# Probing the Black Hole Metric. I. Black Hole Shadows and Binary Black-Hole Inspirals

Dimitrios Psaltis,[1] Colm Talbot,[2] Ethan Payne,[3, 4] and Ilya Mandel[3, 4, 5]

[1] *Steward Observatory and Department of Astronomy,*
*University of Arizona, 933 N. Cherry Ave., Tucson, AZ 85721, USA*
[2] *LIGO Laboratory, California Institute of Technology, Pasadena, CA 91125, USA*
[3] *School of Physics and Astronomy, Monash University, Clayton VIC 3800, Australia*
[4] *OzGrav: The ARC Centre of Excellence for Gravitational Wave Discovery, Clayton VIC 3800, Australia*
[5] *Birmingham Institute for Gravitational Wave Astronomy and School of Physics and Astronomy,*
*University of Birmingham, Birmingham, B15 2TT, United Kingdom*

In General Relativity, the spacetimes of black holes have three fundamental properties: (i) they are the same, to lowest order in spin, as the metrics of stellar objects; (ii) they are independent of mass, when expressed in geometric units; and (iii) they are described by the Kerr metric. In this paper, we quantify the upper bounds on potential black-hole metric deviations imposed by observations of black-hole shadows and of binary black-hole inspirals in order to explore the current experimental limits on possible violations of the last two predictions. We find that both types of experiments provide correlated constraints on deviation parameters that are primarily in the $tt$-components of the spacetimes, when expressed in areal coordinates. We conclude that, currently, there is no evidence for a deviations from the Kerr metric across the 8 orders of magnitudes in masses and 16 orders in curvatures spanned by the two types of black holes. Moreover, because of the particular masses of black holes in the current sample of gravitational-wave sources, the correlations imposed by the two experiments are aligned and of similar magnitudes when expressed in terms of the far field, post-Newtonian predictions of the metrics. If a future coalescing black-hole binary with two low-mass (e.g., $\sim 3M_\odot$) components is discovered, the degeneracy between the deviation parameters can be broken by combining the inspiral constraints with those from the black-hole shadow measurements.

## I. INTRODUCTION

Over the past century, numerous predictions of the theory of General Relativity (GR) have been tested against multiple experiments and astrophysical observations [1]. Even though all these investigations aim to test the same theory, they nevertheless address different combinations of its ingredients, different aspects of its predictions, and are performed on widely different scales. A very stringent constraint imposed in one setting does not necessarily preclude the importance of testing different predictions of the theory in a different setting.

At one level, it is common to distinguish gravitational tests between those that test Einstein's equivalence principle and those that test the field equation [1]. The former search for violations of the weak equivalence principle, of the local Lorentz invariance, and of the local position invariance. The latter assume the validity of Einstein's equivalence principle and, therefore, that spacetime is endowed by a metric and that test particles and photons follow geodesics in this metric. These tests then use observations to map the metric of an object and test whether its parameters are consistent with GR predictions.

Among metric tests, there is a critical distinction between those that measure parameters of time-independent metrics and those that probe the dynamics of the theory. For example, all solar-system tests and many of the tests involving binary pulsars explore equilibrium metrics [2], whereas gravitational-wave tests with pulsar timing [3], direct gravitational-wave obser-

vations [4–6], and cosmological tests [7] explore dynamical metrics. A different distinction between metric tests separates those that involve vacuum metrics from the ones in which the coupling of matter with the gravitational field plays an important role. The latter category includes not only the cosmological tests, for which the feedback of matter to the gravitational field determines the evolution of structure formation, but also, e.g., tests of scalar-field gravity with pulsars, in which the presence of matter in the strong gravitational fields of the neutrons stars gives rise to dipole radiation (which has not been observed) [8, 9].

Even though the breadth of gravitational tests explores many aspects of the underlying theory and its predictions, it is usually impossible to assess the relative significance of any one particular upper limit on possible deviations. This is true because there is no compelling alternative to GR that is based on fundamental physics arguments and that leaves observable signatures at the scales of most experiments. In other words, without a plausible alternative to the theory to guide the tests and our thinking, it is not viable to ask how large a deviation one would expect from any test. Moreover, it is impossible to guess at what scale would such deviations become observable.

For these reasons, a wide variety of gravitational tests have been performed with objects and in settings that probe a vast range of scales: from sub-mm length scales to the size of the observable Universe, from the $GM/Rc^2 \sim 10^{-9}$ potentials [10] of terrestrial experiments to the order-unity potentials of black holes, and

from the $GM/R^3c^2 \sim 10^{-12}$ cm$^{-2}$ curvature scales of neutron stars to the $\sim 10^{-57}$ cm$^{-2}$ curvature scale of the cosmological constant [11].

The differences in the qualitative character between the gravitational tests, the observables used, and the tools employed hamper our ability to cross compare the resulting constraints of any deviations from GR. If a specific modification to the GR field equations is considered, then the complete theory can be used to make predictions for and be compared against all types of observations and astrophysical systems. However, translating empirical constraints on general modifications from one setting to another poses serious challenges. For example, it is very hard to ask in a theory-independent way how the results of cosmological tests affect the predictions for the gravitational-wave emission from coalescing neutron stars. Perhaps more importantly for the design of future experiments, it is often impossible to understand whether existing constraints on modifications from GR already preclude the detection of beyond-GR phenomena in a previously unexplored setting.

In recent years, a new set of gravitational tests has emerged that probe a previously unexplored regime: that of the near-field regions of astrophysical black holes. The LIGO/Virgo detection of gravitational waves from coalescing black holes and neutron stars has led to tests of the dynamics of GR with stellar-mass objects, during the inspiral and ringdown stages of the events [4–6][12]. The constraints from gravitational-wave tests are typically expressed in terms of upper limits on deviations from the GR predictions on a set of parametric post-Newtonian terms in the waveform expression. These terms amalgamate both potential deviations in the metrics of the individual black holes from the Kerr solution (when they are at large distances) as well as potential deviations in the strength, polarization, and angular distribution of the radiated gravitational waves, i.e., the dynamics of the theory.

Moving to the larger masses and smaller curvatures of the supermassive black holes in the centers of galaxies, the monitoring of the orbits of stars within a few thousand Schwarzschild radii from the black hole in the center of the Milky Way, Sgr A*, has led to application of two classical GR tests to this black-hole environment: the measurement of the gravitational redshift (which leads to a test of the equivalence principle) [13, 14] and the detection of orbital precession of the nearest star [15]. As in the case of the precession of Mercury in the Solar System, the constraints from the latter test are expressed in terms of the usual coefficients of the parametric post-Newtonian (PPN) framework [2]. Additional constraints can be imposed on equivalence principle violations via potential changes in the fine-structure constant [16] or on the Yuwaka strength and scale of a putative fifth force [17].

More recently, the Event Horizon Telescope (EHT) has detected the shadow of the black hole in the high-resolution image obtained from the center of the M87 galaxy [18]. Comparing the observed shadow size to that predicted for the mass of the black hole that was known *a priori* from stellar dynamics has led to constraints on the possible deviation of the black-hole spacetime from the Kerr metric [19]. These constraints were expressed in terms of upper limits on parameters of metrics that have been designed to be different from Kerr, while ensuring that no pathologies are present outside the horizons [20–23]. They were further translated into constraints on the post-Newtonian expansions of these metrics, in order to compare them with earlier, weak-field tests.

A common denominator among many of these avenues of testing GR with black holes or other stellar objects is the set of constraints they impose on the equilibrium metrics of objects with different compositions and masses. However, as discussed above, these constraints are expressed in different ways that are specific to each test because, e.g., they are merged with parameters that quantify the dynamics of the theory (as is the case of the gravitational-wave tests) or employ complexity that is necessary to avoid pathologies (as is the case of the shadow tests).

The aim of this series of papers is to combine all existing tests of metrics of astrophysical objects in order to test three important GR predictions for the metrics of black holes, i.e., that: (i) the metric of a black hole, expanded to first order in spin, is identical (when expressed in geometric units) to that of a slowly spinning star, i.e., it is the Schwarzschild spacetime with the first-order frame-dragging terms; (ii) all black holes, independent of mass or curvature, are described by the same metric; and (iii) the black-hole spacetime is described by the Kerr metric.

The approach we will follow here is to use the constraints that are imposed with each type of test and calculate their implications for the values of the various terms in a parametric post-Newtonian expansion of the equilibrium black-hole metrics. It is important to emphasize here that, in many cases involving black-hole tests, we will not be testing post-Newtonian expansions of the metrics. Instead, we will be constraining the parameters for regular, well behaved metrics that deviate from the GR predictions and then use these constraints to place bounds on deviations of the corresponding post-Newtonian parameters of these metrics.

Albeit not comprehensive, this approach allows for a comparison between the results of various tests and identifies the unique aspects of equilibrium metrics that each test is sensitive to. Moreover, this approach facilitates the comparison of the new constraints to those imposed by previous Solar System, pulsar, and cosmological tests, which we will address in forthcoming papers.

In this first paper, we will focus on the tests that involve the black-hole shadow and the inspiral phase of coalescing black holes and aim to address the latter two GR predictions discussed above. Section §II introduces the formalism and coordinate system for the parametric post-Newtonian equilibrium metric used throughout

this series of papers. Section §III presents the black-hole shadow constraints in terms of several parametric metrics that deviate from Kerr without introducing pathologies and translates them into the parametrization of the post-Newtonian expansion of the equilibrium metrics. Section §IV follows the same approach for tests that use LIGO/Virgo observations of gravitational waves emitted during the inspiral phase of compact binary coalescence. In Section §V, we discuss the key results and future prospects.

## II. PARAMETRIC METRICS OF ISOLATED STATIC OBJECTS

As discussed in Section §I, the goal of this work is to translate the constraints imposed by various gravitational tests into bounds on deviations from the GR solution for the equilibrium metric of an isolated object. This first paper will focus on the results of two types of observations: of the black-hole shadow images obtained with the EHT and of the gravitational waves from coalescing black holes detected with LIGO/Virgo.

The shadow of a Kerr black hole is highly circularly symmetric, up to near-extremal values of the spin, and has a diameter that depends very weakly on the spin of the black hole, for all observer inclinations. This is a consequence of a fortuitous near cancellation of the effects of frame dragging and of the quadrupole mass moment of the Kerr spacetime [24]. The extremely weak dependence of the shadow diameters on spin was also shown to be preserved in several metrics that are parametrically different from Kerr [25, 26]. Moreover, the constraints on deviations from the Kerr metric imposed by the diameter measurement of the shadow in M87 were shown to depend very weakly on the assumed spin [19].

In the case of the black hole binaries observed with LIGO/Virgo, the predicted inspiral waveforms do depend on the spins of the black holes, starting at the 1.5 Post-Newtonian (PN) order (see, e.g., Ref. [27]). However, the majority of coalescing black holes observed to date are consistent with having low or moderate spins, $\chi \lesssim$ few tenths [5, 28]. Because both sets of measurements provide very weak bounds on the spins of the black holes, the focus of this work will be on their implications for the equilibrium spacetimes of non-spinning, isolated objects.

In GR, Birkhoff's theorem ensures that the external spacetime of a spherically symmetric, isolated object is unique and is described by the Schwarzschild metric. The theorem is not necessarily valid in other metric theories of gravity; spherically symmetric isolated objects may be described by different metrics depending on the gravitational potential, the curvature, or even the nature of the object. For example, in scalar-tensor modifications to gravity with quadratic couplings [9] or with Chameleon screening [29, 30], the external spacetime of a stellar object depends on the density of matter in its interior. In
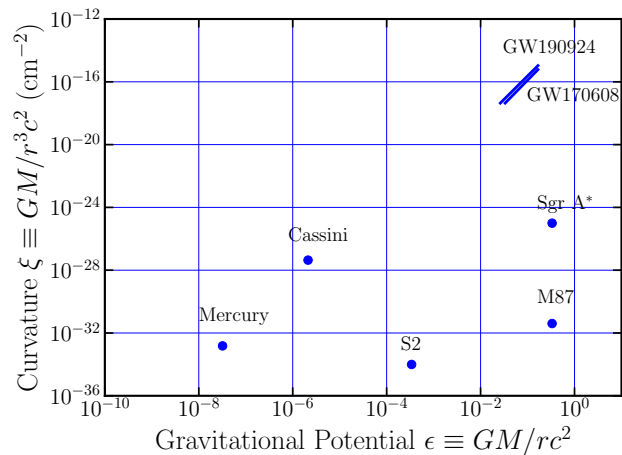


FIG. 1. Characteristic gravitational potential and curvature probed by two Solar-System tests (the perihelion precession of Mercury and the light deflection measured with Cassini) as well as the same quantities probed by the detection of gravitational waves from the LIGO/Virgo events GW170608 and GW190924, by the detection of periapsis precession in the S2 star around Sgr A*, and by the observation of the black-hole shadow in the M87 and Milky Way galaxies (after [11]).

the first class of theories, the modifications are important at the strong fields of neutron stars, whereas in the second class of theories, the modifications become significant in the low-density envelopes of giant stars. However, in all such theories, the unique external spacetime of a non-rotating black hole remains the Schwarzschild solution [31]. For this reason, the external spacetimes of objects of different composition, mass, or nature (i.e., stellar objects vs. black holes) may not be the same. Indeed, demonstrating conclusively that the external spacetimes of two different spherically symmetric static objects, when expressed in gravitational units, are not the same would serve as a direct evidence that GR will have to be modified.

The external spacetime of a spherically-symmetric compact object can be written in many forms and coordinate systems. In order to facilitate comparison with earlier constraints, especially those at weak gravitational fields, we will translate our results into predictions for a general post-Newtonian parametrization of the metric of a spherically symmetric object. In doing so, we will be implicitly assuming that the values of the various parameters depend on the mass, gravitational potential, gravitational curvature, or nature of the object under consideration.

For comparison with other GR tests, the gravitational potential

$$\epsilon \equiv \frac{GM}{rc^2} \qquad (1)$$

TABLE I. Fields Probed by Different Gravitational Tests

| Test | Mass $(M_\odot)$ | Distance (cm) | Potential $\epsilon$ | Curvature $\xi$ $(cm^{-2})$ |
|---|---|---|---|---|
| Cassini (Shapiro) | 1 | $7 \times 10^{10}$ | $2 \times 10^{-6}$ | $4 \times 10^{-28}$ |
| Mercury (perihelion) | 1 | $5 \times 10^{12}$ | $3 \times 10^{-8}$ | $2 \times 10^{-33}$ |
| GW170608 | 19 | $\sim (2-9) \times 10^{7}$ | 1/30-1/6 | $\sim 5 \times 10^{-18} - 6 \times 10^{-16}$ |
| GW190924_021846 | 14 | $\sim (1-8) \times 10^{7}$ | 1/38-1/6 | $\sim 5 \times 10^{-18} - 1 \times 10^{-15}$ |
| S2 Star (periapsis) | $4 \times 10^{6}$ | $2 \times 10^{15}$ | $3 \times 10^{-4}$ | $1 \times 10^{-34}$ |
| Sgr A* (shadow) | $4 \times 10^{6}$ | $2 \times 10^{12}$ | 1/3 | $9 \times 10^{-26}$ |
| M87 (shadow) | $6.5 \times 10^{9}$ | $3 \times 10^{15}$ | 1/3 | $4 \times 10^{-32}$ |

and curvature

$$\xi \equiv \frac{GM}{r^3 c^2} \qquad (2)$$

probed by various gravity tests are shown in Table I and Fig. 1. These include the black-hole shadow tests of M87 and, in the near future, of Sgr A* (evaluated at $r = 3GM/c^2$), the inspiral test of GW170608 and GW190924 (evaluated at the range of separations that are detectable by the LIGO/VIRGO detectors, see below), the periapsis precession test with the S2 star around Sgr A* [15] (evaluated at the periapsis distance), and for two classical Solar-System tests with Cassini and the perihelion precession of Mercury [2]

### A. Parametric Post-Newtonian Metrics for non-Spinning Objects

The traditional GR tests in the Solar System with the Parametric Post-Newtonian (PPN) framework have been performed with a metric written in terms of isotropic coordinates [1], i.e.,

$$ds^2 = g_{tt,I}dt^2 + g_{rr,I}\left(dr_I^2 + r_I^2 d\Omega\right) , \qquad (3)$$

where the subscript $I$ denotes that the coordinates are isotropic.

Following the PPN approach, and denoting the traditional 1PN parameters as $\beta_1$ and $\gamma_1$, the non-zero metric components in isotropic coordinates become

$$g_{tt,I} = -1 + \frac{2}{r_I} - \frac{2\beta_1}{r_I^2} + \frac{3\beta_2}{2r_I^3} + \mathcal{O}(r^{-4})$$

$$g_{rr,I} = 1 + \frac{2\gamma_1}{r_I} + \frac{3\gamma_2}{2r_I^2} + \mathcal{O}(r^{-3}) , \qquad (4)$$

where $G = c = M = 1$. In GR, the values of all PPN parameters for the Schwarzschild spacetime, as defined here, are equal to unity.

In contrast, testing GR with black-hole shadows or inspiral waveforms is more naturally performed using parametric metrics in areal coordinates, i.e.,

$$ds^2 = g_{tt,S}dt^2 + g_{rr,S}dr_S^2 + r_S^2 d\Omega , \qquad (5)$$

where the subscript $S$ denotes the fact that the area of a closed surface at coordinate $r_S$ is always equal to $4\pi r_S^2$. This is because it has been recently shown that, in areal coordinates, the diameter of the shadow of a non-spinning black hole and the inspiral waveform during a coalesence event depend only on the $tt-$components of the metrics [19, 32].

Converting between isotropic and areal radial coordinates is, nevertheless, trivial using the transformation

$$r_I = r_S - \gamma_1 + \frac{(2\gamma_1^2 - 3\gamma_2)}{4r_S} + \mathcal{O}(r_S^{-2}) . \qquad (6)$$

For the parametrization shown in Eq. (4), the result is

$$\begin{aligned} g_{tt,S} = &-1 + \frac{2}{r_S} - \frac{2(\beta_1 - \gamma_1)}{r_S^2} + \frac{2}{r_S^3}\left\{\frac{1}{4}\left[2\left(\gamma_1^2 - 1\right)\right.\right. \\ &\left.-8(\beta_1\gamma_1 - 1) + 3(\beta_2 - 1) + 3(\gamma_2 - 1)\right]\} \\ &+\mathcal{O}(r^{-4}) \end{aligned} \qquad (7)$$

and

$$g_{rr,S} = 1 + \frac{2\gamma_1}{r_S} + \frac{\gamma_1^2 + 3\gamma_2}{r_S^2} + \mathcal{O}(r^{-3}) . \qquad (8)$$

If, as in Ref. [19], the various PN parameters were introduced instead in the expansion of the metric components written in areal coordinates as

$$g_{tt,S} = -1 + \frac{2}{r_S} + 2\sum_{i=1}^{N}(-1)^i \frac{\zeta_i}{r_S^{i+1}} , \qquad (9)$$

then the coefficients of the various order become

$$\zeta_1 = \beta_1 - \gamma_1$$

$$\begin{aligned} \zeta_2 = \frac{1}{4}&\left[2\left(\gamma_1^2 - 1\right) - 8(\beta_1\gamma_1 - 1)\right. \\ &\left.+3(\beta_2 - 1) + 3(\gamma_2 - 1)\right] , \end{aligned} \qquad (10)$$

etc, when expressed in terms of the present parametrization.

In the Solar System, the PPN parameter $\gamma_1$ is constrained via measurements of the deflection of light and of the Shapiro delay for signals that graze the solar surface [2]. The most stringent limit to date has been achieved with the Cassini mission and is $|\gamma_1 - 1| <$

$2.3 \times 10^{-5}$. On the other hand, the PPN parameter $\beta_1$ is constrained, in combination with $\gamma_1$, via measurements of the perihelion precession of Mercury and of the Nordtvedt effect in Lunar Ranging [2]. The most stringent limit to date is $|\beta_1 - 1| < 8 \times 10^{-5}$.

Albeit proven to be very effective [33], the expansion of the metric in a series over inverse powers of the coordinate radius is, of course, formally valid when the series converges. This is expected to be true, as long as there are no pathologies in the metric down to the radius probed by a particular test and the expected deviations from GR increase in strength as the radius from the central object decreases (i.e., not when searching for Yukawa-type corrections). However, even if formally converging, interpreting different astrophysical settings requires a careful assessment of the truncation errors in comparison to the measurement uncertainties. We will address these convergence issues for the black-hole shadow tests in §III and for the inspiral tests in §IV.

### B. Expected Magnitude of Deviations

Because of the lack of compelling alternatives to GR that arise from fundamental physics arguments and lead to astrophysically relevant effects, there is no first-principles approach to estimate the expected magnitudes of corrections to the various PPN parameters. This is especially true for tests with black holes since the Kerr metric is a solution to many simple modifications of the Einstein field equations supplemented with an additional field characterized by constant coupling coefficients [31]. Obtaining non-Kerr black-hole solutions for modifications to the field equations arising from, e.g., the addition of quadratic terms in the Riemann tensor, requires couplings that are described by dynamical fields. One example of such non-Kerr black holes occur in the so-called Einstein-Dilaton-Gauss-Bonnet [EDGB] theories, were originally obtained in Ref. [34], and were further studied in [35–39].

One might naively expect deviations from the Schwarzschild/Kerr metrics only in black holes with masses that are comparable to the scale at which the Einstein-Hilbert action of GR is modified. It is straightforward to show, however, that this is not necessarily the case. As a proof of principle, we will use the non-spinning black-hole solution for EDGB gravity, when the coupling coefficients are linear functions of the dynamical fields [35]. The Lagrangian action of this theory is

$$S = \int d^4x \sqrt{-g} \left\{ \kappa R + a_1 \theta R^2 + a_2 \theta R_{ab} R^{ab} \right.$$
$$+ a_3 \theta R_{abcd} R^{abcd} + a_4 \theta R^*_{abcd} R^{abcd}$$
$$\left. - \frac{\beta}{2} \left[ \nabla_a \theta \nabla^a \theta + 2V(\theta) \right] \right\} , \qquad (11)$$

where $R$, $R_{ab}$, and $R_{abcd}$ are the Ricci scalar, Ricci tensor, and Riemann tensor, respectively, $V(\theta) \simeq (1/2) m_\theta \theta^2$

is the potential of the dynamical scalar field, $a_1, ..., a_4$ are the coupling coefficients of the Gauss-Bonnet terms, and $\beta$ is the coupling coefficient of the dynamical scalar field. Hereafter, we set the latter to unity, as it can be reabsorbed in a redefinition of the scalar field [36].

The presence of the dynamical field forces the non-spinning black-hole solution to deviate from Schwarzschild. When expressed in areal coordinates, the two non-trivial components of the metric for this solution are [35]

$$g_{tt,S} = -1 + \frac{2}{r_S} - \frac{\lambda}{3r_S^3} + \mathcal{O}(r^{-4}) \qquad (12)$$

and

$$g_{rr,S} = 1 + \frac{2}{r_S} + \frac{4 - \lambda}{r_S^2} + \mathcal{O}(r^{-3}) , \qquad (13)$$

where $\lambda = a_3^2/(\kappa M^4)$ and we have explicitly shown the dependence on the mass of the black hole $M$ that can be measured by, e.g., monitoring the orbits of stars at large distances. Comparing this solution to the PPN expansions (7)-(8) leads, for this theory, to

$$\beta_1 = 1$$
$$\gamma_1 = 1$$
$$\beta_2 = \frac{1}{9} \left( 9 + \frac{a_3^2}{\kappa M^4} \right)$$
$$\gamma_2 = \frac{1}{3} \left( 3 - \frac{a_3^2}{\kappa M^4} \right) . \qquad (14)$$

In other words, this particular modification to GR leads to non-spinning black-hole spacetimes that have the same 1PN expansion as the Kerr metric but differ only at the 2PN and higher orders. Moreover, the deviation depends only on the dimensionless ratio $\lambda$.

In principle, the $a_3$ coupling of the quadratic term can be substantially different than the mass of the black hole. However, as long as there is substantial scale separation between the couplings of the quadratic field and the Planck mass, then the metrics of black holes with masses $M \sim (a_3^2/\kappa)^{1/4}$ will show order unity deviations from the Schwarzschild solution at the 2PN (and higher) orders. This argument was discussed explicitly in Ref. [40], where an expression was derived for the effective cut-off scale $\Lambda$ of the EDGB gravity. Deviations from GR reach order unity for a black of mass $M$, when te cut-off scale is

$$\Lambda \sim 7 \left( \frac{M}{M_\odot} \right)^{-2/5} \text{TeV} . \qquad (15)$$

It is important to emphasize here that, among all gravitational tests with black holes, the ones with the smallest masses will lead to the tightest constraints in this theory. This would lead to the conclusion that the stellar-mass black holes involved in gravitational-wave tests generate the most stringent limits on any possible deviations. This

conclusion, however, is an artifact of the assumption intrinsic to this modification that the various couplings are proportional to the dynamical field $\theta$. Different (and perhaps more complex) coupling functions (see, e.g., [34, 38]) could easily reverse this trend, as is the case with many modifications of gravity that involve various screening mechanisms (see, e.g., [29]).

Finally, even if corrections to the fundamental theory appear at scales that are very different than the masses of astrophysical objects, the presence of horizons surrounding black holes might give rise to horizon-scale classical metric perturbations [41]. The consequence of such perturbations would be the presence of time-variable, order unity stochastic deviations of black-hole metric parameters, with observable effects both in gravitational-wave emission [42] and in the images of black-hole shadows [43]. If the timescale of variability of such deviations is longer than the $\sim 10$ h it takes for the EHT to obtain a single snapshot image, then these deviation parameters will be frozen to some arbitrary combination of values.

## III.  BLACK HOLE SHADOW TESTS

The EHT generated a high-resolution image of the center of the M87 galaxy [18] that is characterized by a deep brightness depression surrounded by a ring of emission. This image has been interpreted as the shadow of the central supermassive black hole, cast on the emission from the surrounding plasma. When measured from the reconstructed images, the fractional width of the ring of emission was constrained to be comparable to the nominal resolution of the array [44]. On the other hand, when inferred by fitting phenomenological emission models directly to the interferometric visibility data, the fractional width of the ring was constrained to $\lesssim 0.2$, at least for one of the four days of observations [18].

This property of the emission ring, i.e., that it is very narrow and is truncated at the black-hole shadow, makes it possible to use it as a proxy for measuring the size of the shadow itself. The bias between the diameter of the brightest emission in the ring and that of the shadow was estimated using a large suite of synthetic EHT data based on General Relativistic MagnetoHydro-Dynamic (GRMHD) simulations and the uncertainty in this bias was found to be comparable to the measured size of the ring itself, as expected. Finally, the inferred size of the shadow was found to be within $\sim 17\%$ of the value predicted by the Kerr metric [18] using the *a priori* known mass-to-distance ratio of the M87 black hole that was based on observations of the motions of stars in its vicinity [45]. As proposed in Ref. [46], this constitutes a null-hypothesis test of the various assumptions that enter this inference, i.e., that the brightness depression in the image is indeed the black-hole shadow, that the analysis of stellar dynamics provides an accurate measurement of the black-hole mass [47], and that the black-hole spacetime is described by the Kerr metric.

Ref. [19] used the inferred size of the M87 shadow to perform a *bona fide* metric test: how much could one deform the Kerr metric and still be consistent with the measurement? Using a variety of parametrically deformed metrics, they showed that, as in the case of Kerr, the shadow size depends very weakly on the spin of the spacetime (defined as the specific angular momentum measured at infinity, divided by the mass of the black hole). Moreover, they demonstrated that, for spherically symmetric spacetimes, the radius of the shadow measured by a distant observer depends only on the $tt-$component of the metric, when expressed in areal coordinates, i.e.,

$$r_{\mathrm{sh}} = \frac{r_{\mathrm{ph}}}{\sqrt{-g_{tt}(r_{\mathrm{ph}})}} \;, \qquad (16)$$

where

$$r_{\mathrm{ph}} = \sqrt{-g_{tt}(r_{\mathrm{ph}})} \left( \frac{d\sqrt{-g_{tt}}}{dr}\bigg|_{r_{\mathrm{ph}}} \right)^{-1}$$

$$= 2g_{tt}(r_{\mathrm{ph}}) \left( \frac{dg_{tt}}{dr}\bigg|_{r_{\mathrm{ph}}} \right)^{-1} \qquad (17)$$

is the coordinate radius of the photon orbit. Finally, they calculated the constraints imposed by the existing shadow-size measurements on deformed spacetimes that show no deviations at the first PN order (as is the case for, e.g., the black-hole spacetimes in the modified gravity theory discussed in §2) and are, therefore, fully consistent with all Solar-System tests. Because, as discussed above, there is no guarantee in a general modified gravity theory that the Solar-System constraints are applicable to black-hole spacetimes, this was offered as a proof of principle that shadow observations with current capabilities can be used to impose new and tighter constraints on potential deviations from the Kerr metric for supermassive black holes.

In this paper, we extend the work of Ref. [19] to explore general forms of deviations from Kerr within multiple phenomenological descriptions without imposing the Solar-System bounds and translate the resulting bounds into limits on the PPN parameters of equilibrium black-hole metrics.

### A.  Deformed Metrics without Pathologies

The outline of a black-hole shadow is the locus of the photon trajectories on the screen of a distant observer that, when traced backwards, become tangent to the surfaces of spherical photon orbits hovering just above the black-hole horizon [48]. The coordinate radius of the photon orbit for a Schwarzschild spacetime is at $r_S = 3M$ and is reduced to $r_S = M$ (i.e., the coordinate radius of the horizon) for a prograde photon orbit around a maximally spinning black hole. Because the Kerr spacetime is regular everywhere outside the horizon and the photon orbits always lie outside the latter, Kerr black holes of

all spins are characterized by shadows that can be well defined, at least mathematically (the same is not true for super-spinning Kerr black holes [49]).

However, this is not true in general for deformed Kerr spacetimes. Introducing any naive parametric deformation violates, by construction, the no-hair theorem, which states that the only asymptotically flat vacuum spacetime that is Ricci flat, is free of singularities outside the horizon, and is free of closed time-like loops is the one described by the Kerr metric (we do not consider here charged black holes) [50]. Indeed, early attempts to deform the Kerr metric while keeping it Ricci flat led to spacetimes with significant pathologies [22]. Calculations of black-hole shadows with these deformed spacetimes required excising in an *ad hoc* manner the regions with pathologies and limited the solutions to slowly spinning black holes, such that the radii of photon orbits remained outside the pathologies [24, 51]. Moreover, the presence of these pathologies effectively precluded any GRMHD simulations of accretion in such spacetimes.

The only way to deform the Kerr metric while removing any pathologies outside its horizon (at least for a broad range of deformation parameters) is to ignore explicitly the requirement that the metric is Ricci flat. In recent years, this has led to a number of parametrically deformed metrics that are free of pathologies but allow for deformations to be dialed in with different phenomenological parameters: the Johannsen-Psaltis metric (hereafter JP) [20, 22] and its extensions by Cardoso, Pani, & Rico [52] and Carson & Yagi [53], the Modified Gravity Bumpy Kerr metric of Vigeland, Yunes, and Stein [21] (hereafter MGBK), etc. In a different approach, a general metric can be written in terms of polynomial or rational functions with free coefficients such that, when a particular discrete set of coefficients is chosen, the metric approximates non-Kerr solutions to various modified-gravity field equations; if the non-Kerr solution is free of pathologies, so is the polynomial or rational expansion [54]. This is the approach followed in the Rezzolla-Zhidenko metric (hereafter RZ) [55, 56], which has found some use in numerical explorations of black-hole shadows from known non-Kerr metrics [57, 58].

### B. Post-Newtonian Expansions of Deformed Metrics

When written in areal coordinates, the $tt$-component of the JP metric of a non-spinning compact object is [22]

$$g_{tt}^{\text{JP}} = -\left(1 - \frac{2}{r_S}\right)\left(1 + \sum_{i=2}^{\infty} \frac{\alpha_{1i}}{r_S^i}\right)^{-2}, \quad (18)$$

where $a_{1i}$ is an infinite sequence of deformation parameters, which are equal to zero for the Schwarzschild metric. In Ref. [22], the coefficient $\alpha_{12}$ was set to zero in order to force $\beta_1 = 1$. In this paper, for reasons discussed above, we will allow potential deviations even at the 1PN order.

Written in terms of the PN parametrization discussed in §2, the result is

$$\zeta_1^{\text{JP}} = -\alpha_{12}$$
$$\zeta_2^{\text{JP}} = \alpha_{13} - 2\alpha_{12}$$
$$\zeta_3^{\text{JP}} = -\alpha_{14} + 2\alpha_{13} + \frac{3}{2}\alpha_{12}^2, \quad (19)$$

etc.

Written in areal coordinates, the $tt-$component of the MGBK metric for a non-spinning compact object is [21]

$$g_{tt}^{\text{MGBK}} = -\left(1 - \frac{2}{r_S}\right)\left[1 - \gamma_1(r_S) - 2\gamma_4(r_S)\left(1 - \frac{2}{r_S}\right)\right]. \quad (20)$$

This expansion appears to terminate at the $r_S^{-2}$ order because the deformed Kerr metric was designed such that it is characterized by an approximate Killing tensor of that same order. However, the two functions $\gamma_1(r)$ and $\gamma_4(r)$ are arbitrary and can be expanded in series, e.g.,

$$\gamma_{\text{A}} = \sum_{n=2}^{\infty} \frac{\gamma_{\text{A},n}}{r_S^n}, \quad (21)$$

where A = 1 or 4 and $\gamma_{\text{A},n}$ are infinite sequences of dimensionless deformation parameters and the $n < 2$ terms are equal to zero in order for the metric to be asymptotically flat and have the correct Newtonian limit. Keeping only a few lower-order coefficients, the metric becomes

$$g_{tt}^{MGBK} = -1 + \frac{2}{r_S} + \frac{\gamma_{1,2} + 2\gamma_{4,2}}{r_S^2}$$
$$+ \frac{-2\gamma_{1,2} + \gamma_{1,3} - 8\gamma_{4,2} + 2\gamma_{4,3}}{r_S^3} + \mathcal{O}(r^{-4}), \quad (22)$$

such that the PN parameters in areal coordinates are

$$\zeta_1^{\text{MGBK}} = -\gamma_{1,2} - 2\gamma_{4,2}$$
$$\zeta_2^{\text{MGBK}} = -\gamma_{1,2} + \frac{1}{2}\gamma_{1,3} - 4\gamma_{4,2} + \gamma_{4,3} \quad (23)$$

etc.

Finally, the $tt$-component of the RZ metric for a non-spinning compact object, written in areal coordinates, is [55]

$$g_{tt}^{\text{RZ}} = -\left(1 - \frac{r_0}{r_S}\right)\left[1 - \epsilon(1 - x) \right.$$
$$\left. + (a_0 - \epsilon)(1 - x)^2 + \tilde{A}(x)(1 - x)^3\right], \quad (24)$$

where

$$x \equiv 1 - \frac{r_0}{r_S}, \quad (25)$$

$$\tilde{A}(x) = \frac{a_1}{1 + \frac{a_2 x}{1 + \frac{a_3 x}{\cdots}}}, \quad (26)$$

$r_0$ is the coordinate radius of the infinite redshift surface (heuristically identified with the horizon, if no pathologies exist at larger radii), and $\epsilon$, $a_1$, $a_2$, ... are a sequence of deformation parameters. Writing all radii in terms of the mass of the black hole, as measured at infinity, fixes one of the parameters to

$$\epsilon = -\left(1 - \frac{2}{r_0}\right) . \qquad (27)$$

Under these assumptions, the PN parameters of the RZ metric become

$$\zeta_1^{\mathrm{RZ}} = \frac{1}{2}a_0 r_0^2$$
$$\zeta_2^{\mathrm{RZ}} = \frac{1}{2}\left[1 - \frac{2}{r_0} + a_0 - \frac{a_1}{1 + \frac{a_2}{1 + \frac{a_3}{\cdots}}}\right] r_0^3 , \qquad (28)$$

etc.

### C. Photon Orbits and Shadows of Deformed Metrics

The three different parametrizations (18), (20), and (24) for the deformed metric of a non-spinning object share a common characteristic: they all involve expansions in power-series after the factor $(1 - 2/r_S)$ has been removed from the $tt$-component of the metric. This ensures that the metric has a surface of infinite redshift (a "horizon") for a large range of deformation parameters, which helps in hiding the pathologies introduced by the deformations from the observable universe. However, because of this factoring, a single PN parameter $\zeta_i$ in areal coordinates corresponds to either a finite combination (for the JP and MGBK metrics) or an infinite complex function (for the RZ metric) of the deformation parameters of each metric. Nevertheless, in each case, there is a trivial transformation between the PN parameters $\zeta_i$ and the deformation parameters of the corresponding metric. For this reason, we will use the JP parametrization given by Eq. (19) to place constraints on plausible metric deviations from the measurement of the black-hole shadow diameter. One can then use transformations (19), (23), or (28) to convert them into constraints on the particular deviation parameters for the other metric parametrizations.

Using Eq. (18) with Eq. (17) yields an expression for the coordinate radius of photon orbits, i.e.,

$$r_{\mathrm{ph}}^{\mathrm{JP}} = 3 + \frac{2}{9}\alpha_{12} + \frac{1}{9}\alpha_{13} + \frac{4}{81}\alpha_{14} + ... , \qquad (29)$$

where only linear terms in the deformation parameters have been retained. This is consistent with the power-series expansion of the metric shown in equation (18). Moreover, terms that involve higher powers in the deformation parameters are negligible compared to the corrections introduced by the black-hole spin, which are not measurable with current data and are neglected here.

For example, the first non-linear term is $-2\alpha_{12}^2/243 \simeq 0.008\alpha_{12}$. Evaluating the $tt$-component of the metric at the radius of the photon orbit gives

$$g_{tt}(r_{\mathrm{ph}}) = -\frac{1}{3} + \frac{2}{81}\alpha_{12} + \frac{10}{729}\alpha_{14} + ... , \qquad (30)$$

which remains regular unless the deformation parameters take extremely large values. Note that the $\alpha_{13}$ contribution to this expression is vanishing at linear order.

Expressed in terms of the PN parameters in areal coordinates, the radius of the photon orbit of the JP metric becomes

$$r_{\mathrm{ph}}^{\mathrm{JP}} = 3 - \frac{52}{81}\zeta_1 + \frac{17}{81}\zeta_2 - \frac{4}{81}\zeta_3 + ... \qquad (31)$$

Finally, inserting equation (29) into the general expression (16) for the shadow radius gives

$$r_{\mathrm{sh}}^{\mathrm{JP}} = 3\sqrt{3}\left(1 + \frac{1}{9}\alpha_{12} + \frac{1}{27}\alpha_{13} + \frac{1}{81}\alpha_{14} + ....\right) . \qquad (32)$$

Expressed in terms of the PN parameters in areal coordinates, the shadow radius becomes

$$r_{\mathrm{sh}}^{\mathrm{JP}} = 3\sqrt{3}\left(1 - \frac{19}{81}\zeta_1 + \frac{5}{81}\zeta_2 - \frac{1}{81}\zeta_3 + ...\right) . \qquad (33)$$

It is important to emphasize that the shadow tests performed here and in Ref. [19] do not employ a parametric post-Newtonian metric. Indeed, the coefficients of the various terms in Eq. (33) are different from the expression obtained if calculating the size of a black-hole shadow using directly the PN metric (9). Instead, the shadow tests are performed using metrics that remain regular all the way down to their horizons and constraints are imposed on the parameters of these regular metrics. However, in order to compare the shadow tests to those of earlier, weak-field tests, these constraints are then translated into equivalent constraints on the post-Newtonian parameters of these metrics.

The convergence properties of the series (32) and (33) are difficult to explore formally. Nevertheless, the coefficients of the terms in Eq. (32) are decreasing by successive powers of 3 and those in Eq. (33) by successive powers of $\sim 4-5$. This is expected given that they are all the result of power series expansions in $1/r_S$ and the radius of the photon orbit, which determines primarily the size of the shadow, is $r_S = 3$ for the Schwarzschild metric. In order for higher-order terms to have a significant impact on the size of the shadow, they need to be successively increasing by corresponding powers of $\sim 3-5$.

### D. Metric Constraints from the Measured Size of the M87 Shadow

Our goal is to place constraints on possible deviations from the Kerr metric using the inferred size of the black-hole shadow in M87, given the mass of the black hole

measured at large distances, in the Newtonian limit, as prior information. However, as discussed earlier, the EHT imaging observations of M87 do not directly measure the size of the shadow but rather the size of the bright ring of emission that surrounds it.

In order to connect the two, we use the model described in Ref. [18] that incorporates a number of steps in order to convert the prior mass measurement to a prediction of the size of the bright ring. In particular, we quantify the prior in terms of the angular size in the sky of one gravitational radius for a black hole of mass $M$ at a distance $D$, i.e., $\theta_{\rm g} \equiv GM/(c^2 D)$. We introduce the correction factor $a$ between the angular diameter $\hat{d}_{\rm m}$ of the peak emission and the angular diameter of the shadow $2r_{\rm sh}$ such that

$$\hat{d}_{\rm m} = 2a r_{\rm sh} \ . \tag{34}$$

Finally, we calculate the angular size of the shadow for a given prior mass-to-distance ratio for the black hole, while allowing for a possible fractional deviation $\delta$ in the prediction of the Kerr metric such that

$$\hat{d}_{\rm m} = 2a(1+\delta)r_{\rm sh} = 2a(1+\delta)3\sqrt{3}\left(\frac{GM}{Dc^2}\right) \ . \tag{35}$$

Even though $\delta$ represents any possible deviation, for the particular model discussed above, it is equivalent to

$$\delta = \frac{19}{81}\left(-\zeta_1 + \frac{5}{19}\zeta_2 - \frac{1}{19}\zeta_3 + ...\right) \ . \tag{36}$$

Eq. (35) allows us to infer or constrain the deviation $\delta$ from the Kerr metric predictions, given the prior information on the mass-to-distance ratio $M/D$ of the black hole, the measurement of the diameter $\hat{d}_{\rm m}$ of the bright ring of emission with the EHT, and a model for the correction factor $a$.

The prior $P(\theta_{\rm g})$ on the angular size in the sky of one gravitational radius for the black hole in M87 has been measured using stellar dynamics in Ref. [45] and quantified in Ref. [44]. Here we use the full numerical information on the prior, which peaks at $\theta_{\rm g,0} \sim 3.62$ $\mu$as and is asymmetric (it can be represented approximately as $\theta_{\rm g} = 3.62^{+0.60}_{-0.34}$ $\mu$as).

The correction factor $a$ has been calibrated using $\sim 100$ synthetic images from accretion-flow simulations that span different black-hole spins, magnetic field configurations in the accretion flows, and models for the plasma physics. Even though the particular simulations used in the calibration were performed for the Kerr metric, the primary source of the error budget arises from the thermodynamic properties of the plasma in the inner accretion flow; the properties of the metric enter predominantly in the imprint of gravitational lensing on the image, which is the size and shape of the black-hole shadow. We use here a Gaussian distribution for $a$ with a mean value $6\sqrt{3}a_0 = 11.35$ and a standard deviation of $\sigma_a/a_0 = 11.4\%$ (see Table 4 and Fig. 26 of
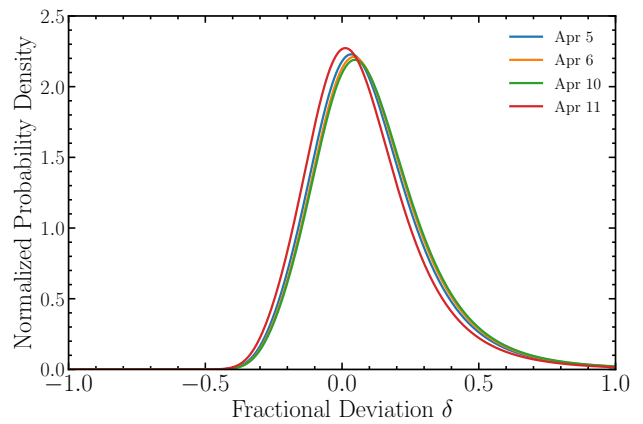


FIG. 2. The posterior over the fractional deviation $\delta$ between the size of the shadow predicted using the Kerr metric for the M87 black hole and the size inferred for the 4 days of the 2017 EHT observations. These posteriors are based on the stellar dynamics measurement of the mass-to-distance ratio for the black hole taken as a prior and incorporate the uncertainty introduced by the difference between the size of the black-hole shadow and that of the bright ring of emission, as well as the uncertainty due to the unknown black-hole spin and observer inclination. For the Apr 5 observations, the result is $\delta = 0.03^{+0.20}_{-0.16}$ (68% credible level) and is approximately the same for the remaining days.

Ref. [18]). Note that the width of this distribution incorporates the small ($\pm 4\%$) spread in the predicted shadow size due to the unknown black-hole spin and observer inclination [24].

The EHT observed the black hole in the center of M87 for four days across the span of a week in April 2017. Interferometric data were collected in two frequency bands (HI and LO) and images were generated for each day and each frequency band separately [44]. Moreover, two different geometric model images as well as images generated from GRMHD simulations were fit directly to the interferometric visibility data, in order to infer the geometric parameters of the emission rings [18]. Both the image-domain and visibility-domain analyses led inferred image sizes that are consistent with each other and across the different days of observation and different frequency bands. Here, we use the posteriors for the measurement of the diameter of the ring of emission inferred using the `xs-ringgauss` model for the combined HI+LO datasets, in all four days of observations. We denote the most likely value of each measurement by $\hat{d}$ and its uncertainty by $\sigma_{\rm d}$ (see Table 3 of Ref. [18]).

The posterior over the deviation parameter $\delta$ is given by

$$P(\delta|\hat{d}, \sigma_d) = C \int d\delta \int da \int d\theta_{\rm g} \ \mathcal{L}[\hat{d}, \sigma_d|\theta_{\rm g}, a, \delta] \\ \times P(\delta)P(\theta_{\rm g})P(a) \ , \tag{37}$$

where $C$ is an appropriate normalizaton constant and

$\mathcal{L}[\hat{d}, \sigma_d | \theta_g, a, \delta]$ is the likelihood of measuring a ring of size $\hat{d}$ given the model parameters $\theta_g$, $a$, and $\delta$. Hereafter, we will assume a flat prior in the fractional deviation $\delta$, with limits that are much larger than unity. Assuming a Gaussian distribution for both the likelihood function and the model correction factor $a$, we perform two of the integrals analytically such that

$$P(\delta|\hat{d}, \sigma_d) = \int d\theta_g P(\theta_g) \frac{1}{\sqrt{2\pi}\Sigma}$$
$$\exp\left[-\frac{108 a_0^2(1+\delta)^2\theta_g^2 - 12\sqrt{3}a_0(1+\delta)\theta_g\hat{d} + \hat{d}^2}{2\Sigma^2}\right] (38)$$

where

$$\Sigma = \sqrt{108(1+\delta)^2\theta_g^2\sigma_a^2 + \sigma_d^2} \,. \qquad (39)$$

After folding in the prior over $\theta_g$ and performing the last integral numerically, we obtain, for each of the four days of observations, the posteriors shown in Fig. 2.

For the Apr 5 2017 observations, the result is $\delta = 0.03^{+0.20}_{-0.16}$ (68% credible level) and is approximately the same for the remaining days. This implies that the limit on the deviation parameters becomes

$$-0.55 \leq -\zeta_1 + \frac{5}{19}\zeta_2 - \frac{1}{19}\zeta_3 + ... \leq 0.98 \,. \qquad (40)$$

It is possible to use inequalities (40) in order to constrain any individual parameter $\zeta_i$, allowing for only that parameter to attain a non-zero value, as was done, e.g., in Ref. [19]. However, because the shadow for a non-spinning compact object is circularly symmetric, its image provides only a single data point: its radius. For a general parametric extension of the metric with an infinite sequence of deviation parameters, it is evident from the above inequality that this single data point can only constrain the linear combination (40) of the infinite series of parameters. Because of this complete degeneracy, it is impossible to quantify credible levels for each of the deviation parameters separately by marginalizing over the remaining parameters.

Nevertheless, the coefficients of each parameter in the sum of Eq. (40) appear to be decreasing by a factor of $\sim 4 - 5$ between successive orders. As a result, if the deviation parameters are comparable to each other or are also decreasing with increasing order, then the series will be converging within a very small number of terms. In fact, even if the 3PN parameter $\zeta_3$ is of order unity, the resulting correction to the predicted shadow size will be at the $\sim 1\%$ level, i.e., smaller than the measurement uncertainty of $\lesssim 20\%$. Assuming that only the first two PN parameters, $\zeta_1$ and $\zeta_2$, provide contributions that are measurable with the current observations, inequality (40) results in the correlated upper limits shown in Fig. 3.
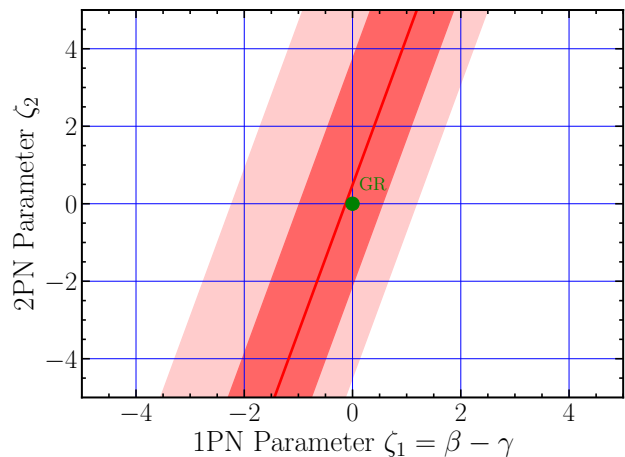


FIG. 3. Correlated bounds on the 1PN and 2PN parameters $\zeta_1$ and $\zeta_2$ imposed by the measurement of the size of the black-hole shadow in M87. The bounds were calculated with the JP parametrization of deformations from the Kerr metric and the shaded regions show the regions of the parameter space with deviation $\delta$ that is within the 68% and 95% credible intervals for the April 5 observations. The bounds assume that all higher-order deviation parameters provide negligible corrections to the predicted shadow size, which is true when $|\zeta_i| \lesssim 1$ for $i \geq 3$.

## IV. BLACK HOLE INSPIRAL TESTS

The LIGO/Virgo detectors [59, 60] have observed gravitational waves emitted during the inspiral, coalescence, and ringdown phases of several black-hole binaries and used them to place constraints on potential deviations from various GR predictions for black-hole spacetimes [4–6]. The tests performed to date can be broadly categorized into three large groups: (i) those involving the inspiral phase, when the coalescing black holes are at distances larger than their horizons; (ii) those involving the ringdown phase of the remnant black hole that is shedding its short-lived gravitational hair, and (iii) the tests involving the polarization and propagation of gravitational waves from the binary to the Earth. The various categories are, of course, not independent from each other and substantial information can be obtained by exploring, e.g., whether the black-hole parameters inferred from the inspiral phase of an event with GR waveforms are consistent with those inferred from the ringdown phase.

In this paper, we will focus only on the tests involving the inspiral phase for three reasons. First, the majority of the individual black holes coalescing in the various detected events appear to have small spins (see [28] for a detailed discussion of the spin distribution and the evidence for non-zero spins in some of the observed systems), in contrast to the remnant black holes ringing down, for which the inferred spins are $\sim 0.7$. Second, during the inspiral phase, the coalescing black holes are at separations larger than the effective radius of the in-
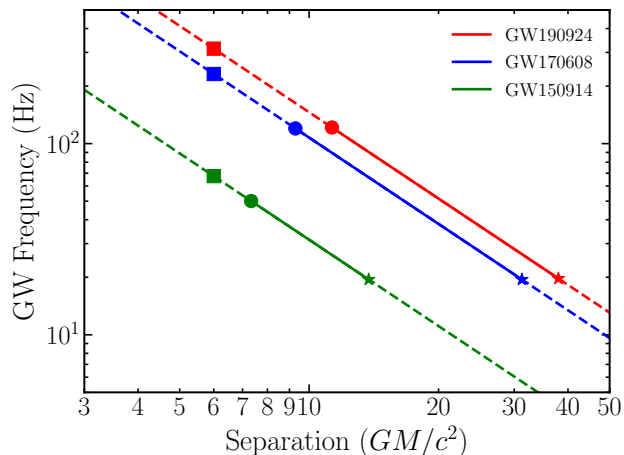
FIG. 4. The frequencies of the emitted gravitational waves as a function of the binary separation, for the three black-hole binary events that we explore here. The stars denote the lowest frequency analyzed, in each case this is 20 Hz; the filled squares denote the frequency at which the separation is $\sim 6GM/c^2$; the filled circles denote the frequencies at which we terminate the analyses of the early inspiral phases of the events. The solid lines identify the range of frequencies and corresponding binary separations that contribute to the constraints on deviations from the GR predictions explored here.

nermost stable circular orbit in the system, so that their individual spacetimes are mildly perturbed away from the equilibrium solutions. Moreover, the relatively large separations allow for useful constraints to be obtained even with a post-Newtonian expansion of the waveforms. Finally, modeling the inspiral phases allows us to place constraints on deviations from the GR predictions for the equilibrium black-hole metrics that can be directly compared to those obtained using the EHT observations, as discussed in the previous section.

To date, the LIGO Scientific Collaboration and Virgo Collaboration have confirmed the observation of gravitational waves from 47 compact binary coalescences with a false alarm rate below 1 per year, including 44 binary black hole systems, two binary neutron star systems, and GW190814, which is likely a binary black hole system [28]. In this paper, we reanalyze the LIGO/Virgo data for three sources, focusing exclusively on tests of GR during the inspiral phase of three events: (i) GW150914, the first and highest signal-to-noise ratio binary black hole merger observed to date; (ii) GW170608, the low-mass system with the strongest constraints on GR deviations to date; (iii) GW190924_021846, the lowest mass binary black hole system observed to date. We analyze these three out of the large number of available systems in order explore the mass dependence of the constraints.

Figure 4 shows the frequency of the emitted gravitational waves as a function of the binary separation for the three sources we analyze here. For the purposes of this figure we set the total mass of GW150914 to $M = 65M_\odot$,

of GW170608 to $M = 19M_\odot$, and of GW190924_021846 to $14M_\odot$ [28]. The stars denote the gravitational-wave frequencies ($\sim 20$ Hz) at which the signal amplitudes become discernible above the noise of the detectors. The filled squares represent the frequencies at which the orbital separation is $6GM/c^2$, i.e.,

$$f_6 = \frac{1}{6\sqrt{6}\pi} \left( \frac{GM}{c^3} \right)^{-1} . \tag{41}$$

Finally, the filled circles represent the frequencies at which we terminate the analyses of the early inspiral phases of the events in order to ensure that our constraints are driven solely by the inspiral regimes (see below). The solid lines identify the range of frequencies and corresponding binary separations that primarily inform the tests of deviations from GR predictions based on the inspiral phases.

### A. Post-Newtonian Tests of Black-Hole Binary Inspirals

In the study of potential deviations from the GR waveforms during the inspiral phases, the LIGO/Virgo collaboration has been employing, among others, the `IMRPhenomPv2` waveform model [61], which is a processing modification to the aligned-spin waveform model `IMRPhenomD` [27]. For this model, the phase evolution during the inspiral phase up to the third PN order is exactly the `TaylorF2` post-Newtonian model

$$\Phi_{\rm GR}(f) = 2\pi f t_c - \phi_c - \pi/4$$
$$+ \frac{3}{128\eta} (\pi f M)^{-5/3} \sum_{i=0}^{7} \phi_i (\pi f M)^{i/3} , \tag{42}$$

where $i = 2p$ is twice the PN order (in the counting system used here) and $t_c$ and $\phi_c$ are the time and phase at coalescence. Here $M = m_1 + m_2$ is the total mass of the system and $\eta \equiv m_1 m_2/M^2$ is the symmetric mass ratio. Beyond this, there are additional phenomenological correction terms at higher order. Neglecting spin corrections, the first few $\phi_i$ terms in GR are

$$\phi_0 = 1$$
$$\phi_1 = 0$$
$$\phi_2 = \frac{3715}{756} + \frac{55}{9}\eta$$
$$\phi_3 = -16\pi$$
$$\phi_4 = \frac{15293365}{508032} + \frac{27145}{504}\eta + \frac{3085}{72}\eta^2 . \tag{43}$$

The deviation from the GR predictions are then usually parametrized as a fractional shift $\delta\hat{\phi}_i$ in $\phi_i$ as described in [62]. Note that, throughout this section, we have reverted to showing explicitly the dependence of the various terms on the mass of the binary, because of the impact of the actual value of the mass on the correlations between the inferred parameters, as we describe below.
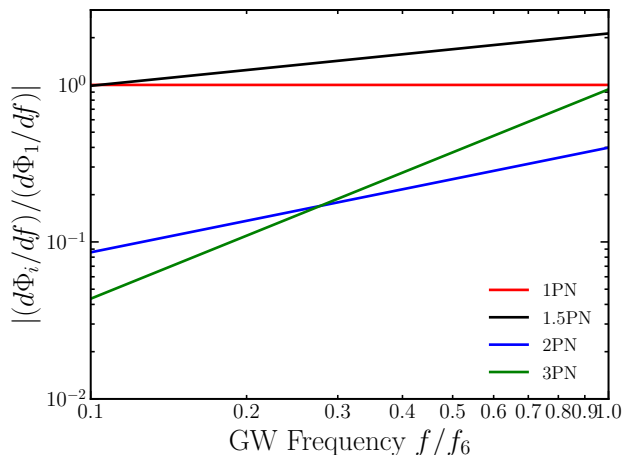
FIG. 5. The relative importance of different PN orders in the calculation of the frequency derivative of the phase of the waveform, as a function of the wave frequency. The PN order is $p = i/2$, as in Eq. (44), and all contributions are normalized to the 1PN order. The wave frequency is expressed in units of $f_6$, the frequency at a separation equal to $6GM/c^2$, at which point the inspiral phase ends. All terms have been calculated for a binary with equal masses, i.e., $\eta = 1/4$, and zero spins.

The particular values of the phase are, of course, arbitrary, and can be masked by an adjustment in the phase $\phi_c$. Instead, the inference of the model parameters is driven primarily by the evolution of the phase of the gravitational wave with frequency. In the Fourier domain, where the IMRPhenomPv2 model is defined, this is captured by the first derivative of the phase of the waveform with respect to frequency, i.e.,

$$\frac{d\Phi(f)}{df} = \sum_{i=0}^{7} \frac{d\Phi_i(f)}{df} , \qquad (44)$$

with

$$\frac{d\Phi_i(f)}{df} = \frac{M}{128\eta}(i-5)\pi^{(i-5)/3} (fM)^{(i-8)/3} \phi_i(1+\delta\hat{\phi}_i) , \qquad (45)$$

where again $i = 2p$ is twice the PN order.

In order to explore the convergence of this series, we first focus on the GR case, i.e., set $\delta\hat{\phi}_i = 0$, and express the frequency in terms of $f_6$ by setting $\hat{f} \equiv f/f_6$. Figure 5 shows the relative importance of the various PN orders with respect to the 1PN order. At the lowest frequencies accessible with LIGO/Virgo, e.g., $\hat{f} \sim 0.1$ for GW170608, the relative contribution of the 2PN and 3PN orders are comparable to each other and $\sim 10\%$ with respect to the 1PN order. As expected, their relative contribution increases (to $\sim 1/2$) as the event approaches the end of the inspiral phase and the PN expansion becomes less accurate.

Incorporating potential deviations from the GR waveforms leads to (showing only the integer PN orders, which describe potential deviations from Kerr of the equilibrium black-hole metrics; see below)

$$\frac{d\Phi(f)}{d\hat{f}} \sim -\frac{8.9}{\hat{f}^2}(1+\delta\hat{\phi}_2) - \frac{3.5}{\hat{f}^{4/3}}(1+\delta\hat{\phi}_4) - \frac{8.3}{\hat{f}^{2/3}}(1+\delta\hat{\phi}_6) , \qquad (46)$$

where, for clarity, we have set $\eta = 1/4$. If we evaluate this expression at $f = (1/10)f_6$, i.e., approximately when the LIGO and Virgo interferometers start detecting the signal from GW170608, we obtain

$$\frac{d\Phi(f)}{d\hat{f}}\bigg|_{\hat{f}=1/10} \simeq -1002 \left( 1 + 0.89\delta\hat{\phi}_2 + 0.08\delta\hat{\phi}_4 \right.$$
$$\left. +0.04\delta\hat{\phi}_6 + ... \right) . \qquad (47)$$

The convergence becomes increasingly slower as the binary separation decreases towards $6GM/c^2$ and the frequency of the wave increases towards $f_6$.

We show later in this section that the constraints on the various deviation parameters are actually driven by the second derivative of $\Phi$ with respect to frequency, because changes in the first derivative can be masked by an adjustment in the coalescence time $t_c$. Had we considered the convergence of the second derivative with frequency, we would have reached a similar conclusion.

The rate of convergence for the waveforms of coalescing binary black holes (see, e.g., Fig. 5) is similar to the rate of convergence of the series (33) used in tests involving the size of the black-hole shadow. This might appear counter intuitive, since the separations of the binaries in the inspiral phase are in the range $\sim (6-40)GM/c^2$, whereas the radius of the photon orbit is at $3GM/c^2$. The reason lies in the fact that the coefficients of the various PN orders in the phase evolution of the waveforms are increasing rapidly with PN order: for a binary system with equal masses, Eq. (43) gives $(\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, ...) \simeq (0, 6, -50, 46, 154, -652, ...)$.

Previous tests of the compact binary phase evolution have mostly allowed for just one parameter to deviate at a time (but see [4] for the analysis of GW150914 in which multiple parameters were varied simultaneously). This is done in order to improve the constraints at specific post-Newtonian orders and reduce computational challenges. When all deviation parameters are allowed to vary, their posteriors are highly correlated and, for the 1PN and higher terms, span a range comparable to that imposed by the priors (see [4]). In fact, if all post-Newtonian coefficients are allowed to deviate from the GR values, the problem is under-determined. The number of free parameters governing the intrinsic phasing of the inspiral, which includes the masses, the spins, and the deviation terms terms $\delta\hat{\phi}_i$, is greater than the number of post-Newtonian coefficients. Therefore, even if all coefficients were to be measured perfectly in the absence of noise, the physical parameters and deviation terms can only be constrained to lie on a degenerate hypersurface whose dimensionality is given by the number of excess parameters relative to

the number of measured post-Newtonian terms. However, in order to compare the LIGO/Virgo constraints to those imposed by the measurement of the black-hole shadow in M87, we are interested specifically in the correlations between the various deviation parameters.

Another complication in using the inspiral gravitational waveforms to constrain the equilibrium spacetimes of black holes, which is our aim here, is the need to have a model for the radiative (or dissipative) properties of the underlying theory. In fact, it is well known that several modifications to GR give rise to altered waveforms, even though the black-hole metrics in these modified theories remain identical to the Kerr solution [63]. This is the reason why the post-Newtonian constraints imposed by the LIGO/Virgo events on, e.g., the phase evolution of the detected gravitational waves amalgamate potential deviations of both the equilibrium spacetimes (i.e., the conservative properties of the theory) and of the gravitational-wave emission and propagation (the dissipative and radiative properties of the theory).

In order to break the complete degeneracy between varying the physical parameters of the black-hole binaries and the deviation terms of the metrics, we explore a different cross section of the parameter space of possibilities compared to previous work (e.g., [6]). In accordance with our goal to use gravitational-wave observations to constrain the equilibrium spacetimes of black holes, we will consider the case where the radiative aspects of the theory are the same as in GR and allow only for the black-hole spacetimes to deviate from the GR solutions. The emission of gravitational waves, which drives the phase evolution of the waveforms, will again start at the 2.5PN order, as in GR, and the gravitational-wave amplitude will be proportional to the second time derivative of the quadrupole moment of the spacetime. However, this derivative will be determined by the time-dependent relative positions of the two black holes in the orbit, which in turn depend on the equilibrium spacetimes of the individual black holes. This is how the PN parameters that we are concerned with in this study enter the calculation of the waveforms and can be constrained with observations [64].

### B. Modeling Inspiral Data with ppE Waveforms

In accordance with the reasoning described above, we will follow the procedure outlined in Refs. [32, 65, 66] for calculating the waveforms of gravitational waves during the inspiral phase based on the parametric post-Einstein (ppE) phenomenological approach [67] and using effective one-body dynamics [68][69]. This enables us to compare model waveforms to LIGO/Virgo data by allowing for simultaneous deviations at more than a single PN term while reducing the extent of correlations discussed above. Because we will be using the ppE formalism, we are also able to incorporate amplitude and phase information in our model.

As shown in Ref. [32], the waveform evolution during the inspiral phase depends only on the $tt$-components of the black-hole metrics as written in areal coordinates. In the ppE formalism, the waveform during the inspiral phase of a binary, written in the frequency domain, is given by [67]

$$\tilde{h}_{\mathrm{ppE}}(f) = \mathcal{A}_{\mathrm{GR}}(f)\left(1 + \alpha_{\mathrm{ppE}}u^{a_{\mathrm{ppE}}}\right)$$
$$\exp\left\{i\left[\Phi_{\mathrm{GR}}(f) + \beta_{\mathrm{ppE}}u^{b_{\mathrm{ppE}}}\right]\right\} \quad (48)$$

where $A_{\mathrm{GR}}$ and $\Phi_{\mathrm{GR}}(f)$ are the GR predictions for the amplitude and phase of the waveform, $u \equiv (\pi\mathcal{M}f)^{1/3}$, $\mathcal{M} = M\eta^{3/5}$ is the chirp mass, $M = m_1 + m_2$ is the total mass, $m_1$ and $m_2$ are the masses of the two black holes, and $\eta = m_1 m_2/M^2$ is the symmetric mass ratio, as before. The various parameters with the "ppE" subscripts describe potential deviations from the GR predictions. In order to incorporate the individual PN orders, we write

$$\tilde{h}_{\mathrm{ppE}}(f) = \mathcal{A}_{\mathrm{GR}}(f)e^{i\Phi_{\mathrm{GR}}(f)}\prod_{p=1}^{3}\mathcal{A}_{\mathrm{ppE},p}e^{i\Phi_{\mathrm{ppE},p}} \quad (49)$$
$$\mathcal{A}_{\mathrm{ppE},p} = (1 + \alpha_{\mathrm{ppE},p}u^{a_{\mathrm{ppE},p}})$$
$$\Phi_{\mathrm{ppE},p} = \beta_{\mathrm{ppE},p}u^{b_{\mathrm{ppE},p}}. \quad (50)$$

The two ppE parameters that affect the phase evolution of the gravitational wave, which we will use here, are the amplitude $\beta_{\mathrm{ppE}}$ and the power-law index $b_{\mathrm{ppE}}$. The latter is fixed for each successive post-Newtonian order $p$ by [32]

$$b_{\mathrm{ppE},p} = 2p - 5 . \quad (51)$$

The former depends on the particular parametrization of the PPN metric in areal coordinates. For the metric (7), application of Eq. (11) of [32] for the $p-$th PN order gives

$$\beta_{\mathrm{ppE},p} = (-1)^{p+1}\frac{5(p+1)(2p+1)}{8(2p-8)(2p-5)\eta^{2p/5}}\zeta_p . \quad (52)$$

Applied to the 1PN and 2PN orders, we obtain

$$\beta_{\mathrm{ppE},1} = \frac{5\zeta_1}{24\eta^{2/5}} = \frac{5(\beta_1 - \gamma_1)}{24\eta^{2/5}}$$
$$\beta_{\mathrm{ppE},2} = -\frac{75\zeta_2}{32\eta^{4/5}}$$
$$= -\frac{75}{128\eta^{4/5}}\left[2\left(\gamma_1^2 - 1\right) - 8(\beta_1\gamma_1 - 1)\right.$$
$$\left. + 3(\beta_2 - 1) + 3(\gamma_2 - 1)\right] , \quad (53)$$

Note that there is one important difference between the phase evolution during the inspiral phase of the ppE waveform (48) and the `IMRPhenomPv2` model (42) that arises when terms at the 2.5PN and higher order are considered: the `IMRPhenomPv2` model involves logarithmic correction terms in frequency, whereas the ppE waveforms do not. Nevertheless, the magnitudes of these logarithmic corrections are subdominant for the cases we consider here and can be neglected.

The amplitude terms are directly related to the phase terms via [32]

$$a_{\rm ppE} = 2p \tag{54}$$

and

$$\alpha_{\rm ppE} = \frac{(-1)^{p+1}(p+1)(2p-1)}{3}\frac{\zeta_p}{\eta^{2p/5}} \ . \tag{55}$$

Using Eqs. (48)-(53), it is also straightforward to connect the fractional deviations $\delta\hat{\phi}_i$, inferred from the LIGO/Virgo measurements [6] to the ppE coefficients and to the various PN parameters of the metric (7). The result for the 1PN and 2PN orders is

$$\zeta_1 \equiv \beta_1 - \gamma_1 = \frac{(743 + 924\eta)}{1344}\delta\hat{\phi}_2 \tag{56}$$

and

$$\zeta_2 = -\frac{(3058673 + 5472432\eta + 4353552\eta^2)}{10160640}\delta\hat{\phi}_4 \ . \tag{57}$$

We use the waveform (48) to explore the correlations in the expanded parameter space through the ensemble Markov chain Monte Carlo sampler `emcee` [70] as implemented in `Bilby` [71]. In order to reduce the burn-in time, we choose our starting ensemble by taking samples from the posterior distributions obtained for the case where GR is assumed to be correct [72] and draw starting samples for the non-GR parameters from tightly peaked distributions around zero. We ensure convergence of the MCMC ensemble by visually inspecting the chains.

We employ the same prior distribution on the GR parameters as in [72] and impose uniform priors on the post-Newtonian terms $\zeta_p \in [-1000, 1000]$. We analyze four seconds of data for GW150914 and sixteen seconds of data for GW170608 and GW19092_021846, with the trigger time placed two seconds before the end of the analyzed data. We neglect the effect of calibration uncertainty in our analysis as it has been shown to be negligible for these events [73]. We use noise power spectral density estimates obtained with the `BayesLine` algorithm [74, 75]. In order to ensure that our constraints are driven solely by the inspiral regime, we impose upper frequency cutoffs of 120 Hz, 120 Hz, and 50 Hz for GW170608, GW190924_021846, and GW150914 respectively.

Figure 6 shows the one- and two-dimensional marginalized posterior distributions for the $\zeta_p$ PN deviation parameters and the chirp mass for GW170608 and GW190924_021846. (The results are not shown for GW150914 since too few cycles from the inspiral regime were observed and the deviation parameters are much less constrained.) All deviation parameters are highly correlated with each other and with the chirp mass of the system. Since gravitational-wave interferometers are more sensitive to the phase evolution of the source than to the amplitude, the posteriors here are not significantly more

constraining than when we allowed for just the phase to deviate from the GR prediction.

In this analysis, although we allow for the black holes to have non-zero spins, we have neglected the effects of spins on the corrections to the GR waveforms. In order to explore the impact of this assumption on our results, we repeat the analysis while fixing the black hole spins to zero. In Figure 7, we compare the one- and two-dimensional marginalized posterior distributions obtained for GW170608 when including and excluding spin effects in the GR waveform template. Due to degeneracies between black hole mass and spin in the evolution of the gravitational waveform, fixing the spins to zero leads to marginally tighter constraints on other parameters. This translates into slightly narrower credible intervals on the non-GR parameters.

We aim to develop a simple understanding of the correlations using the following arguments. The similarity of two waveforms $\tilde{h}_1(f)$ and $\tilde{h}_2(f)$ is expressed through their normalized match

$$\mathcal{O}(\tilde{h}_1, \tilde{h}_2) = \frac{I_{12}}{\sqrt{I_{11}I_{22}}} \ ,$$
$$I_{12} \equiv \Re \int df \frac{\tilde{h}_1(f)\tilde{h}_2^*(f)}{S_n(f)}, \tag{58}$$

where $S_n(f)$ is the frequency-dependent noise power spectral density of the interferometer. If the normalized match is close to unity, then the noise-weighted residuals from match-filtering waveform $\tilde{h}_1$ against $\tilde{h}_2$ are low. In other words, if $\tilde{h}_1$ represents the actual signal in the data and $\tilde{h}_2$ the waveform model used for parameter estimation, the likelihood of observing the data given the model is high.

Assuming that the two waveforms differ only in the phase evolution, as is the dominant effect measured by the interferometers, and expanding the phases around a fiducial frequency $f_0$, we obtain

$$\tilde{h}_i(f) = A\exp\left[i\Phi_i(f)\right]$$
$$\sim A\exp\left[i\Phi_i(f_0) + i\frac{d\Phi_i}{df}\Big|_{f_0}(f - f_0) + \frac{i}{2}\frac{d^2\Phi_i}{df^2}\Big|_{f_0}(f - f_0)^2 + ...\right] \tag{59}$$

In the stationary phase approximation and, ignoring the overall phase of the integral, which can be incorporated into $\phi_c$ and is marginalized over during the analysis, the overlap integral becomes

$$I_{12} = \int \frac{df}{S_n(f)}\exp\left[i\left(\frac{d\Phi_1}{df} - \frac{d\Phi_2}{df}\right)(f - f_0)\right]$$
$$+ \int \frac{df}{S_n(f)}\exp\left[\frac{i}{2}\left(\frac{d^2\Phi_1}{df^2} - \frac{d^2\Phi_2}{df^2}\right)(f - f_0)^2\right] + \tag{60}$$

The first of these two integrals amounts to a time offset (cf. the term $2\pi f t_c$ in the waveform phase in Eq. (42)) and is also marginalized over. As a result, to leading
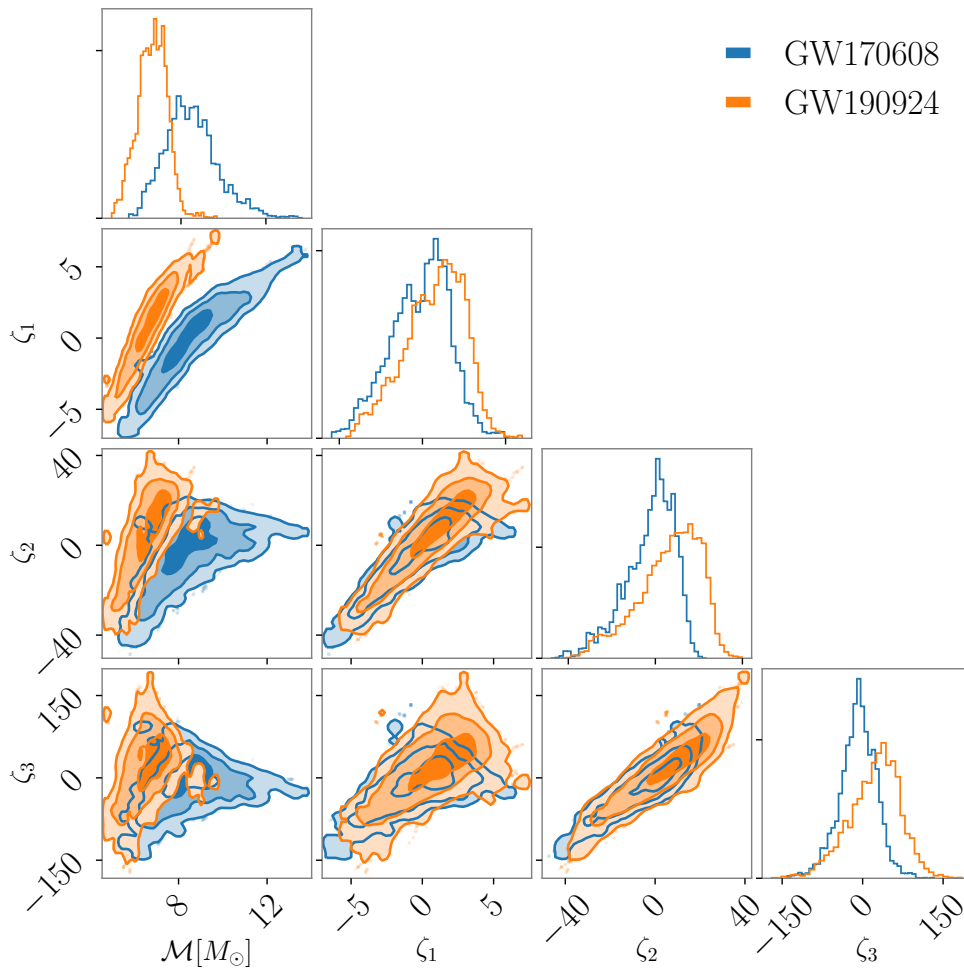
FIG. 6. Correlated uncertainties between the chirp mass $\mathcal{M}$ and the parameters $\zeta_p$ that describe deviations from the GR metrics at various post-Newtonian orders. These results were obtained for two relatively low-mass binary black hole mergers, GW170608 and GW190924, which have total masses of $\sim 19 M_\odot$ and $\sim 14 M_\odot$ respectively. There are tight mass-dependent correlations between the ppE parameters. The mass dependence of these correlations creates the wedge structure in the two-dimensional posterior distributions.

order, the likelihood is maximized when the rapidly oscillating term inside the second square bracket is close to zero, i.e., when the second derivatives of the phase functions are similar at the frequencies of interest.

Requiring that the second derivative of the GR waveform at a frequency $f_0$ is equal to the second derivative of the waveform when the deviation parameters $\zeta_1$ and $\zeta_2$ are allowed to take non-zero values results in the anticorrelation

$$\zeta_2 \sim \frac{2}{5\pi^{2/3}(f_0 M)^{2/3}}\zeta_1$$
$$\sim 4.2 \left(\frac{f_0}{100\ \text{Hz}}\right)^{-2/3}\left(\frac{M}{19 M_\odot}\right)^{-2/3}\zeta_1 \,, \quad (61)$$

where we have evaluated the various expressions for $\eta = 1/4$ and chose $f_0 = 100$ Hz, roughly the frequency at which the LIGO and Virgo noise spectral density is

minimal, as the fiducial frequency value. This captures the correlation seen in the results of the comparison of the model to the data (Fig. 6). The mass dependence in this expression also accounts for the wedge shapes of the correlations between the various deviation parameters. A similar understanding of the correlations could be obtained with a traditional Fisher matrix analysis.

Surprisingly, for the masses of the two sources analyzed here, the direction of this correlation is nearly parallel to the one between the PN parameters of the metric, as inferred from the measurement of the size of a black-hole shadow (cf. Eq. [40]). The slope of the correlation, however, depends on the mass of the system. More massive systems spend too few cycles in the inspiral regime to be useful for probing post-Newtonian coefficients. However, lower-mass binary black holes are particularly promising in breaking the degeneracy with EHT observations and
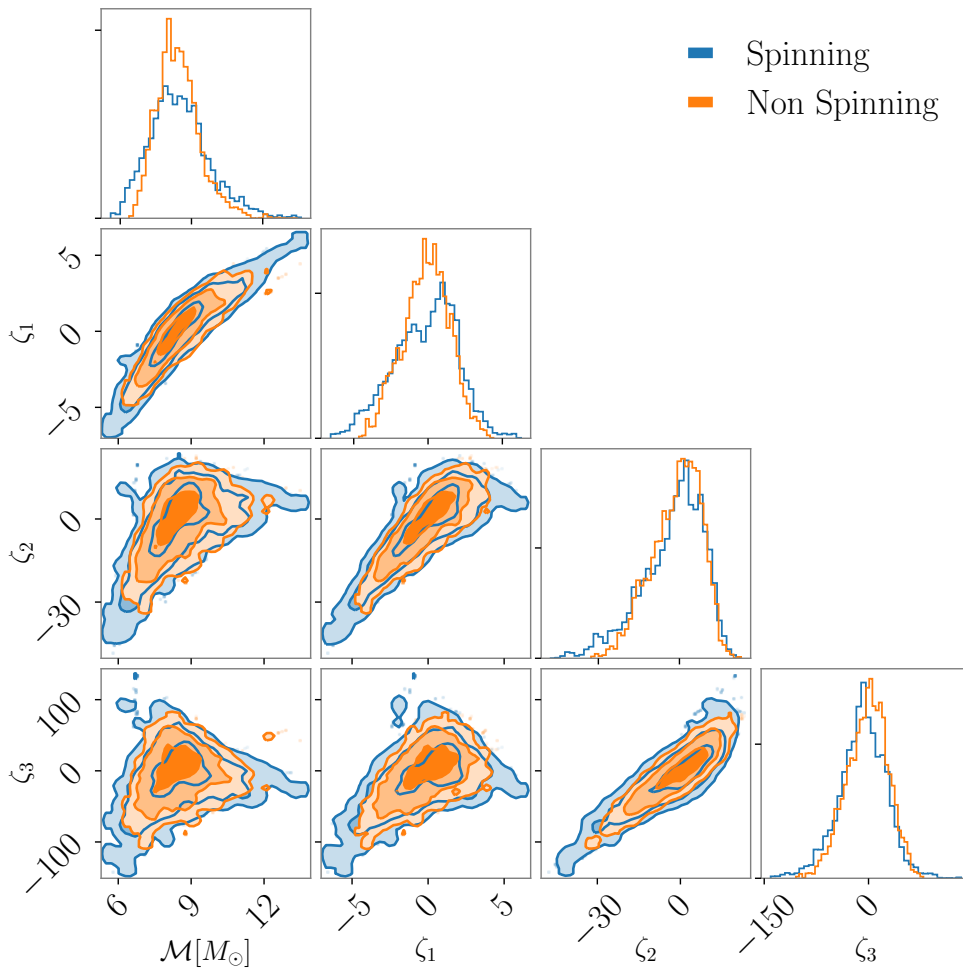
FIG. 7. Comparing correlated uncertainties between the chirp mass $\mathcal{M}$ and the parameters $\zeta_p$ that describe deviations from the GR metric at various post-Newtonian orders for GW170608, when including or excluding spin effects in the GR waveform. Including spin effects in the GR waveform does not significantly change the posterior distributions for the non-GR parameters.

providing a complementary test. Binary neutron stars have even lower masses, but neutron stars may exhibit different couplings than black holes in alternate theories of gravity. Alternatively, a change in the detector's typical sensitive frequency band, represented by $f_0$, would allow future third-generation ground-based detectors and particularly the LISA space instrument to provide constraints with very different correlations.

## V.  DISCUSSION

In this paper, we explored the constraints imposed on potential deviations from the Kerr metric by the observation of the black-hole shadow in the M87 galaxy and the detection of gravitational waves during the inspiral phase of binary black-hole coalescence. There are a number of similarities and differences between these two types of tests of the Kerr metric.

The shadow observations probe the equilibrium spacetimes of black holes whereas the detections of gravitational waves also probe the dynamics of the theory. For this reason, when comparing the two types of tests, we only consider the constraints placed by gravitational-wave observations on the metrics of the black holes. On the other hand, the shadow observations probe length and mass scales that are 8 orders of magnitude larger and curvature scales that are 16 orders of magnitude smaller than those measured by LIGO/Virgo data. It is therefore conceivable that the stellar-mass black holes probed by gravitational-wave observations and the supermassive black holes probed by shadow observations might not be described by the same metric. In that case, the constraints imposed by the two types of tests cannot be combined, but only offer complementary information.

There is limited information available from either type of observation on aspects of the spacetime that are controlled by the black-hole spins. For the case of the shad-
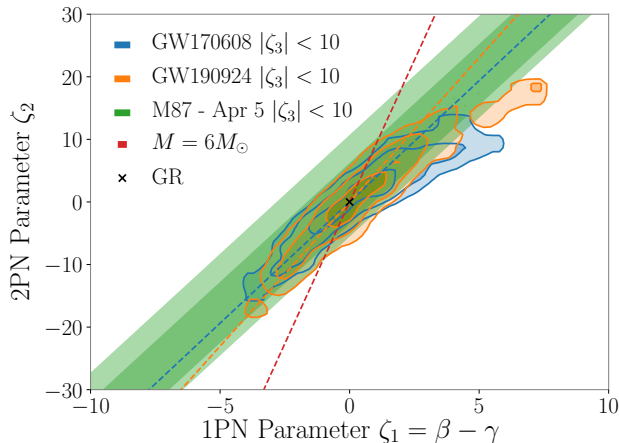
FIG. 8. Comparison of the correlated posteriors over the two PN parameters $\zeta_1$ and $\zeta_2$ as inferred from the black-hole shadow test (green slanted contours) and the two coalescing events GW170608 (blue) and GW190924_021846 (orange). The blue and orange contours show the 1-, 2-, and 3-$\sigma$ credible regions. For both tests, we assume that the 3PN parameter $|\zeta_3| < 10$ and all higher-order PN parameters introduce negligible corrections. Because of a fortuitous coincidence, the inspiral tests for binary black holes with a total mass of $\sim 20 M_\odot$ and detectors with a peak sensitivity around $\sim 100$ Hz lead to correlated uncertainties that are parallel to those of the black-hole shadow tests. The dashed lines indicate the expected correlation from Eq. (61). The red curve indicates the expected correlation for a binary black hole system with a low total mass of $6 M_\odot$. The detection of such a binary system can assist in breaking the degeneracies between the deviation parameters.

FIG. 9. Posterior distributions for the fractional deviation $\delta$ (Eq. (36)) of the $tt-$component of the metric from the GR solution, as imposed by the EHT measurement of the size of the black-hole shadow in M87 (green), and the LIGO/Virgo measurements of the inspiral phases of events GW170608 (blue) and GW190924 (orange). Both EHT and LIGO/Virgo measurements are consistent with no deviations from GR and lead to bounds of comparable magnitude.

ows, this is a consequence of a fortuitous near cancellation of the effects of frame dragging and of the quadrupole moment of the spacetimes, both of which depend on spin [24]. For the gravitational-wave observations, there seems to be a paucity of merging binary black holes in the Universe with substantial spins [28], although merger products have dimensionless spins of $\sim 0.7$ and their impact on the ringdown portion of the waveform has been explored [6]. Neglecting the effects of black-hole spins, as we have done here, substantially reduces the complexity of the problem.

We have chosen here to express the constraints on possible deviations from the Kerr metric in terms of the coefficients of the parametric post-Newtonian expansions of the metric predictions for weak-field tests (even though for the case of the shadow tests we use metrics that are remain regular all the way to the horizons). This might appear to not be warranted by the fact that both tests probe strong gravitational fields, for which the post-Newtonian expansion converges slowly. Indeed, the radius of the photon orbit is at $\sim 3GM/c^2$, the radius of the shadow is $\sim 5GM/c^2$, and the frequencies detected by LIGO/Virgo probe typical separations $6 - 30GM/c^2$ during the inspiral phases. However, the relatively low accuracy of the
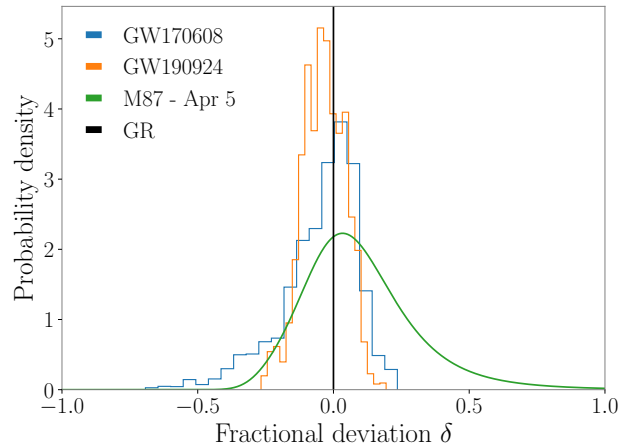
observations allows us to construct such post-Newtonian expansions without significant concerns. For example, the size of the black-hole shadow in M87 has been measured to be consistent with the Kerr predictions to within an accuracy of $\sim 17\%$ [18]. If naturalness forces all post-Newtonian terms to have coefficients of the same order of magnitude, then successive terms are a factor of $\sim 4 - 5$ smaller than the previous ones; at the 3PN order the correction is at the percent level, i.e., below the current observational uncertainty. The same argument can also be made for the post-Newtonian expansion of the gravitational waveforms and their observations, though these coefficients are known to increase at higher PN orders (see §IVA). The situation is, of course, starkly different in the case of Solar System tests only because the fractional observational uncertainties there are as small as $10^{-5}$ and therefore require an expansion that converges more rapidly (as the PPN expansion does) [2].

When we consider only the constraints from the gravitational-wave observations on the metrics of non-spinning black holes and not on the dynamics, both the shadow size tests and the gravitational-wave tests depend only on the $tt$-component of their metrics written in areal coordinates [19, 32]. Because of another fortuitous coincidence, the correlations between the deviation parameters constrained from the shadow test are degenerate with the correlations constrained from the gravitational-wave tests, for binary black holes with a total mass of $\sim 20$ $M_\odot$ and detectors with a peak sensitivity around 100 Hz. This similarity in the degeneracy between the two types of tests is shown explicitly in Fig. 8 for the 1PN and 2PN deviation parameters as is the expected

degeneracy (61) that we derived in the previous section.

The similarity in the degeneracy between the various PN parameters in the two types of tests allows us to project both the EHT and the LIGO/Virgo correlated constraints on a single parameter, i.e., the fractional deviation $\delta$ defined in Eq. (36), which is directly related to the deviation of the shadow size from the GR prediction. The resulting posterior distributions for this parameter obtained for both the shadow-size and the inspiral tests are shown in Fig. 9. The GR prediction is indicated by the solid black vertical line. This Figure shows that that current imaging and gravitational-wave data provide similar constraints on any deviations of slowly spinning black-hole spacetimes from the Kerr metric across the mass and curvature scales probed by the two experiments. We note that implicit in the analysis that led to Figure 9 is that the deviations at fourth and higher post-Newtonian orders are negligible as are deviations in the dynamics of the theory, which are described by the half-integer orders in the gravitational waveforms, and non-linear combinations of the $\zeta_i$. For this Figure, we additionally impose that $|\zeta_i| < 10$ for all non-GR correction terms; not including this constraint approximately doubles the width of the posteriors for the GW events.

A future detection of a low-mass binary black hole, as well as observations with future ground-based and space detectors that are sensitive to a different range of gravitational-wave frequencies, would provide constraints with degeneracies that are not parallel to those of the shadow test. This is shown explicitly in Fig. 8, where the expected line of degeneracy for a $3M_\odot + 3M_\odot$ black-hole binary is shown. Under the assumption that the black-hole metrics do not depend on the mass or curvature of the objects involved, such an inspiral test can be combined with the shadow tests in order to break the degeneracies and constrain the individual deviation parameters.

[1] C. M. Will, Theory and Experiment in Gravitational Physics (1993).

[2] C. M. Will, Living Reviews in Relativity **17**, 4 (2014).

[3] N. Wex, arXiv e-prints , arXiv:1402.5594 (2014).

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

[5] B. P. Abbott, et al, LIGO Scientific, and Virgo Collaborations, Phys. Rev. D **100**, 104036 (2019); M. Isi, M. Giesler, W. M. Farr, M. A. Scheel, and S. A. Teukolsky, Phys. Rev. Lett. **123**, 111102 (2019).

[6] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, et al., arXiv e-prints , arXiv:2010.14529 (2020), arXiv:2010.14529 [gr-qc].

[7] P. G. Ferreira, Ann. Rev. Astron. Astrophys. **57**, 335 (2019).

[8] R. V. Wagoner, Phys. Rev. D **1**, 3209 (1970).

[9] T. Damour and G. Esposito-Farese, Phys. Rev. Lett. **70**, 2220 (1993); T. Damour and G. Esposito-Farèse, Phys. Rev. D **54**, 1474 (1996).

[10] Here, $M$ denotes the mass scale of the test, $R$ the length scale, $G$ is the gravitational constant, and $c$ is the speed of light.

[11] T. Baker, D. Psaltis, and C. Skordis, Astrophys. J. **802**, 63 (2015).

[12] Using the gravitational-wave events as sirens, additional tests have been performed on the propagation of gravitational waves across the large distances from the sources to the Earth, in order to constrain the mass of the graviton [4] or the speed of gravitational waves that is allowed to be different from the speed of light in some modified gravity theories inspired by cosmological observations [80].

[13] Gravity Collaboration, Phys. Rev. Lett. **122**, 101102 (2019).

[14] T. Do and et al., Science **365**, 664 (2019).

[15] Gravity Collaboration, Astron. Astrophys. **636**, L5 (2020).

[16] A. Hees, T. Do, B. M. Roberts, A. M. Ghez,

S. Nishiyama, R. O. Bentley, A. K. Gautam, S. Jia, T. Kara, J. R. Lu, H. Saida, S. Sakai, M. Takahashi, and Y. Takamori, Phys. Rev. Lett. **124**, 081101 (2020).

[17] A. Hees, T. Do, A. M. Ghez, G. D. Martinez, S. Naoz, E. E. Becklin, A. Boehle, S. Chappell, D. Chu, A. Dehghanfar, K. Kosmo, J. R. Lu, K. Matthews, M. R. Morris, S. Sakai, R. Schödel, and G. Witzel, Phys. Rev. Lett. **118**, 211101 (2017).

[18] Event Horizon Telescope Collaboration, Astrophys. J. Lett. **875**, L1 (2019); Astrophys. J. Lett. **875**, L6 (2019).

[19] D. Psaltis, et al, and the EHT Collaboration, Phys. Rev. Lett. (2020).

[20] T. Johannsen and D. Psaltis, Phys. Rev. D **83**, 124015 (2011).

[21] S. Vigeland, N. Yunes, and L. C. Stein, Phys. Rev. D **83**, 104027 (2011).

[22] T. Johannsen, Phys. Rev. D **87**, 124017 (2013).

[23] T. Johannsen, Phys. Rev. D **88**, 044002 (2013).

[24] T. Johannsen and D. Psaltis, Astrophys. J. **718**, 446 (2010).

[25] T. Johannsen, Astrophys. J. **777**, 170 (2013).

[26] L. Medeiros, D. Psaltis, and F. Özel, arXiv e-prints , arXiv:1907.12575 (2019).

[27] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, Phys. Rev. D **93**, 044007 (2016).

[28] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, et al., arXiv e-prints , arXiv:2010.14527 (2020), arXiv:2010.14527 [gr-qc]; arXiv e-prints , arXiv:2010.14533 (2020), arXiv:2010.14533 [astro-ph.HE].

[29] A. Joyce, B. Jain, J. Khoury, and M. Trodden, Phys. Rep. **568**, 1 (2015).

[30] C. Burrage and J. Sakstein, Living Reviews in Relativity **21**, 1 (2018).

[31] K. S. Thorne and J. J. Dykla, Astrophys. J. Lett. **166**, L35 (1971); M. A. Scheel, S. L. Shapiro, and S. A. Teukolsky, Phys. Rev. D **51**, 4236 (1995); D. Psaltis, D. Perrodin, K. R. Dienes, and I. Mocioiu, Phys. Rev. Lett. **100**, 091101 (2008); T. P. Sotiriou and V. Faraoni, Phys. Rev. Lett. **108**, 081103 (2012).

[32] Z. Carson and K. Yagi, Phys. Rev. D **101**, 084050 (2020), arXiv:2003.02374 [gr-qc].

[33] C. M. Will, Proceedings of the National Academy of Science **108**, 5938 (2011), arXiv:1102.5192 [gr-qc].

[34] P. Kanti, N. E. Mavromatos, J. Rizos, K. Tamvakis, and E. Winstanley, Phys. Rev. D **54**, 5049 (1996); Phys. Rev. D **57**, 6255 (1998).

[35] N. Yunes and L. C. Stein, Phys. Rev. D **83**, 104002 (2011).

[36] K. Yagi, N. Yunes, and T. Tanaka, Phys. Rev. D **86**, 044037 (2012).

[37] D. Ayzenberg and N. Yunes, Phys. Rev. D **90**, 044066 (2014); R. McNees, L. C. Stein, and N. Yunes, Classical and Quantum Gravity **33**, 235013 (2016).

[38] G. Antoniou, A. Bakopoulos, and P. Kanti, Phys. Rev. Lett. **120**, 131102 (2018); Phys. Rev. D **97**, 084037 (2018).

[39] H. O. Silva, J. Sakstein, L. Gualtieri, T. P. Sotiriou, and E. Berti, Phys. Rev. Lett. **120**, 131104 (2018); D. D. Doneva and S. S. Yazadjiev, Phys. Rev. Lett. **120**, 131103 (2018).

[40] L. C. Stein and K. Yagi, Phys. Rev. D **89**, 044026 (2014), arXiv:1310.6743 [gr-qc].

[41] S. B. Giddings, International Journal of Modern Physics D **25**, 1644014 (2016).

[42] S. B. Giddings, Classical and Quantum Gravity **33**, 235010 (2016), arXiv:1602.03622 [gr-qc].

[43] S. B. Giddings and D. Psaltis, Phys. Rev. D **97**, 084035 (2018).

[44] Event Horizon Telescope Collaboration, Astrophys. J. Lett. **875**, L4 (2019).

[45] K. Gebhardt, J. Adams, D. Richstone, T. R. Lauer, S. M. Faber, K. Gültekin, J. Murphy, and S. Tremaine, Astrophys. J. **729**, 119 (2011).

[46] D. Psaltis, F. Özel, C.-K. Chan, and D. P. Marrone, Astrophys. J. **814**, 115 (2015).

[47] Had we instead used the mass smaller black-hole mass measurements based on gas dynamics [81], we would have concluded that the black hole spacetime is not described by the Kerr metric, a situation that was assigned a negligible prior.

[48] J. M. Bardeen, in Black Holes (Les Astres Occlus), edited by C. Dewitt and B. S. Dewitt (1973) pp. 215–239; S. Chandrasekhar, The mathematical theory of black holes (1983); E. Teo, GRG **35**, 1909 (2003); R. Takahashi, Astrophys. J. **611**, 996 (2004).

[49] C. Bambi and K. Freese, Phys. Rev. D **79**, 043002 (2009).

[50] W. Israel, Physical Review **164**, 1776 (1967); Communications in Mathematical Physics **8**, 245 (1968); B. Carter, Phys. Rev. Lett. **26**, 331 (1971); S. W. Hawking, Communications in Mathematical Physics **25**, 152 (1972); D. C. Robinson, Phys. Rev. Lett. **34**, 905 (1975).

[51] T. Johannsen and D. Psaltis, Astrophys. J. **716**, 187 (2010).

[52] V. Cardoso, P. Pani, and J. Rico, Phys. Rev. D **89**, 064007 (2014).

[53] Z. Carson and K. Yagi, Phys. Rev. D **101**, 084030 (2020), arXiv:2002.01028 [gr-qc].

[54] Albeit useful in numerical studies of known metrics, such parametrizations do not guarantee the absence of pathologies when the various coefficients are chosen outside the discrete sets that are known to describe regular metrics.

[55] L. Rezzolla and A. Zhidenko, Phys. Rev. D **90**, 084009 (2014).

[56] R. Konoplya, L. Rezzolla, and A. Zhidenko, Phys. Rev. D **93**, 064015 (2016).

[57] Z. Younsi, A. Zhidenko, L. Rezzolla, R. Konoplya, and Y. Mizuno, Phys. Rev. D **94**, 084025 (2016).

[58] Y. Mizuno, Z. Younsi, C. M. Fromm, O. Porth, M. De Laurentis, H. Olivares, H. Falcke, M. Kramer, and L. Rezzolla, Nature Astronomy **2**, 585 (2018).

[59] LIGO Scientific Collaboration and J. o. Aasi, Classical and Quantum Gravity **32**, 074001 (2015), arXiv:1411.4547 [gr-qc].

[60] F. Acernese et al., Classical and Quantum Gravity **32**, 024001 (2015), arXiv:1408.3978 [gr-qc].

[61] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, Phys. Rev. Lett. **113**, 151101 (2014).

[62] M. Agathos, W. Del Pozzo, T. G. F. Li, C. Van Den Broeck, J. Veitch, and S. Vitale, Phys. Rev. D **89**, 082001 (2014), arXiv:1311.0420 [gr-qc].

[63] E. Barausse and T. P. Sotiriou, Phys. Rev. Lett. **101**, 099001 (2008).

[64] We do not explore the possibility that back-scattering of gravitational waves off the background – so-called tails of tails terms – could contribute to higher non-integer deviations in the gravitational-wave signature: the 2.5PN term corresponds to a fixed phase offset that is marginalized over, and we restrict our analysis to terms at or below 3PN.

[65] A. Cárdenas-Avendaño, S. Nampalliwar, and N. Yunes, Classical and Quantum Gravity **37**, 135008 (2020), arXiv:1912.08062 [gr-qc].

[66] S. Tahura and K. Yagi, Phys. Rev. D **98**, 084042 (2018), arXiv:1809.00259 [gr-qc].

[67] N. Yunes and F. Pretorius, Phys. Rev. D **79**, 084043 (2009).

[68] A. Buonanno and T. Damour, Phys. Rev. D **59**, 084006 (1999), arXiv:gr-qc/9811091 [gr-qc].

[69] Even though the effective one-body approach was proven to be valid only for the radiative properties of GR, this is consistent with the approach in this work of allowing only for the equilibrium spacetimes to be different than GR.

[70] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, PASP **125**, 306 (2013), arXiv:1202.3665 [astro-ph.IM].

[71] G. Ashton, M. Hübner, P. D. Lasky, C. Talbot, K. Ackley, S. Biscoveanu, et al., Astrophys. J. Suppl. Ser. **241**, 27 (2019), arXiv:1811.02042 [astro-ph.IM].

[72] I. M. Romero-Shaw, C. Talbot, S. Biscoveanu, V. D'Emilio, G. Ashton, C. P. L. Berry, et al., Mon.

[73] Not. R. Astron. Soc. 10.1093/mnras/staa2850 (2020), arXiv:2006.00714 [astro-ph.IM].

[73] E. Payne, C. Talbot, P. D. Lasky, E. Thrane, and J. S. Kissel, arXiv e-prints , arXiv:2009.10193 (2020), arXiv:2009.10193 [astro-ph.IM].

[74] B. P. Abbott, et al, LIGO Scientific, and Virgo Collaborations, Physical Review X **9**, 031040 (2019), arXiv:1811.12907 [astro-ph.HE].

[75] T. B. Littenberg and N. J. Cornish, Phys. Rev. D **91**, 084034 (2015), arXiv:1410.3852 [gr-qc].

[76] M. Vallisneri, J. Kanner, R. Williams, A. Weinstein, and B. Stephens, J. Phys. Conf. Ser. **610**, 012021 (2015), arXiv:1410.4839.

[77] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, M. Agathos, K. Agatsuma, N. Aggarwal, et al., arXiv e-prints , arXiv:1912.11716 (2019), arXiv:1912.11716 [gr-qc].

[78] Gravitational Wave Open Science Center, Strain data release for gwtc-1: A gravitational-wave transient catalog of compact binary mergers observed by ligo and virgo during the first and second observing runs (2018).

[79] LIGO Scientific Collaboration and Virgo Collaboration, Ligo virgo strain data from gwtc-2 catalog (2020).

[80] T. Baker, E. Bellini, P. G. Ferreira, M. Lagos, J. Noller, and I. Sawicki, Phys. Rev. Lett. **119**, 251301 (2017).

[81] J. L. Walsh, A. J. Barth, L. C. Ho, and M. Sarzi, Astrophys. J. **770**, 86 (2013).