiv:1202.5911 [gr-qc].

[44] K. Chatziioannou, N. Yunes, and N. Cornish, Phys. Rev. D **86**, 022004 (2012), arXiv:1204.2585 [gr-qc].

[45] E. Berti, A. Buonanno, and C. M. Will, Phys. Rev. D **71**, 084025 (2005), arXiv:gr-qc/0411129.

[46] N. Yunes and L. C. Stein, Phys. Rev. D **83**, 104002 (2011), arXiv:1101.2921 [gr-qc].

[47] S. Alexander and N. Yunes, Phys. Rep. **480**, 1 (2009), arXiv:0907.2562 [hep-th].

[48] N. Yunes and S. A. Hughes, Phys. Rev. D **82**, 082002 (2010), arXiv:1007.1995 [gr-qc].

[49] M. Vallisneri, Phys. Rev. D **77**, 042001 (2008), arXiv:gr-qc/0703086.

[50] T. Cokelaer, Class. Quantum Grav. **25**, 184007 (2008).

[51] G. M. Harry and LIGO Scientific Collaboration, Class. Quantum Grav. **27**, 084006 (2010).

[52] W. G. Anderson, J. T. Whelan, P. R. Brady, J. D. E. Creighton, D. Chin, and K. Riles, *Beam Pattern Response Functions and Times of Arrival for Earthbound Interferometers*, Tech. Rep. LIGO-T010110-00-Z (LIGO, 2001).

[53] J. Abadie, B. P. Abbott, R. Abbott, M. Abernathy, T. Accadia, F. Acernese, C. Adams, R. Adhikari, P. Ajith, B. Allen, and et al., Class. Quantum Grav. **27**, 173001 (2010), arXiv:1003.2480 [astro-ph.HE].