Possible Causes of False General Relativity Violations in Gravitational Wave Observations

Anuradha Gupta[1⋆] K. G. Arun[2] Enrico Barausse[3,4] Laura Bernard[5] Emanuele Berti[6] Sajad A. Bhat[7,2] Alessandra Buonanno[8,9] Vitor Cardoso[10,11] Shun Yin Cheung[12,13] Teagan A. Clarke[12,13] Sayantani Datta[14,2] Arnab Dhani[8,15] Jose María Ezquiaga[10] Ish Gupta[15] Nir Guttman[12,13] Tanja Hinderer[16] Qian Hu[17] Justin Janquart[18,19] Nathan K. Johnson-McDaniel[1] Rahul Kashyap[15,20] N. V. Krishnendu[21] Paul D. Lasky[12,13] Andrew Lundgren[22] Elisa Maggio[8] Parthapratim Mahapatra[2] Andrea Maselli[23,24] Purnima Narayan[1] Alex B. Nielsen[25] Laura K. Nuttall[22] Paolo Pani[26] Lachlan Passenger[12,13] Ethan Payne[27,28] Lorenzo Pompili[8] Luca Reali[6] Pankaj Saini[2] Anuradha Samajdar[18,19] Shubhanshu Tiwari[29] Hui Tong[12,13] Chris Van Den Broeck[18,19] Kent Yagi[14] Huan Yang[30, 31, 32] Nicolás Yunes[33] B. S. Sathyaprakash[15,20,34]

1 Department of Physics and Astronomy, The University of Mississippi, University, Mississippi 38677, USA
2 Chennai Mathematical Institute, Siruseri 603103, India
3 SISSA, Via Bonomea 265, 34136 Trieste, Italy and INFN Sezione di Trieste
4 IFPU - Institute for Fundamental Physics of the Universe, Via Beirut 2, 34014 Trieste, Italy
5 Laboratoire Univers et Théories, Observatoire de Paris, Université PSL, Université Paris Cité, CNRS, F-92190 Meudon, France
6 William H. Miller III Department of Physics and Astronomy, Johns Hopkins University, Baltimore, Maryland 21218, USA
7 Inter-University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411007, India
8 Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlenberg 1, Potsdam 14476, Germany
9 Department of Physics, University of Maryland, College Park, MD 20742, USA
10 Niels Bohr International Academy, Niels Bohr Institute, Blegdamsvej 17, 2100 Copenhagen, Denmark
11 CENTRA, Departamento de Física, Instituto Superior Técnico – IST, Universidade de Lisboa – UL, Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal
12 School of Physics and Astronomy, Monash University, VIC 3800, Australia
13 OzGrav: The ARC Centre of Excellence for Gravitational-Wave Discovery, Clayton, VIC 3800, Australia
14 Department of Physics, University of Virginia, Charlottesville, Virginia 22904, USA
15 Institute for Gravitation and the Cosmos and Physics Department, Penn State University, University Park PA 16802, USA
16 Institute for Theoretical Physics, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands, EU
17 Institute for Gravitational Research, School of Physics and Astronomy, University of Glasgow, Glasgow, G12 8QQ, United Kingdom
18 Institute for Gravitational and Subatomic Physics (GRASP), Utrecht University, Princetonplein 1, 3584 CC Utrecht, Netherlands
19 Nikhef – National Institute for Subatomic Physics, Science Park, 1098 XG Amsterdam, The Netherlands
20 Department of Astronomy and Astrophysics, Penn State University, University Park PA 16802, USA
21 International Centre for Theoretical Sciences, Survey No. 151, Shivakote, Hesaraghatta, Uttarahalli, Bengaluru, 560089, India
22 University of Portsmouth, Portsmouth, PO1 3FX, United Kingdom

23 Gran Sasso Science Institute (GSSI), I-67100 L'Aquila, Italy
24 INFN, Laboratori Nazionali del Gran Sasso, I-67100 Assergi, Italy
25 Department of Mathematics and Physics, University of Stavanger, NO-4036, Norway
26 Dipartimento di Fisica, "Sapienza" Universitá di Roma & Sezione INFN Roma1, Piazzale Aldo Moro 5, 00185, Roma, Italy
27 Department of Physics, California Institute of Technology, Pasadena, CA 91125, USA
28 LIGO Laboratory, California Institute of Technology, Pasadena, CA 91125, USA
29 Physik-Institut, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland
30 Department of Astronomy, Tsinghua University, Beijing 100084, China
31 Perimeter Institute for Theoretical Physics, Waterloo, ON N2L2Y5, Canada
32 University of Guelph, Guelph, Ontario N1G 2W1, Canada
33 Illinois Center for Advanced Studies of the Universe & Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA
34 School of Physics and Astronomy, Cardiff University, Cardiff, CF24 3AA, United Kingdom

⋆ agupta1@olemiss.edu

Abstract

General relativity (GR) has proven to be a highly successful theory of gravity since its inception. The theory has thrivingly passed numerous experimental tests, predominantly in weak gravity, low relative speeds, and linear regimes, but also in the strong-field and very low-speed regimes with binary pulsars. Observable gravitational waves (GWs) originate from regions of spacetime where gravity is extremely strong, making them a unique tool for testing GR, in previously inaccessible regions of large curvature, relativistic speeds, and strong gravity. Since their first detection, GWs have been extensively used to test GR, but no deviations have been found so far. Given GR's tremendous success in explaining current astronomical observations and laboratory experiments, accepting any deviation from it requires a very high level of statistical confidence and consistency of the deviation across GW sources. In this paper, we compile a comprehensive list of potential causes that can lead to a false identification of a GR violation in standard tests of GR on data from current and future ground-based GW detectors. These causes include detector noise, signal overlaps, gaps in the data, detector calibration, source model inaccuracy, missing physics in the source and in the underlying environment model, source misidentification, and mismodeling of the astrophysical population. We also provide a rough estimate of when each of these causes will become important for tests of GR for different detector sensitivities. We argue that each of these causes should be thoroughly investigated, quantified, and ruled out before claiming a GR violation in GW observations.

Received Date
Accepted Date
Published Date

Contents

---

1 Introduction

Einstein's general theory of relativity (GR) stands as the most successful theory of gravity to date. Rigorously tested in weak-field, low-speed, and linear gravity regimes, GR has consistently withstood all scrutiny. Gravitational waves (GWs) are predictions of GR and offer a unique avenue for exploring spacetime dynamics in extreme gravitational conditions. Despite the widespread use of GWs from compact binary coalescences (CBCs) for testing GR, no deviations from the theory have been found so far (e.g., [1–12]).

The sensitivity of GW detectors has been continuously improving and LIGO and Virgo detectors are currently witnessing their fourth observing run (O4) with Advanced LIGO and Virgo sensitivity [13] which later will be joined by KAGRA [14]. These detectors will be further upgraded for the fifth observing run (O5) during 2027-2029 [15] with A+ sensitivity [16], and they will eventually be joined by LIGO-India [17, 18]. Looking further into the future beyond

O5, there is a possibility for detectors with A$^{\#}$ sensitivity [19] that are expected to be twice as sensitive as A+. Moreover, there are concrete plans to build next generation (XG) detectors, such as Cosmic Explorer [20] and Einstein Telescope [21], that are expected to be at least 10 times more sensitive than the current detectors in O4. The first space-borne mission, LISA [22], is scheduled to be launched in the mid-2030s, and it might be followed by other missions such as TianQin [23, 24], Taiji [25], DECIGO [26, 27] and LGWA [28].

With these improvements in sensitivity, thousands of CBCs are expected to be observed with high signal-to-noise ratios (SNRs) [16]. A subset of these mergers will cover extreme regions of the parameter space, including highly spinning and/or strongly precessing binaries, binaries with eccentricity, binaries involving dense matter, etc. Such binaries will have the capability to test GR stringently and constrain beyond-GR effects, if present in the data. For example, higher black hole spins lead to higher curvature outside the horizon [29], which allows one to place constraints on a variety of higher-derivative or curvature-corrected theories [30, 31]. More so, the near-horizon region of black holes could potentially access energies as large as the Planck scale that could alter the black hole ringdown spectrum if GR is modified near the event horizon [32, 33]. There is also the possibility that GR may be violated not in the ultraviolet (UV), but rather in the infrared (IR) regime of the theory, aimed at offering an alternative explanation of the dark sector. In this "IR" scenario, extending the reach of GW detectors to lower frequencies may help observe possible deviations from GR in the inspiral phase of CBCs [34–37].

The majority of tests of GR currently performed rely on waveform models that are compared with the GW data. Often these tests are formulated as *null tests* where one looks for possible departures from GR by introducing deviation parameters on a given waveform model. No statistically significant deviation from GR has been observed at the level of individual events or for the whole population [5]. However, there were a couple of events in GWTC-3 [38] that suggested GR deviations, though further investigations are needed since these deviations could be due to the use of imperfect waveform models or inadequately understood noise artifacts in the data [39].

Due to the complexity of the physics of compact binary mergers as well as the detector noise modeling, it is extremely important that there is a consensus in the community about the necessary conditions that will warrant a much more comprehensive list of tests to be carried out to vet (or rule out) a potential GR violation claim. There are two aspects to this issue. The first is to identify all possible causes which might lead to a false GR violation. The second is a checklist to be executed upon encountering a strong candidate for GR violation. The objective of this paper is to tackle the first aspect and enumerate an extensive list of scenarios that may appear as violations of GR, when in fact they are not. The second aspect requires us to construct a checklist of items that address other issues such as the statistical significance of the violation, the status of the detector, or if the violation is in contradiction with other experiments or astrophysical observations. A companion paper will address these issues and a possible formulation of a GR violation detection checklist. It is worth noting that a similar effort has been made in Section 7 of [40], albeit in the context of tests of GR using LISA. Our goal here is to broadly classify different effects that can mimic a GR violation in the context of present- and next-generation ground-based interferometric observational facilities.

There are at least three distinct scenarios that can mimic a GR violation (see Fig. 1): noise artifacts in data, waveform systematics, and astrophysical aspects, each of which is discussed at length below. Much work has already been done to understand aspects of these scenarios on tests of GR. Broadly speaking, these three scenarios also have the possibility to impact other scientific conclusions based on GW data, such as constraints on astrophysical sources or cosmological models. In many cases, efforts to understand the impact of these scenarios on astrophysics or cosmology can also illuminate potential impacts on tests of GR.

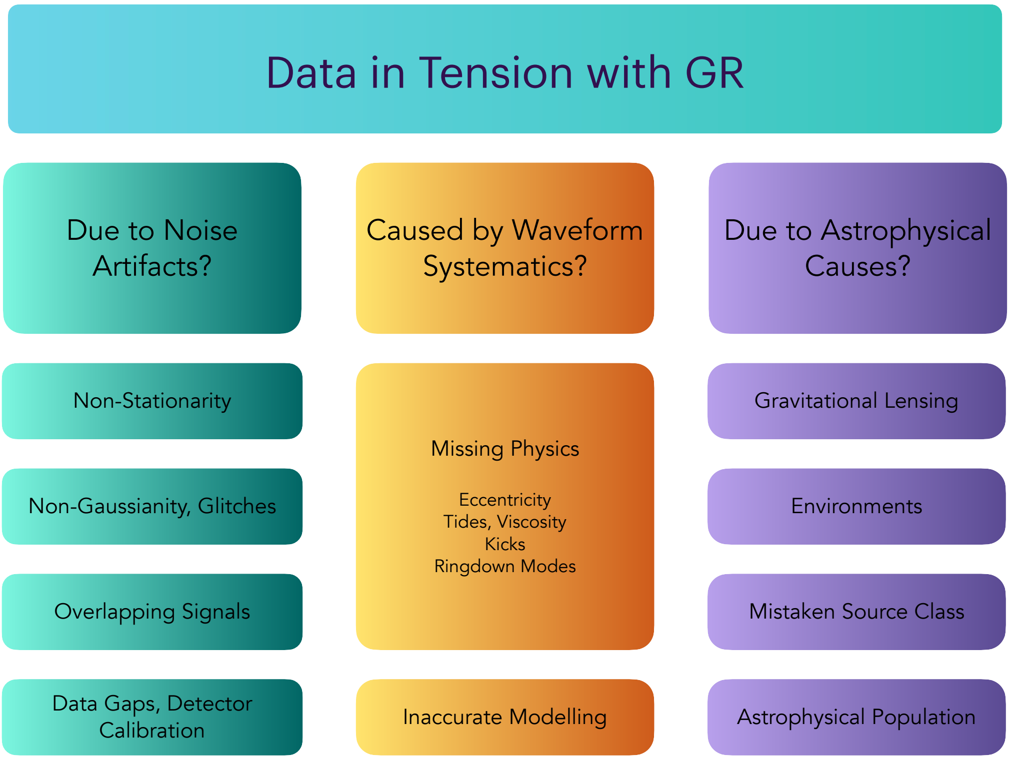

Figure 1: The diagram illustrates the principal false causes of GR violation in GW data. They are classified into three main classes: (a) noise artifacts, (b) waveform systematics, and (c) astrophysical effects.

To keep the discussion coherent, we group the causes only into these three scenarios even if this classification, or the distinction between any two causes, may seem somewhat arbitrary. For example, we keep the *overlapping signals* under noise artifacts even if this is not, strictly speaking, an instrumental noise source. Similarly, we divide issues related to waveform systematics into two main themes (*missing physics* and *inaccurate modeling*), even if the distinction between the two is not always obvious. By "missing physics" we mean cases when a particular effect is not included at all, or only partially included in the waveform models (e.g., tides and higher-order ringdown modes), while "inaccurate modeling" refers to intrinsic limitations of the waveform models in fully describing the known features of GR (e.g., waveform truncation errors).

While most of the scenarios discussed below could lead to confusion with a GR violation in a given event or subset of events, any GR deviation should be consistent across the dataset, e.g., a given theory should explain why there is evidence for deviations in certain events and not in others in a similar region of the parameter space. The ever-increasing number of events expected in the future will help sort out these situations.

2 Noise Systematics

Current interferometric GW detectors are limited by fundamental noise sources [13] which causes the noise to appear as stationary and Gaussian only over short time scales and ranges of frequency [41]. In reality, however, noise from the detectors is neither Gaussian nor stationary (see, e.g., [41–43]). It can be relatively easy to spot times of extremely bad data quality in GW data, but the challenge lies with times of subtle data quality issues. The origin of noise sources is notoriously difficult to pinpoint, even for obvious cases of poor data quality. However, it is essential that we understand our noise, remove any bias that noise introduces, and accurately infer the parameters of the observed sources.

In this Section, we discuss the three main sources of noise (namely, non-stationary, non-Gaussian, and overlapping signals) observed in ground-based detectors that can affect our inference of transient GW signals. We also discuss the systematic error due to the gaps in data and calibration of the GW instruments that may also introduce some bias in the inference results.

2.1 Non-stationarity

Non-stationarity is a broadband form of noise which causes the statistical properties of the background to change with time. Non-stationarity occurs on the order of tens of seconds in the current LIGO detectors and can be caused by both instrumental and environmental sources [42,44], such as detector's lockloss [45], variable seismic motion, thunderstorms [43], magnetic effects from lightning strokes [46], and radio frequency interference [42, 43]. This form of noise has been shown to affect the estimation of source parameters [47,48]. Modelled searches typically estimate a detector's power spectrum over several minutes [49–51], which can cause the matched filter to miss the variable nature of the noise, affecting the search sensitivity. One method to account for this is to construct a statistic which tracks the variation of the power spectrum and to normalize the ranking statistic used by the detection pipeline [51–53]. The method presented in [52] is also used to assess the stationarity of the data around candidate GW events [43]. This is because non-stationary noise can impact binary neutron star signal parameters [54, 55] since noise estimates, usually calculated over minutes, fail to capture variations on shorter time scales. As signals from sufficiently massive binary black holes are usually shorter than the typical time scale of non-stationary noise, these sources are not thought to be affected. However, due to their long duration in the sensitivity band, sub-solar mass binary black hole searches will be affected, especially in the XG era where the signal duration could be up to several days.

To date, this form of noise has not seriously affected the conclusions drawn from any of the LIGO-Virgo-KAGRA collaboration's GW events. However, it could be an issue in the future, and certainly for XG detectors which will be more sensitive to noise variability and observe hours-long signals, breaking the assumption of stationarity. As such, future methods for detecting and interpreting GW signals should account for the variable nature of the detector noise.

2.2 Noise Transients or Glitches

Transient noise artifacts, also known as glitches, are also a common problem in interferometric GW detectors. Glitches can mask or mimic a signal and add to the noise background of transient GW searches (see, e.g., [42,43,56]). Glitches occur frequently in all detectors; in the third observing run, the rate of glitches was between 0.29 to 0.32 per minute for LIGO-Hanford, 1.10 to 1.17 per minute for LIGO-Livingston and 0.47 to 1.11 per minute for Virgo [38]. The inferred population properties of glitches have been shown to typically exhibit characteristics similar to CBC signals with large mass ratios and large spins, in contrast to the observed astrophysical properties, which tend to have near equal masses and moderate spins [57]. This is because CBC signals with large mass ratios and large spins can have more 'irregular' waveform morphologies compared to equal-mass, non-spinning CBCs from the twisting-up effects due to precession. Therefore, this class of signals has a better chance of fitting well with the terrestrial disturbance produced by glitches that lack a CBC signal's typical chirping-up characteristics.

The morphology of glitches, in particular their time duration and the frequency space they affect, can be highly variable between different glitch classes. For example, blip glitches (e.g., [58]) are fractions of a second in duration, covering a large bandwidth (e.g., tens to hundreds of Hz) and can mimic a GW signal of high mass compact binaries. We still do not know the origin of these types of glitches as they do not have a known environmental or instrumental coupling, but they appear to have different subcategories that may be caused by different physical mechanisms. In the third observing run, these types of glitches occurred 4 times per hour at LIGO-Livingston and twice per hour at LIGO Hanford [43]. However, scattering glitches (e.g., [59]) caused by microseism noise, can be a few seconds long, and present as arches in the time-frequency plane, affecting frequencies below 100 Hz. These glitches manifest due to a small fraction of laser light scattering off a test mass, hitting a moving surface, and recombining with the main beam. These types of glitches are most prevalent when the ground motion is high. As such they can seriously contaminate hours of data, but not be a concern for weeks at a time.

Tracking the occurrence and emergence of new glitch types can be a challenge. Both LIGO and Virgo take advantage of machine learning frameworks, combined with citizen scientists, to classify glitches based on their morphology in the time-frequency plane. GravitySpy [60] has been in operation since the second observing run, and citizens have helped to classify LIGO glitches into 23 distinct classes [43]. GWitchHunters [61] helps to classify glitches from the Virgo detector, and has been open to the public since November 2021. Both projects will prove extremely valuable in identifying and understanding glitches in the fourth and future observing runs.

Glitches overlapping or being in the vicinity of a real GW signal can be a huge problem. In fact, in the third observing run 24% of GW events had a glitch within the analysis window for one or more detectors [38]. These glitches did not impact the detection of these events, but they had to be mitigated before the source parameters could be accurately estimated. A prime example of this issue first arose in the interpretation of GW170817 where a short glitch occurred 1.1 seconds before the coalescence of the event, lasting only 5 ms [62]. Nonetheless, this noise had to be removed before the parameters of the event could be accurately determined. Macas *et al.* [63], for example, shows that certain types of glitches can cause the sky localization to be incorrectly determined for certain types of signals, which can even affect follow-up with large field of view telescopes (i.e., 20 $\deg^2$).

There are a number of ways in which noise can be removed or subtracted from the data. Should the noise be broadband in origin then noise subtraction over the course of hours or days is needed. This can be achieved using auxiliary channels which monitor noise sources at different points around an interferometer. A coupling function can then be determined to understand how much a certain type of noise affects the GW channel, and the noise subtracted [64, 65]. This method is optimal when the data are Gaussian and stationary. More recent work has focused on machine learning techniques to cope with data with non-stationary noise couplings [66].

For short instances of transient noise that may be in the vicinity of an event, there are a few methods which are currently used. A window function can be applied to zero out the glitch; this method is known as gating [49, 67]. Gating has the benefit of being quick, however uncontaminated data will also be removed using this method, as the window function needs to be smoothly applied to avoid adding filtering artifacts to the data. Hence, this method is not appropriate if the glitch is not well localized in time and is close to an event's coalescence time. A more robust method is to model a glitch with a time-frequency wavelet reconstruction and use this to subtract it from the data; this method is applied using the BayesWave algorithm [68]. This method has been used to great effect in the third observing run [38]. Another method, called gwsubtract, uses data from an auxiliary witness to the noise to subtract the noise from the GW channel [64, 69]. This was done for the first time around the event GW200129 [38], which seems to exhibit characteristics consistent with spin induced orbital precession [70]. However, Payne *et al.* [71] find that residual data quality issues leftover from this cleaning process may be the origin of the precession observed in GW200129. Moreover, in a ringdown analysis of GW200129 [39] found a deviation from GR in the peak of the GW amplitude while employing a nonprecessing `SEOBNRv4HM_PA` model [72–74] but they ascribe it to waveform systematics (modeling of spin precession) or data-quality issues (glitch mitigation procedures). Regardless, this example of GW200129 highlights the complexities and care that need to be taken when removing glitches from GW data and interpreting results from inference analyses.

Glitches will always remain a feature of GW data because as the detector sensitivity improves noise artifacts that were sub-dominant will become more relevant. It is unfeasible to remove them all. New methods are being developed to effectively deduce both source and population parameters by integrating realistic but imperfect data. For example, Ashton *et al.* [75] uses Gaussian processes to replace the traditional GW likelihood. This method, in principle, can model arbitrarily colored noise, non-stationarity, and glitches, to augment the approach to estimate the parameters of sources. In addition, Heinzel *et al.* [76] presents a method for inferring the population of GW sources contaminated by blip glitches. They are able to infer the shape parameters of a GW population, whilst simultaneously inferring the population of the glitch background events.

In order to be confident that a signal is indeed a violation of GR, characteristics that may arise due to the noise identified here need to be understood. Work has started in this regard, for example with [77]. They investigated how an overlapping binary black hole signal with three different glitches can affect tests of GR before and after the glitches were mitigated. Moreover, they only considered a glitch in a single detector out of three and still found a GR deviation when the glitch was not mitigated. The authors also point out that their study is not sufficient to give quantitative statements about the effects of certain glitch classes or mitigation methods on tests of GR. Therefore, their work needs to be extended to assess the amount of GR deviation in different realizations of Gaussian noise, the effect of non-stationarities in the noise background, and the effect of data cleaning methods on mimicking GR deviations.

2.3 Contamination from Overlapping Signals

As the sensitivity of ground-based GW detectors improves, the chances of observing time-overlapping signals will also increase [78]. This may demand a shift in our detection and parameter estimation strategies since current pipelines, designed for single GW signals, may yield biased results when applied to overlapping signals. However, several studies have shown that the detection [79, 80] and parameter estimation [78, 81–83] of overlapping signals are not a significant concern. For example, [79] and [80] showed that it is possible to detect and discern overlapping signals from binary neutron stars using a matched filtering algorithm. Relton *et al.* [84] conducted a more thorough study with both modeled and unmodelled search analyses and found that both analyses can detect overlapping signals from binary black holes when merging > 1 s apart. However, unmodelled analysis can identify overlapping signals merging within < 1 s while modeled analysis can only identify only one of the two overlapping signals. Himemoto *et al.* [83] thoroughly explored the parameter space and concluded that overlapping signals do not lead to large biases in parameter estimation provided the coalescence times and redshifted chirp masses of the two overlapping signals differ by at least 10^{-2} s and $10^{-4}M_\odot$ for binary neutron star mergers and 10^{-1} s and $10^{-1}M_\odot$ for binary black hole mergers, respectively.

Nonetheless, overlapping signals do pose biases in both detection and parameter estimation of sources and methods have been proposed to correct those biases [85–88]. For example, Wu & Nitz [85] pointed out that overlapping signals reduce the search sensitivity by changing the noise's amplitude spectral density and proposed an updated search campaign on overlapping signals using the single-detector signal subtraction method. Johnson *et al.* [89] pointed out that the presence of overlapping signals may require us to revisit the definition of the likelihood as well as the assumption that source confusion can be treated as stationary Gaussian noise. Possible remedies to the bias in source parameter inference have been suggested, either from a Fisher Matrix study [86] or adapting the signal model accordingly in the Bayesian likelihood [87]. Langendorff *et al.* [88] used normalizing flows as an avenue to deal with the computational burden coming from multiple-signal analyses in case of overlaps.

Moreover, Hu & Veitch [90] studied the effects of waveform inaccuracy and overlapping signals on tests of GR and demonstrated that when combining results from multiple signals with overlaps, the deviation from GR decreases when waveform models are perfect (no waveform inaccuracy), but inaccurate waveform modeling can lead to a false deviation of GR for overlapping signals. Dang *et al.* [91] extended this study to higher post-Newtonian (PN) deformation parameters. They concluded that although a non-negligible number of overlapping signals can lead to false GR violations at the individual event level, when the results are combined, the biases tend to smoothen out, leading to a preference for GR at the population level inference. (We discuss the effects of population-level analyses on tests of GR in more detail in Section 4.4.)

All these studies focussed on overlaps arising in the data of XG detectors, since the probability of observing overlapping signals in the era of A+ sensitivity [16] or Voyager [92] is very small [78]. However, it is likely that a quiet GW signal below the detection threshold is present along with the dominant GW signal in the data [93]. This will not pose a problem for estimating individual source parameters, but issues may arise when combining multiple signals, where sub-threshold events collectively act as background or confusion noise [94, 95]. Although [94, 95] considered signals in the XG era only, we might need to consider the effect of a confusion-noise-like background in O5 or A$^\#$ era in the context of testing GR. Moreover, quieter signals may result in imperfect subtraction of the GW model from data when following the definition of likelihood to infer source properties under the assumption of stationary, Gaussian noise. Consequently, combining results across multiple signals to infer population properties could gradually accumulate biases from each single-signal analysis, potentially mimicking noise properties [89] and introducing deviations from GR.

2.4 Gaps in the Data

The data we expect to collect from XG detectors is likely to contain gaps, due to loss of lock at the inferometers that could be caused by a plethora of instrumental or anthropomorphic reasons. The sensitivity band of current detectors is such that GW signals are in the band for about 30 minutes at most. The likelihood of a data gap in such a short window is small, and if it occurs, it is likely to decrease the SNR significantly, since the recovery time (for the instrument to reacquire lock and start data taking again) is comparable to the signal duration. This scenario changes drastically with XG detectors because the low-frequency sensitivity is greatly increased, allowing for the observation of signals for many hours to days. The likelihood of a data gap in this window is larger, and if it occurs, it is likely to both decrease the SNR of the event and deteriorate the analysis of the GW source.

Not much work has gone yet to study the effect of data gaps in XG detectors, but some work already exists for data gaps in space-based detectors, from which we can extrapolate some conclusions. Previous work has shown that data gaps can deteriorate and bias parameter estimation for certain sources [96, 97], in particular when the data gap coincides with the merger phase. In general, we would expect that a data gap during the merger would inhibit our ability to constrain deviations from GR at high PN order, while gaps in the early inspiral will be the same for low (or negative) PN order modifications to GR. In particular, if the data has a gap, but our analysis does not account for it, parameter correlations between non-GR and GR parameters are likely to introduce biases that may lead to a false GR violation. Certain methods, such as Bayesian data augmentation [98], however, can be used to include missing data periods as auxiliary variables when sampling the posterior distribution of model parameters that have shown promise at eliminating biases.

2.5 Detector Calibration Error

The GW strain data d are not directly recorded by the interferometer. Instead, it is reconstructed from the voltage $v(f)$ measured by photodetectors and a response function $R(f)$ that relates the digital readout and GW strain, i.e., $d(f) = R(f)v(f)$ [99]. The calibration process includes a series of measurements to construct a reference model for the response function [99–101]. Bias in any step of this process can lead to errors in the measured strain data, and systematic errors in parameter estimation could arise if the calibration error is not taken into account. Vitale *et al.* [102] investigate the consequences of calibration error in Bayesian inference of source parameters. By comparing the ratio between the calibration systematic errors and the statistical uncertainties, their results show that the parameters that suffer the largest biases are those mostly related to the amplitude of GW signals: on average, calibration systematic errors are of $\mathcal{O}(1/100)$ of the statistical uncertainty for sky location parameters and $\mathcal{O}(1/1000)$ for masses. This is potentially because phase errors are localized in frequency and do not accumulate over the inspiral. It implies that calibration errors could have a minor effect in parameterized tests of GR that modify the phase of waveform. Vitale *et al.* [102] also conclude that $< 20\%$ of amplitude calibration error or $< 10 - 20^\circ$ of phase calibration error should not lead to significant biases for all but the strongest signals in the advanced LIGO era, consistent with [103] and [104]. Furthermore, they report that the calibration systematic error is not strongly correlated with SNR as the calibration affects both noise and signal. However, whether such a level of calibration systematics is tolerable in the XG era where SNR can reach $\mathcal{O}(1000)$ is worth investigating. It is still of great importance to improve calibration techniques along with the high sensitivity in the XG era [44, 105].

It is possible to quantify and mitigate calibration errors in detection and data analysis. The uncertainty of the response function can be indicated by the photon calibrators which apply a known radiation pressure directly on the test masses within the detector [99,106–108]. Abbott *et al.* [109] reported $< 10\%$ calibration uncertainty in the strain amplitude and $< 5^\circ$ in phase during the first LIGO-Virgo observing run , and in the third observing run these uncertainties were reduced to $< 7\%$ and $< 4^\circ$, respectively [110]. Note that these are overall uncertainties and systematic errors alone are even smaller. These estimates of calibration uncertainties are used as priors to marginalize uncertainties in GW strains during parameter estimation, which effectively mitigates the calibration error [111, 112]. However, this technique might conceal tiny deviations from GR, since it marginalizes over some level of uncertainties on amplitude and phase. Hence, the effect of calibration errors on tests of GR needs to be studied for current and future GW detectors, so that it can be ruled out (or included) as one of the possible causes for false GR violations.

3 Waveform Systematics

3.1 Missing Physics in Waveform Models

The current state-of-the-art waveform models used in tests of GR still lack certain physical effects, such as eccentricity of the binary's orbit, overtones, and non-linearities in the ringdown phase of the binary merger, etc. Including each of these known physical effects individually is crucial for precision GR tests, but their collective inclusion is essential for unbiased assessments of GR. Here we discuss missing physical effects that could lead to a false GR violation.

3.1.1 Eccentricity

The eccentricity of a binary's orbit depends on the formation history of the binary. Binaries formed through isolated formation channels in the galactic field are expected to have negligible eccentricity when observed in the frequency band of ground-based detectors, whereas binaries inside dense stellar environments such as globular clusters and nuclear star clusters might have moderate to high eccentricities when observed by these detectors. In an isolated formation channel [113], the binary goes through various mass transfer episodes between its components, and as the components evolve and undergo supernova explosions, the binary orbit could gain some eccentricity due to supernova kicks. However, due to the emission of gravitational radiation [114, 115] the binary's orbit shrinks, and the binary sheds away all its eccentricity over the long inspiral, leaving it with negligible eccentricity close to merger [114]. For example, if a binary with an initial orbital eccentricity of 0.2 emits GWs whose dominant mode has a frequency of 0.1 Hz, the eccentricity reduces to $\sim 10^{-3}$ when it reaches a dominant mode GW frequency of 10 Hz. That is why binaries detected by LIGO/Virgo are expected to be quasi-circular. On the other hand, a fraction of dynamically formed binaries can still have some eccentricity (and as high as ~ 1 at 10 Hz) when observed in the frequency band of the LIGO/Virgo detectors [116–124]. Further, environmental effects such as accretion and dynamical friction can also increase the eccentricity of binaries [125].

The problem of misinterpreting eccentricity as a potential GR violation is currently a twofold problem. First, of missing physics; namely, the inclusion of both eccentricity, argument of periapsis (although see [126]), and precession in an inspiral-merger-ringdown waveform model. Distinguishing eccentricity from precession without waveforms that include both [127] introduces systematic biases in the estimated binary parameters [128–132] that could be misconstrued as false violations of GR [133–137]. Second, the current analysis methods are producing inconsistent results [126,128,129,138–140] for the same events such as GW190521 [141].

Once the above two problems are solved, the problem of eccentricity reverts back to being one of waveform systematics discussed in more detail in Section 3.2.2 below. We anticipate larger waveform systematics in systems with higher eccentricities. However, these are not the ones for which eccentricity will manifest as a violation of GR, due to the large-amplitude modulations that are inconsistent with a quasi-circular inspiral.

3.1.2 Tidal Effects

Neutron stars and their mergers are characterized not only by strong gravity but also by extreme matter conditions. To explore how matter affects the space-time deformations around these stars, we need to understand the relation between the dynamical properties of matter and the behavior of strong gravity. Analytic methods are used to model the early inspiral phase of a neutron star binary merger, where neutron stars are approximated as massive point particles with small corrections due to finite-size effects [142–144]. However, close to the merger finite size effects become significant and numerical relativity (NR) simulations are required to capture them accurately [145–148]. Effective one body models achieve a nonperturbative resummation of the PN information on tidal effects into a complete framework [145, 149–154]; some reduced-order-model versions incorporate NR-calibrated tidal models [148, 155, 156] as also used in Phenomenological models.

The tidal deformation of bodies is directly proportional to the Riemann tensor and its derivatives, produced primarily by the energy-momentum distribution of the companion [157], which becomes the second derivatives of the Newtonian potential for the electric-type quadrupole effect in the Newtonian limit. However, such effects are observable in the GWs only if they produce significant mass and current type multipole deformations of the neutron stars in a binary system. The dominant deformations come from the electric-type, $l = 2$ tidal deformation, which imprints primarily in the GW phase evolution. However, it is important to note that these tidal effects are relatively small and become more pronounced as the binary approaches merger. While these effects are subtle, their detection has already provided invaluable insights [62], and with the advent of more advanced detectors (such as XG), we can look forward to even more precise measurements in the near and far future [158–161].

The effects of the tidal field on neutron star matter are studied using observed GWs [2], however, such results are susceptible to waveform systematics and incomplete modeling of neutron star physics. For example, Refs. [162–164] show that the inference of tidal parameters with XG detectors can be significantly affected due to waveform systematics. Not including subdominant tidal effects, such as dynamical tides, which become important in the inspiral regime, can also lead to substantial biases in the estimation of tidal parameters [153, 154, 165, 166]. Likewise, the effects of spins on dynamical tides [167–170], other spin-tidal couplings [148, 171], spin-induced multipole effects [172–175], nonlinear tides [176], higher-order relativistic corrections, and the GW features of tidal disruption in cases with precessing spins [177] are examples of areas requiring further investigations. Further, XG detectors will be sensitive to the octupolar electric and quadrupolar magnetic tidal deformabilities, and not including them in the waveform might bias the measurements [178].

Resonant mode excitations may contribute distinct features in the waveform from the tidal effect considered in [157]. As the inspiraling orbit passes through the frequency of a certain characteristic mode, the resonant excitation of the mode must be compensated by the loss of the same amount of orbital energy, speeding up the following orbital evolution. The excitation of gravity modes [179–181], the interface mode [182–184] and gravitomagnetic mode [185–188] have been studied, where for the latter two cases the phase modulation may reach the level of $\mathcal{O}(10^{-2}) - \mathcal{O}(10^{-1})$ radians in the frequency band of ground-based detectors.

Inaccurate or missing physics in analytical and NR modeling due to thermodynamical transformation of nuclear matter during inspiral and post-merger leads to waveform systematics. Such effects include, but not limited to, viscosity [189–192], thermal effects [193–197], phase transition to hyperon condensates or quark matter and other such transformations (see, e.g., [198–203] and also see Section 4.3.2 for discussion of proposed exotic matter that has not been observed but, may have compactness close to black holes). As shown in [204–206], the viscous effect introduces a new dissipative channel that modifies the GW phase at 4PN order and higher. If not included in the modeling, a signal containing such a 4PN effect could be misinterpreted as a GR deviation at that PN order and at neighboring PN orders.

Similar effects during the post-merger evolution are subject to systematic bias which requires emphasis on accurate post-merger waveform model development. Currently, only a few post-merger models exist and can detect such effects only in the XG detectors [207–211]. There are also sources of bias in parameter estimation that are exclusive to data analysis challenges arising from noise systematics. For a minority of events, multiple overlapping signals and confusion background created by CBC mergers could potentially lead to a bias in tidal deformability as described in Section 2.3.

Additionally, GR predicts relations between the spin-induced quadrupole moment and the (quadrupolar, electric) tidal deformability [8, 212–214] and between tidal deformabilities of different multipolar order and parity [215] or between different tidal parameters in gravitational waveforms for binary neutron star mergers [216, 217] which are only mildly sensitive to the neutron star equation of state. These relations have been used in GW data analyses to reduce the number of search parameters [218, 219] but small equation-of-state variation in these relations can induce systematic biases. One could, however, use constraints on nuclear physics from neutron star observations available at the time to keep updating and reducing the amount of variation in the relations. For example, such variation has been reduced by 50% after GW170817 and current systematic errors on the tidal deformabilities are subdominant than statistical errors until the A$^{\#}$ era [220]. Another way to reduce systematic biases due to the variation in quasi-universal relations is discussed by [159]. It should be noted that the subpercent accuracy in the universal relations will become important in deducing the correct equation of state and hence in the tests of GR for the sensitivities corresponding to XG detectors. Since alternative theories predict different relations, an independent measurement of the quantities in the universal relations can therefore be used as null tests of GR, circumventing potential degeneracy with unknown nuclear physics [212–214, 221–223]. While the spin-induced quadrupole moment is expected to be small for neutron stars, the magnetic tidal deformability could be measured by XG detectors [178] and might need to be included in the waveform models.

Besides testing GR, these relations can be used to disentangle source misidentification (discussed in detail in Section 4.3.2), since each model of exotic compact objects other than neutron stars would display their own quasi-universal relation [222, 223]. Notably, the tidal deformability parameter may carry information about the nuclear equation of state and hence offer a unique tool to distinguish conventional neutron stars from the ones with exotic signatures. Analyzing binary neutron star mergers with exotic matter while using waveforms of conventional neutron star binaries could lead to false indications of GR violations. This needs to be investigated thoroughly, so that this effect could be ruled out or observed.

Assuming that our NR-assisted waveform models are accurate and free of systematic biases including those arising from the unknown equation of state, any deviation from the predictions will be indicative of either GR not being the complete theory of gravity or deviations in the coupling of matter to gravity, a subset of which is the test of the strong equivalence principle [224–229]. Therefore, only after ruling out the systematic effects arising from these inaccuracies, robust conclusions can be drawn about deviations from GR.

3.1.3 Kick-induced Effects

The anisotropic emission of GWs during a CBC carries away linear momentum and results in a recoil or *kick* of the merger remnant [230, 231]. The kick leaves the following imprints in the GW signal: the Doppler effect [232] and the aberration effect [233] on the post-merger signal along with an additional contribution of a (linear) memory effect [232] to the whole GW signal [234]. Since the black hole kicks are non-relativistic, the kick-induced effects are small and might not be important for current GW detectors but could be crucial for XG detectors [234, 235]. For loud ringdown signals (SNR$\gtrsim$ 100, [235]) in the XG era, these kick-induced effects, if not accounted appropriately in the waveform model [236, 237], might contaminate those tests of GR that depend on the post-merger signal and kick [238] of the remnant (see, e.g., [39, 234, 239–242]).

3.1.4 Beyond Fundamental Modes in Ringdown Signal

The gravitational radiation from a perturbed black hole is in the form of quasi-normal modes [243, 244]. At sufficiently late times following a binary black hole merger, it is expected that the remnant can be very well approximated by a perturbed Kerr black hole. Moreover, it is well known that the radiation at this stage is dominated by just the fundamental quasi-normal mode, since it is the slowest damped quasi-normal-mode (QNM) [245–247]. The frequency and damping time of a mode are in one-to-one correspondence with the remnant mass and spin. In principle, assuming GR and using NR simulations, the latter quantities could be predicted from the properties of the progenitor binary, which can be extracted from the premerger signal. In practice, waveform systematics in the premerger phase could jeopardize this ringdown consistency test [248]. For example, large unmodelled eccentricity could lead to an inconsistency in the final mass and spin, and hence to a false GR deviation [135]. In the spirit of the original black-hole spectroscopy program [245–247, 249], it is therefore better to test GR using ringdown signals only, and an "agnostic" selection of multiple modes to model the ringdown [250].

Recently, there have been efforts to increase the range of validity of linear perturbation theory by modeling the early postmerger signal using overtones and mirror modes [250–262]. These studies show that the inclusion of these additional QNMs improve the remnant mass and spin estimates using a ringdown model. They also show that there will be biases in the remnant parameters if a ringdown model is used to describe early postmerger without the inclusion of such QNMs. Such biases in parameter estimation can show a deviation from the predictions of GR. Isi and & Farr [263] investigated the impact of an incomplete ringdown model on parameter recovery by analyzing a synthetic signal mimicking a binary black hole ringdown (see also [250] for a discussion). Their findings reveal biased parameter measurements in instances of very high ringdown SNR. Dhani & Sathyaprakash [255] displayed the modulations in the odd-m modes in the waveform and how the inclusion of mirror modes in the ringdown waveform model can explain these modulations.

The BH spectroscopy program in GW literature aims to test the Kerr nature from the observed QNM spectrum. These tests are typically referred to as "no-hair" theorem tests too. However, since the tests are based on QNMs as the only observables, they are not sensitive to the type of BH hair — namely, primary hair[1] or secondary hair [264]. Therefore, any modification to the Kerr QNM spectra would fall under these tests. There are claims in the literature that overtones have been detected [265–267] and used to test the "no-hair" theorem with GW150914 [241]. However, there is a disagreement in the literature regarding the significance of the measurement of the first overtone in GW150914 [266, 268–271]. There are also theoretical arguments suggesting caution in the use of overtones for no-hair theorem tests [250, 259, 272–275]. The above authors show, using toy models, black hole perturbation theory and NR simulations, that even though the estimates of the final mass and spin of the black hole can be improved starting the ringdown analysis at earlier times by the addition of overtones, a linear model including only overtones is not appropriate at early times (see also [276]). Therefore, they contend that overtones are unphysical and that their role in a waveform model is to "fit away" other features in the signal, namely, transients related to the initial data, power-law tails at late times, and nonlinearities.

However, for less symmetric binaries than GW150914 (as commonly expected among current and future catalogs) the original black-hole spectroscopy program can be realized using higher-order modes in addition to the least damped QNM, i.e., $(l, |m|) = (3, 3), (2, 1), (4, 4)$, can be used to perform independent tests of the no-hair theorem [242, 270, 277–283]. Given current estimates of the merger rates, XG detectors are predicted to perform percent-accuracy tests for a few events per year [278, 283–285].

To conduct any of the above tests of GR using the perturbative ringdown model, one must make a choice on the start time of the ringdown to begin fitting exponentially damped sinusoids. The analysis should begin as soon as the perturbative prescription is relevant. On one hand, waiting too long to begin the analysis will make testing GR impossible because the strain amplitude has decayed exponentially (e.g., [286, 287]). However, beginning the analysis too early could result in overfitting to non-linear features in the signal (e.g., [250, 288]). To undertake robust tests of GR, some criterion for the analysis start time should be established through, e.g., searching for the earliest time at which one can measure self-consistent QNM

---

[1] In this context, primary hair refers to extra charges that are independent of the BH mass and spin (e.g., the electric charge in the Kerr-Newman solution), whereas secondary hair refers to extra charges that are fixed in terms of the mass and spin.

parameters with time [259, 260, 262]. A further source of systematics is the decomposition of QNMs in spherical rather than spheroidal harmonics; if unmodelled, the spherical-spheroidal mode mixing introduces biases for highly spinning remnants [250].

Another important effect of the nonlinearity in the ringdown stage is the presence of second-order QNMs [289–291], which are generated through mode-mode couplings. The frequency of a second-order QNM is twice as the associated "parent" linear QNM. Its amplitude and phase are also uniquely determined by the linear mode [292–294], as a nontrivial prediction of GR at the nonlinear level. The dominant nonlinear modes may be observable with XG detectors, although event rates are uncertain [295].

An approach complementary to null tests using QNM frequencies and damping times is to test QNM amplitude-phase relations predicted by NR simulations within GR. This test was successfully applied to GW190521 in [296], finding that measurement errors for this event are still large, but would strongly improve for the louder detections routinely expected for XG detectors.

Finally, because of its short duration, one should be careful with the statistical methods and their underlying assumptions while analyzing the ringdown signal. Seemingly innocuous data processing choices such as the uncertain starting time, duration of the signal, and noise estimation techniques can lead to materially different inferences [241, 268, 269, 297–299]. While the ringdown signal is typically analyzed in the time domain, frequency domain methods have also been proposed [257, 269, 300, 301] with the approach of [300] shown to be formally equivalent to the time-domain approach [263]. Even then, [300] comes to a different conclusion regarding the ringdown of GW190521 compared to [4] or [302]. This highlights the need to better understand systematics and data analysis techniques in the analysis of ringdown signals.

3.2 Inaccurate Modeling of Known Physics in Quasi-Circular Waveform Models

3.2.1 Higher-order Modes, Precession, and Memory

Gravitational waveforms can be decomposed in the basis of spin weighted spherical harmonics with spin weight $s = -2$, $Y^{lm}_{-2}(\iota)$, where ι is the inclination angle. In this basis, for nonprecessing systems, the dominant contribution to the GW amplitude comes from the $(l, |m|) = (2, 2)$ harmonics. The $(2, 1)$ and $(3, 3)$ harmonics are subdominant and suppressed by a prefactor that goes to 0 for symmetric (equal mass) binaries [303–307]. These modes only contribute for systems that are not face-on/off ($\iota \neq 0, 2\pi$), and become particularly important for unequal mass binaries. The presence of these higher-order modes causes characteristic modulations in the amplitude and phase of the waveform.

The effect of higher-order modes becomes even more important in the presence of spin-induced precession. Spin-induced precession occurs when the spin angular momentum vectors of the binary components are not aligned with the orbital angular momentum vector, leading to the precession of the orbital angular momentum (or, equivalently, the orbital plane of the binary) as well as the spin vectors about the total angular momentum of the binary. The effect of precession is best understood by considering two frames of reference [308–310]—the *inertial* frame in which the binary appears to be precessing, and the *co-precessing* frame that follows the instantaneous motion of the orbital plane where the effects of precession disappear. The inertial modes can then be approximately described as the sum of nonprecessing modes with the same l value and all possible m values, each rotated using Wigner D-matrices which depend on the instantaneous position of the orbital plane [311]. Thus, due to spin-induced precession, subdominant precessing modes will have contributions from both dominant and subdominant nonprecessing modes, increasing the precession effect due to the presence of higher-order modes in the waveform [312].

A consequence of using nonprecessing modes to approximate the co-precessing-frame signal is that these obey the reflection symmetry $h_{\ell m} = (-1)^{\ell} h^*_{\ell -m}$, which no longer holds for precessing binaries [313, 314]. Most state-of-the-art waveform models, with the exception of `NRSur7dq4` [237] and `IMRPhenomXO4a` [315, 316], currently rely on this approximation. While the impact of anti-symmetric contributions to the waveform modes is typically small, neglecting these effects could result in biased measurements of the spin magnitude and orientation at high SNR [317, 318].

Currently, state-of-the-art nonprecessing waveforms like `IMRPhenomXHM` [319] include the harmonics $(l, |m|) = (2, 1), (3, 3), (3, 2), (4, 4)$, and `SEOBNRv5HM` [320], in addition to these, also includes $(l, |m|) = (4, 3)$ and $(5, 5)$. Their precessing counterparts are `IMRPhenomXPHM` [321] and `SEOBNRv5PHM` [322], respectively. The widely used NR surrogate waveform model, `NRSur7dq4`, has been trained with simulations with mass ratio less than 4, and contains all spherical-harmonic modes with $l \leq 4$.

Many studies have explored the improvement in the inference of source parameters due to the inclusion of spin-induced orbital precession and higher-order modes [323–326]. Particularly, for edge-on systems, including higher-order modes improves parameter estimation by breaking the luminosity distance-inclination angle degeneracy, whereas modulations due to spin-induced precession break the degeneracy between the spin and mass parameters. Additionally, the amplitude of the higher-order modes also brings information about the mass ratio of the source.

We should note that none of these models discussed above contain the memory modes that depend on the binary's past history. The most well-known of these is the displacement memory effect which is dominant in the $l = 2, m = 0$ mode, and the next leading memory effect, known as the spin memory, is dominant in $l = 3, m = 0$ mode for the non-precessing binaries (see e.g., [327] and [328]). There are other higher-order memory effects, but these can be extremely sub-dominant. Most of these are discussed in [329] and references therein. While these are small effects, they will need to be included to prevent biases, and have now been included in a surrogate model for nonprecessing (quasicircular) binary black holes constructed using the waveforms obtained from Cauchy-characteristic evolution [330]. The effect of non-linear memory on the binary black hole parameter estimation is studied in [331] where the dominant displacement memory in the $l = 2, m = 0$ mode starts to affect the parameter inference at SNR > 60 for the current generation ground-based detectors (such as LIGO A$^{\#}$). Moreover, the effect of memory has been studied in the case of neutron star-black hole and binary neutron star mergers [332, 333], where it is argued that the memory can affect parameter estimation for the XG detectors. Studies show that it will be difficult to detect the presence of memory in individual sources with the current LIGO, Virgo, and KAGRA detectors at O4 sensitivity or even O5 sensitivity, but it could be detected in a population using the stacking procedure (e.g., [334]). Thus, it is necessary to understand the effect of memory on parameter estimation and tests of GR at the population-level.

Therefore, analyzing a GW signal that has a significant magnitude of spin-induced precession, higher order mode content, and memory effect with an inaccurate or incomplete waveform model may not only deteriorate parameter estimation, but also show biases in the inference of other source parameters (see, e.g., [312]). A recent study has investigated systematics due to waveform mismodeling by comparing `SEOBNRv5PHM` and `IMRPhenomXPHM`. It was found that systematic biases can impact the current and future GW-detector networks, affecting the inference of realistic binary black hole population properties, as well as, the science cases of individual loud signals [248], and more in general binaries with large mass ratios and high precession. Such systematic biases may eventually find their way into the measurement of a beyond-GR parameter depending on the nature of its correlation with the other source parameters, inducing a false violation of GR. Hence, it is essential to use accurate waveform models with spin-precession effects, *sufficient* number of higher-order modes, and memory effects while testing a GW signal for a violation of GR.

3.2.2 Sub-optimal Calibration and Agreement With NR Waveforms

State-of-the-art waveform models are built by combining and resumming information from different analytical methods, such as PN approximation and gravitational self-force theory, and then calibrating/validating against NR simulations and merger-ringdown waveforms in the test-particle limit, which are obtained by solving the Teukolsky equation. The assessment of the accuracy of the waveform models from the two main waveform families (notably `EOB` and `IMRPhenom` models) can be found in [248, 315, 320, 322, 335–337]. Due to the number of calibration parameters and the large number of NR simulations at disposal, it is especially important to devise a computationally efficient and flexible calibration procedure. For instance, in calibrating the `SEOBNRv5HM` model [320], the authors quantified the agreement with NR waveforms in a Bayesian fashion and employed nested sampling to obtain posterior distributions for the calibration parameters. State-of-the-art waveform models use best-fit estimates across the physical parameter space for their calibration parameters. Providing instead a probability distribution, modeled for example through a multidimensional Gaussian mixture, would allow accounting for uncertainty estimates due to sub-optimal fits, and could mitigate waveform systematics at high SNR [338]. Other proposed methods to marginalise over waveform modeling uncertainties include Gaussian process regression [339–342], or introducing frequency-dependent amplitude and phase corrections, as in the case of detector calibration uncertainty [164]. While these methods may obscure small deviations from GR, particularly around the merger phase, significant deviations that exceed the estimated modeling uncertainties should still be detectable.

Calibration parameters typically enter in waveform models as higher-order PN coefficients, which are currently unknown. Including higher-order analytical information, while pushing the calibration parameters at even higher orders, could improve the accuracy of current waveform models, but requires careful studies on how to incorporate and resum this information [320, 335, 343] Nonetheless, neglecting higher-order PN terms carries an error which might become relevant with updates to current detectors and XG detectors, but could be mitigated by marginalizing over higher-order PN coefficients as new model parameters [344]. Incorporating results from the post-Minkowskian (PM) approximation [345–348], a weak fields expansion in G at all orders in the velocity, is also promising, particularly for highly eccentric binaries for which relativistic velocities can be reached at each periastron passage even in the weak field regime. While PM results have not yet been incorporated in state-of-the-art waveform models for bound orbits, remarkable agreement has been obtained comparing PM-improved EOB models to NR for scattering orbits [349–352].

The calibration procedure imposes that the waveform model agrees, as much as possible and for the entire coalescence, with the NR waveform. This is often quantified by computing the unfaithfulness (or mismatch) $\mathcal{M}$ between the model and NR waveform. As detectors become more sensitive and the SNR increases, the accuracy requirements become more stringent, thus demanding smaller unfaithfulness values. Accuracy requirements are usually formulated in terms of an indistinguishability criterion [353–357], which states that if two waveforms fulfill the condition

$$\mathcal{M} < \frac{D}{2\,\mathrm{SNR}^2}, \tag{1}$$

for a given power spectral density (PSD) and SNR, then these waveforms are considered indistinguishable, and differences in the recovered parameters are expected to be smaller than statistical errors. Here D is an unknown coefficient, usually set to the number of intrinsic parameters of the source [356] or tuned with synthetic injections at increasing SNR [357]. Being sufficient, but not necessary, this criterion is generally too conservative, and, if it is violated, differences are not necessarily measurable, or may appear in a subset of parameters in which one is not typically interested [357, 358]. Toubiana & Gair [359] recently proposed a correction to the standard indistinguishability criterion by revisiting some of the hypotheses under which it is derived, and employed it to quantify apparent deviations from GR due to waveform inaccuracies [360].

The state-of-the-art multipolar, aligned-spin `SEOBNRv5HM` model shows a median unfaithfulness of 1.01×10^{-3} against 442 NR waveforms when using the O5 PSD [361] and considering binary total masses in the range $[20\text{—}300] M_\odot$. Using this model would lead to a false deviation from GR when measuring the QNM (complex) frequencies of a massive binary black hole with a mass ratio of 2, as observed in LISA with an SNR $\mathcal{O}(100)$ [360]. This issue occurs because for such massive binary black holes, the majority of the SNR lies in the merger-ringdown stage. By contrast, a stellar-mass binary black hole with mass ratio 6, observable in O5, would not incorrectly lead to a violation of GR at SNR 75 [282], because in this case a large portion of the SNR is accumulated during the inspiral stage. Normally, the accuracy of waveform models gets worse toward merger, where the presence of higher-order modes becomes more and more important, while their modeling is quite challenging. The recent study of [362] investigated the impact of inference biases from sub-optimal waveform calibration on a realistic population of binary black holes in XG detectors. They considered two quasi-circular, nonprecessing waveform models of the same family (namely, `IMRPhenomD` [363] and `IMRPhenomXAS` [364]) and estimated a mismatch requirement of $\sim 10^{-5}$ for 99% of the events with SNR > 100 not to be biased.

Inaccuracies in NR waveforms, due to, e.g., numerical truncation errors and issues with GW extraction and extrapolation, are typically at least one order of magnitude smaller than errors between semi-analytic models and NR [357]. Nonetheless, they are expected to become relevant with updates to current detectors and XG detectors, especially for binaries with asymmetric masses and orbits inclined with respect to the line of sight [357, 365, 366].

4 Astrophysical Aspects

There are several astrophysical aspects of the source, its surroundings, and the emitted GW signal that have not been accounted for in the state-of-the-art waveform models. These aspects, if present in the real GW signal, might affect the tests of GR and can lead to false GR violations. Here we discuss those astrophysical aspects that we can think of.

4.1 Gravitational Lensing

As GW detectors get upgraded and new ones join the network, more and more distant mergers can be observed. This increases the chance of having a matter density crossing the GW travel path, possibly leading to gravitational lensing. Depending on the lens properties and the lens-source geometry, different effects can be observed. For the best-aligned and most massive cases, we are in the geometric optics limit and lensing leads to several copies or "images" of the initial signal. These images have the same frequency evolution but are delayed in time, (de)magnified, and can undergo an overall phase shift. When the time delay is large enough, these images are distinct, and we face strong lensing [367, 368]. For ground-based detectors, typical lenses are galaxies and galaxy clusters [369]. For smaller time delays, corresponding to less aligned systems and lighter lenses, one has millilensing, where the various images overlap and sum to a non-trivial signal in-band [370]. This is expected to be due to heavy black holes, or dark matter over-densities, for example. Finally, when the GW wavelength is comparable to or greater than the size of the lens, we need to perform the full wave-optics treatment [367], and lensing leads to frequency-dependent beating patterns known as microlensing. For ground-based detectors, typical lens sources are individual stars, black holes, or dark-matter overdensities [371]. It is also important to note there can be interplay between these different types of lensing. When strong lensing happens, one or more of the images may undergo micro or millilensing because of individual objects present in the strong lens [372–374].

False GR deviations could be expected when GR signals are distorted. For strong lensing, one can have such an effect for specific values of the overall phase shift. In particular, it can take only three distinct values: 0, $\pi/2$, or π, corresponding to a minimum, saddle point, or maximum of the Fermat potential, and referred to as Type I, II, and III images, respectively [368, 375]. Under all circumstances, Type I and III images are indistinguishable for the GR case because they correspond to no shift or a sign flip in the polarization, which cannot be detected [375]. For Type II images, on the other hand, detectability is possible when the GW displays higher-order modes. In this case, the phase has different pre-factors for different frequency modes and is not degenerate with the (frequency independent) lensing phase shift anymore [375]. This can be used to detect strong lensing based on a single image, although it requires rather large SNRs and very asymmetric, precessing or eccentric systems [375–378]. When analyzing Type II images under the unlensed assumptions, one can face losses in SNR, possibly missing the event with template searches [376], or biases in parameter estimation [377, 378]. Therefore, one can expect this non-trivial feature to also be picked up when searching for GR deviations. For example, this is the case with modified dispersion relations that change the frequency evolution of the GW phase in a way possibly similar to lensing [379]. The link between Type II images and GR deviations is also highlighted in [380], where the authors show that some GR deviations are flagged by Type II search pipelines.

The cases of millilensing and microlensing are even more favorable in leading to spurious GR deviations being detected since they both lead to a non-trivial signal in the detection band, although the nature of the resulting image is different between the two cases [367, 370, 371]. When analyzing such signals with traditional GR templates, one expects imperfect modeling of the signal, leading to coherent power left in the data [381]. This is also confirmed in [382] for some tests of GR. In this study, the authors show that milli and microlensed signals can lead to spurious deviations from GR, sometimes with a high significance. However, it is also important to note that adapted lensing pipelines also clearly see these events as being lensed. Therefore, the GR deviation would probably not be confirmed as it would be explained via lensing, underlying the importance of accounting for possible astrophysical effects on the GW signals when looking for GR deviations. The link between GR deviations and micro and millilensing is also further confirmed in [380], where the authors show that some deviations of GR lead to false positives in micro and millilensing searches. In the case of a multi-messenger lensing event in which the GW lensed signal is in the wave optics regime but the electromagnetic signal is in geometric optics (which is to be expected given their higher frequency), the speed of propagation of GWs could appear to be superluminal due to the waveform distortions [383], although no information actually arrives faster than light [384].

A crucial approximation in these studies is the exclusion of the effect of parallel-transporting the polarization tensor across the lensing geometry and the treatment of GWs as scalar waves which become increasingly violated as one moves from the weak gravity limit. Recent studies [385, 386] have pointed out the consequences of such an approximation and started treating GWs as a tensor field. It is pointed out that there is no notion of a unique "propagation direction" as can be defined in the geometric optics limit as well as the wave optics treatment for a scalar wave. Similarly, strong gravity effects could add extra phenomenology [387].

Therefore, all types of lensing—micro, milli, and strong—can potentially lead to spurious GR deviations being detected if neglected. Hence, should such deviations be seen, it would be crucial to verify possible astrophysical origins of the modification in the GW signal, and in particular if the GW event is not lensed.

4.2 Environmental Effects

The current waveform models can be referred to as *vacuum templates* as they only describe GWs from isolated binary systems in a vacuum environment, neglecting realistic astrophysical surroundings of the source. However, in reality, the binary is always in an astrophysical environment that impacts the binary's orbital evolution and hence results in a GW signal from the binary different than the vacuum template. There are many scenarios in which the GW signal from an environment-embedded binary system could be different from its corresponding vacuum signal. These are, but not limited to, (i) the source resides in a dense environment [388–391] such as dense cores of massive stars [392–394], accretion disks of active galactic nuclei [32, 395–400], and star clusters (see, e.g., [401]), (ii) the source resides in a dark matter halo [32, 402–407], and (iii) the source is immersed in a strong electromagnetic field [408, 409]. Moreover, the peculiar acceleration of the source with respect to the observer, i.e., time-varying Doppler shift [410–413] and the acceleration of the universe, i.e., time-varying redshift [410, 414, 415] itself could lead to GW signals being different from vacuum templates.

The situation where there is a massive bosonic field that is amplified around black holes via superradiance (see, e.g., [416]) is also sometimes considered an environmental effect and can similarly lead to deviations from a vacuum binary black hole signal. However, there are significant differences in this case compared to the environmental effects considered above. Most importantly, in this case the size of the deviation is set by the universal properties of the boson and the properties of the binary, not the specifics of where the binary formed. This makes the deviations more similar to a deviation from GR (which also depends on the properties of the binary and some universal parameters). However, there are many ways to distinguish binaries of black holes with boson clouds from GR deviations. Some of these are discussed in Sec. 4.3.2, since the emitted GW signal in such a scenario will be similar to the one from binaries of black hole mimickers (e.g., there will be tidal effects from the boson clouds). Additionally, since the superradiant growth of the clouds is only possible for certain pairs of black hole masses and spins (see, e.g., [417]), it should be easy to distinguish this case from modified gravity (or black hole mimickers) when considering the population. The time dependence of the tidal deformability and non-black hole multipole moments due to perturbations or even disruption of the clouds due to the effects of the other black hole (see, e.g., [418–420]) should allow one to distinguish the boson cloud case from black hole mimickers even for individual sources. Additionally, one can obtain constraints on the boson mass from the contributions from the superradiant instability to the stochastic background of GWs [421, 422]. Furthermore, boson clouds are expected to emit a nearly periodic and long-duration GW signal [421, 422] and no evidence of such signals is found in current GW data, which provides constraints on the ultra-light scalar boson field mass (see, e.g., [423–426]).

Returning to environmental effects proper, the detailed modeling of different environmental effects on the binary's GW signal is challenging and requires computationally expensive NR simulations [393]. However, in the literature, these effects have been approximated as a correction to the vacuum GW signal's PN phase evolution. For example, at the leading order, dynamical friction due to gas accretion can be modeled as a −5.5PN correction whereas collisionless (collisional) accretion can be modeled as a −4.5PN (−5.5PN) correction [391, 427–429]. The accretion and dynamical friction due to a scalar dark matter cloud give rise to a −4PN and −5.5PN correction, respectively, to the phase at the leading order [430]. Electromagnetic effects have been computed at next-to-leading order (at 3PN) by taking into account the whole electromagnetic structure of a star. The leading magnetic corrections at 2PN order (assuming a constant and aligned magnetic dipole) to the GW phase are found to be comparable to a 1.5PN point-particle effect [431,432]. Phase correction due to the line-of-sight peculiar acceleration of the source has been computed up to 3.5PN order [411,433] while the acceleration of the universe leads to a −4PN correction to the phase at leading order [414,415].

It has been argued that the magnitude of the environmental [32, 434] and cosmological [410] effects are expected to be quite small and hence could be neglected for ground-based detectors. However, there could be scenarios where these effects are non-negligible, e.g., stellar-mass compact binaries would merge around a supermassive black hole and one can still get a significant deviation from the vacuum template in the bands of LIGO/Virgo/KAGRA detectors [433]. Moreover, near supermassive black holes, in galactic nuclei, triple systems of stars are common and they mostly are hierarchical in nature [435–437], i.e., a tight inner binary is orbiting a tertiary on a wider orbit which forms the outer binary. In these *hierarchical triples*, the tertiary brings interesting features to the GW signal emitted by the inner binary, e.g., the oscillation of eccentricity and inclination of the inner binary's orbit due to the Kozai-Lidov mechanism [438, 439]. Such oscillations could modify the frequency evolution of the inner binary and this needs to be taken into account in waveform modeling [440, 441].

A recent study by Santoro *et al.* [442] showed that particularly large environmental effects can significantly bias the parameter estimation if vacuum templates are used for the analysis, even when not directly detectable by LIGO-like instruments. Although this bias requires extremely dense environments that are not predicted by standard astrophysical models, it would be important to find out if such biases in parameters could lead to false GR violations for more sensitive XG detectors.

Likewise, ringdown templates are simple and based on predictions from vacuum GR. Modifications of GR usually lead to extra polarizations or include degrees of freedom with different modes, introducing a simple handle to test for beyond-GR physics. However, environmental effects, such as accretion disks, dark matter halos or any form of matter outside of black holes introduces low-frequency modes or drastic changes to higher overtones, de-stabilizing the spectrum [32, 443, 444]. Concrete examples suggest that spectral instability of the dominant mode introduces changes in the waveform only well after coalescence, but the relevance of overtone instability for time-domain waveforms still needs to be well understood [445].

However, it is worth mentioning that environmental effects will be possibly important only for certain events, while likely negligible for the majority. Thus, any competing beyond-GR interpretation of environmental effects should coherently explain this non-trivial dependence on the source.

4.3 Mistaken Source Class

4.3.1 Beyond Compact Object Mergers on Bound Orbits

Parabolic or hyperbolic scattering [446] as well as head-on collision of compact objects [447–449] may give rise to GW signals which may resemble that of a quasi-circular CBC close to the peak of the signal. Therefore, for relatively short-duration signals, there is a risk of confusing a compact binary merger with one of the above classes of sources, leading to biases on the source parameters and thereby affecting tests of GR. In the case of GW190521, studies have discussed the degeneracy between a precessing compact binary in quasi-circular orbit with a binary that undergoes head-on collision [450] and a merger of two nonspinning black holes on hyperbolic orbits [451]. It is argued that the lack of premerger features in certain precessing configurations in quasi-circular CBC may mimic a head-on collision leading to underestimation of mass parameters and overestimation of luminosity distance when a quasi-circular CBC waveform is employed for parameter estimation. Obviously, such biases will directly affect most tests of GR.

However, precise estimates of final spin can help in distinguishing head-on collision from a quasi-circular CBC. For example, if the inferred remnant black hole spin is high (e.g., ~ 0.7 as was the case for GW190521), this could make the head-on collision unlikely as very special configurations may need to be invoked to explain this. As the head-on collisions are themselves very special configurations, additional requirements such as this (large remnant spin) may weigh down their possibility in a model selection problem. Further, due to the special symmetries of the head-on collision, the spherical harmonic modes excited in a head-on collision may differ from those in a quasi-circular CBC. For instance, unlike quasi-circular CBCs, in head-on collisions $\ell = 2, m = 0$ mode may be as strong as $\ell = 2, m = 2$. Such features may also help in a model selection problem. A dedicated study that looks into the effect of degeneracy between quasi-circular CBC and head-on collision or parabolic/hyperbolic encounters and how that impacts tests of GR will be very useful. To do this we require more accurate analytical or numerical waveform modeling of head-on collision and parabolic/hyperbolic encounters.

4.3.2 Black Hole Mimickers

There are various exotic compact objects that are massive and compact enough that gravitational waveforms from binaries of such objects could be close to those from a binary black hole (see, e.g., [452, 453]). The simplest such objects can be described by GR minimally coupled to a non-Standard Model field (e.g., an ultralight scalar field describing dark matter [454]). More complicated models for such objects involve nonminimally coupled fields, where it may make more sense to treat the additional scalar field as part of the gravity sector. However, even in the case where gravity is still GR, the specifics of the waveform would still differ from that of a binary black hole in GR, and one would thus obtain a false deviation from GR when applying a test of GR based on a binary black hole waveform model. The most theoretically well-modelled such objects are boson stars (see, e.g., [455]), which are formed from a massive complex scalar or vector field, that may be self-interacting, as is necessary to obtain more compact stars (that are thus more similar to black holes)—see, e.g., [456]. However, there are many other models, including quite exotic objects, like gravastars [457], which have an interior made of de Sitter space. A concrete framework for these exotic objects might require GR deviations [458], but they can be described also using exotic matter within GR (e.g., [459]).

For all of these cases, there will be the same matter effects on the inspiral that one finds in the PN approximation for binary neutron stars (some of which are discussed in Section 3.1.2), albeit with different values. In particular, there will be effects of nonzero tidal deformabilities (see, e.g., [456, 459–461]), tidal disruption (at least for sufficiently unequal-mass binaries; see, e.g., [462]), and the excitation of resonant modes in the objects (see, e.g., [463]), as well as effects from multipoles that are different from those in black holes (see, e.g., [173, 464]) and possibly a lack of the relatively large GW absorption (a.k.a. tidal heating) one obtains with black holes (see, e.g., [465]). However, since recent studies [206, 466] have shown that neutron stars can have larger GW absorption than black holes if they have a sufficiently large shear viscosity, it is possible that the same is true for some potential black hole mimickers, though this is not the case for, e.g., standard models of boson stars. There will also be differences in the merger-ringdown part of the signal (see, e.g., for simulations of orbiting binary boson stars [467–470]). If the merger of a binary of exotic compact objects forms an ultracompact object (i.e., an object that has a light ring outside its surface), then the ringdown is nearly indistinguishable from that of a black hole and a train of modulated pulses—known as GW echoes—is emitted in the late postmerger stage [32, 471]. From the analysis of current GW events, no evidence for postmerger echoes has been found with unmodelled and modelled searches [4, 5, 472–477], despite claims of echo detections in [478–481]. Moreover, for perfectly reflecting objects the presence of echoes is disfavored by the current upper bounds on the stochastic background in the advanced LIGO frequency band [482].

If one has a single population of exotic stars that are formed from a single fundamental field, then the non-GR effects in the inspiral will be solely determined by the masses of the objects, and there will be a maximum mass of stable stars, just as in the neutron star case. Thus, if one can measure these effects (and the masses of the stars) accurately (using, e.g., a more refined version of the analysis given in [483]), then one can check if the signals are indeed consistent with coming from a population of binaries of such stars. While alternate theories of gravity with an intrinsic scale will have a roughly similar behavior, where the GR deviation decreases with increasing mass of the black holes, it seems unlikely that an alternative theory of gravity would be able to mimic the situation of exotic stars to a high degree of accuracy. Moreover, if there is a population of exotic binaries as well as binary black holes, then one may observe binary black holes with very similar masses, spins, and distances as the exotic binaries, where a modified theory would predict that one would also observe deviations for the black hole binaries. Thus, while it is likely that the two situations could be confused with initial observations, it should be straightforward to distinguish them with high-accuracy observations. However, the ability of a given set of observations to distinguish specific exotic star models and specific alternative theories would need to be tested with explicit calculations.

For instance, black holes can have nonzero tidal deformabilities in certain alternative theories, such as those that introduce higher-order-in-curvature corrections in the action [460, 484]. However, in such models the dimensionless tidal deformabilities are proportional to inverse powers of the black hole mass, $1/M^n$, where n is a positive integer that depends on the theory ($n = 4$ or 6 in the calculations cited). This is not a good match for the mass dependence of any of the boson star models considered in [460], and while it might be possible to find an exotic star model that gives a better match, the stars would still have a maximum mass, while the black holes in the alternative theory have nonzero tidal deformabilities for all masses. The black holes also have differences in the spin-induced multipoles (see, e.g., [485]) that would also have to be reproduced by the exotic stars, which is unlikely to be possible to more than moderate accuracy. For instance, for some families of boson stars, the spin-induced moments have minimum values larger than their Kerr values (similar to the minimum values of tidal deformability), and show a different spin dependence than one obtains for alternative theories (see, e.g., [486]). Additionally, there will be differences in the GW absorption comparing black holes in this theory and black hole mimickers with no horizon (which will generally have a much smaller GW absorption cross section than black holes). However, one also expects that the GW absorption in such theories will differ from that in GR due to the differences in the static tidal response, given the relation between this and GW absorption/tidal heating (see, e.g., [487]). Moreover, there are also changes to the binary's dynamics that do not come from finite size effects in such theories (see, e.g., [488]), albeit only occurring at high PN orders.

Thus, individual signals from binaries of exotic compact objects could be confused with a GR deviation in many tests (which do not include the expected non-black hole modifications to the waveform). However, binaries of black hole mimickers will in general be able to be distinguished from a modification to GR, even one that predicts nonzero tidal deformabilities for black holes, at sufficiently high SNRs and when analyzing the population of signals, or possibly when performing multiple independent tests of a single signal.

4.4 Statistical Assumptions of Astrophysical Population

Combining information from multiple signals is a powerful method to perform stronger tests of GR. However, assumptions on the underlying astrophysical population and the statistical methods adopted to perform the joint analysis can affect the results.

Biases due to waveform modelling systematics can pile up when stacking multiple events in a catalog. Several studies [90,136,489,490] show that even if systematics are under control at the level of the individual events, the accumulation of biases in a population analysis can produce false deviations from GR if the catalog is large enough. Depending on the actual population of resolved signals and on the way the events are combined, false deviations can appear with as little as ~ 30 events with SNR> 20 in the most pessimistic scenarios [490]. Moreover, restricting the study to golden events with high SNR is even more vulnerable to false deviations once these events become routine in XG detectors [90, 136], although techniques to mitigate the biases have been proposed [489].

Furthermore, combining events requires concrete assumptions about the impact of the astrophysical population and the detectability of GW sources that violate GR. Many parameterized tests of GR infer the presence of expected correlations between individual source parameters (such as the total mass of a binary black hole system) and the deviation parameter [491]. These correlated features within the inferred posterior distributions for individual events imply that specific choices regarding the astrophysical population distribution can skew these results to different regions of the parameter space.

In a recent study, Payne *et al.* [492] demonstrate that neglecting the astrophysical population leads to inferences which are $\sim 0.4\sigma$ less consistent with GR within GWTC-3 for parameterized tests of GR. However, they show that such biases can be mitigated by jointly inferring the astrophysical population properties while combining the distributions of GR violation parameters. Furthermore, Magee *et al.* [493] illustrate that neglecting the loss in detectability of signals with GR violations places constraints on PN deviations that are up to 10% too narrow when ignoring the selection bias in the population. These studies highlight the need to carefully consider the underlying statistical methodologies used when attempting to test GR. In the same vein, astrophysical inaccuracies or biases in the properties of a source population (e.g., imperfect mass distributions) could also lead to false GR deviations. For example, this can happen if events are detected in regions of the parameter space disfavored by astrophysical population models, such as hierarchical mergers from black holes residing in the theorized upper-mass gap [494]. Tests of GR will need to adopt the ever-growing astrophysical population knowledge to remain sufficiently unbiased.

Combining events to test GR also requires assumptions on the GR deviations that are being tested. If the GR modification is common among all the events (as in the case of, e.g., a nonzero graviton mass or a nonzero time variation $\dot{G}$ of Newton's constant), one can multiply the individual, marginalized likelihoods on the deviation parameter to obtain the combined likelihood for the catalog [90, 490, 495, 496]. On the other hand, if the GR deviations are independent for each event (as may be the case if black holes have "hair"), one can multiply the individual Bayes factors in favor of GR to obtain the total evidence from the catalog [90, 490, 495]. In a more general framework where the distribution of GR deviations across the catalog is a known function of the event parameters (such as masses, spins, and compactness), one would need to perform a full Bayesian hierarchical inference on the population [495, 497].

Studies have shown that testing GR at the population level under one of the three assumptions listed above (that all events share the same beyond-GR parameter; that modified theories introduce a new unrelated parameter for each detection; or that GR deviations across the catalog are a known function of the event parameters) can lead to the wrong conclusions if the underlying GR deviation does not satisfy the assumption [495, 497]. Moreover, the accumulation of biases across the catalog due to waveform systematics can change significantly depending on which method is chosen to combine multiple events [90, 490]. Recent work by [498] suggests that performing a full Bayesian analysis should be the most robust approach, but it still requires assumptions that can make the inference inherently model-dependent [495].

As shown by [499], the finite size of the observed catalog will produce cosmic-variance effects that can cause to incorrectly infer deviations from GR, but a bootstrapping technique can be used to mitigate this effect.

5 When Does a Cause Become Important?

Not all effects discussed in this paper are created equal, with some being always important for understanding false GR violations, such as non-stationary noise artifacts and glitches (see Sections 2.1 and 2.2) while some will not be important until XG detectors or beyond, such as unaccounted effects of the physics of gas and dust in the environment of binary black hole mergers (see Section 4.2). In this Section, we gauge when each of these causes will become important in terms of the generation of GW observatory.

It is worth stressing that some level of systematics is unavoidable. For example, waveform models are intrinsically imperfect: even without missing any physics and removing current waveform systematics, there will always be intrinsic limitations due to truncation errors in perturbative schemes, calibration inaccuracy with NR waveforms, phenomenological modelling of the merger, unavoidable numerical errors in NR simulations. Thus, we will have to always face some degree of waveform systematics, noise artifact, or astrophysical uncertainty, whose potential impact will grow for high SNR events. The point here is to control such systematics as much as possible, to a level that make them negligible with respect to a putative GR deviation.

We summarize the discussion in Table 1. We note that this is intended as a rough guide as exact predictions for the size of relative effects can depend on a number of factors, and one expects improvements in the coming years (e.g., one expects waveform systematics to improve in the coming years, however, we do not consider this here). Below we give our reasoning for why we think these causes will be important (or not) for a given detector sensitivity.

Cause	**O4**	A+	A$^\#$	XG
Non-Stationary Noise	✓	✓	✓	✓
Non-Gaussian Noise/Glitches	✓	✓	✓	✓
Overlapping Signals	✗	✗	✗	✓
Data Gaps	✗	✗	✗	✓
Detector Calibration	✗	✗	✗	✓
Eccentricity	✓	✓	✓	✓
Tidal Effects	✗	✓	✓	✓
Kick-induced Effects	✗	✗	✗	✓
Ringdown Modes	✓	✓	✓	✓
Precession and Higher-order Modes	✓	✓	✓	✓
Memory	✗	✗	✓	✓
Sub-optimal Waveform Calibration	✗	✗	✓	✓
Lensing	✗	✗	✗	✓
Environmental Effects	✗	✗	✗	✓
Source Misclassification	✓	✓	✓	✓
Astrophysical Population Assumptions	✓	✓	✓	✓

Table 1: Summary of the causes discussed in this paper that can potentially mimic a GR deviation while performing tests of GR. The tick means the effect should be accounted for in the waveform models and/or analysis methods when analyzing data of a GW detector of a given sensitivity. The cross means the effect is sub-dominant to show up as a false GR violation with that detector sensitivity.

5.1 Noise Systematics

Non-stationarities, non-Gaussianities, overlapping signals Non-stationary and non-Gaussian noise artifacts are an ever-present analysis burden in the current generation of observatories as discussed in Sections 2.1 and 2.2. While the extent to which these artifacts will alter with upgrades to current observatories or persist in future-generation observatories remains uncertain, it is difficult to imagine that they will subside to any degree. It therefore behooves analysts to understand and mitigate these noise sources as post-processing steps before making any claim of a GR violation. On the other hand, the effect of contamination from overlapping signals, whether they be super- or sub-threshold to detection, will only increase and get worse as the sensitivity of instruments gets better.

Data Gaps For current-generation detectors, data gaps are not expected to pose a problem for tests of GR due to the typically short duration of signals in the band and the low likelihood of data gaps occurring precisely during those times. For XG observatories, however, data gaps could become more problematic, as the signal duration increases to many hours to days, and the likelihood of gaps increases.

Detector calibration For the current generation of observatories, uncertainties due to detector calibration do not introduce biases in parameter estimation when assuming general-relativistic waveforms, and therefore are not expected to introduce problems in tests of GR (e.g., [102] and see Section 2.5). For XG observatories, assuming an $\approx 1\%$ relative error on the amplitude, and $\approx 1^\circ$ error in phase, detector calibration error leads to mismatch errors of approximately 10^{-5}, which may be problematic for tests of GR [500]. Of course, this is only a dominant source of uncertainty if other sources (e.g., waveform systematics) can be mitigated below this level.

5.2 Waveform Systematics

Eccentricity Employing non-precessing, eccentric waveforms, some papers have claimed the evidence for eccentricity in observed GW signals [128, 129, 138–140]. Although this is contentious (see discussion in Section 3.1.1), it points to the fact that effects of eccentricity are already relevant in current observations, and therefore already pose a difficulty when performing tests of GR. This will continue to be a problem, and may be further exacerbated, as observatories become more sensitive.

Tidal Effects Tidal signatures may be present in several observed neutron star binary mergers (e.g., [62, 501]), although a confident detection of tidal signature is yet to occur. While misspecification of tidal effects is unlikely to appear as a GR violation in current detectors, a clean tidal signature may be present in A+ observatories for dynamical tidal effects [502], and XG detectors for linear tides (e.g., [503, 504]).

Kick-induced Effects The kick-induced effects are too small to be detected with the current GW detectors but could potentially be observed in XG era [234, 235]. The XG detectors are expected to observe $\sim 4-5$ events per year for which these effects will be constrained to better than $\sim 10\%$ [234].

Ringdown Tests of GR and the no-hair theorem are already performed using the ringdown of loud GW signals (e.g., [1]) where the challenges that arise with specifying the ringdown start time and avoiding overfitting to nonlinearities are already present. These challenges will only intensify as the ringdown signals become louder in future observatories (e.g., [286]).

Precession and Higher-order Modes Several events in the existing GWTC have strong evidence of higher-order modes due, e.g., to large mass ratios such as GW190412 [505] and GW190814 [506]. There are several events that have evidence of spin precession, such as GW190521 [507] and GW200129 ([70], although see [71,508]). It is therefore important to account for spin precession and higher-order modes in current analyses, and the inclusion of higher modes will become even more important as the sensitivity of observatories continues to improve.

Memory Displacement memory is too small to be detected in individual events with the sensitivities of current detectors [334, 509–511]. A memory signal is expected to influence parameter estimation results in loud events with SNR greater than 60, expected during the $A^{\#}$ era [331], implying at this stage memory needs to be properly accounted for in waveforms models. Memory will have a significant influence in XG observatories; for example, Cosmic Explorer is predicted to have 3 to 4 events per year where memory is detectable for an individual event [334], amplifying the need to properly account for memory effects.

Waveform Calibration If we consider NR simulations to be the ground truth, then current waveform calibration errors refer to systematic biases introduced because the waveform approximants do not exactly match the NR simulations. But even NR waveforms carry uncertainties associated with, e.g., resolution effects and finite radius extraction. Such waveform calibration errors on the order of a few percent in amplitude, and a couple of degrees in phase, are subdominant to stochastic noise processes for binary neutron star observations at approximately 100 Mpc in A+ observatories [164]. Waveform uncertaintes are currently smaller than this, implying they are not a potential source of bias for tests of GR. This is not necessarily true in the $A^{\#}$ and XG era when even NR waveforms will not be sufficiently accurate for unbiased parameter estimation recovery [337,500]. This latter point motivates the continual need for more accurate NR simulations and waveform extraction methods, as well as waveform approximations.

5.3 Astrophysical Aspects

Lensing In current and future detectors like advanced LIGO and A+, the estimated rate of strong lensing events for binary neutron stars is approximately 0.1%, while for binary black holes it is expected to be around 0.2%. These figures are consistent across various studies [512–514]. Following this, advanced LIGO is anticipated to detect approximately 0.1 lensing events per year, whereas A+ is projected to observe 1 event annually. However, with XG detectors, $\mathcal{O}(100)$ events could be detected per year. It is important to note that these rates serve as a lower bound for millilensing and microlensing, since they could occur together with strong lensing in events. Therefore, lensing effects will not be a significant issue only until XG era.

Environmental Effects Astrophysical environments in which one may anticipate binary systems merging (and which may leave an imprint on the GW signal) include thick ($\bar{\rho} \sim 10^{-8}$ g/cm^3) and thin ($\bar{\rho} \sim 0.1$ g/cm^3) accretion disks around active galactic nuclei [32], cold dark matter spikes ($\bar{\rho} \sim 10^{-6}$ g/cm^3) [404], superradiant-boson clouds ($\bar{\rho} \sim 0.1$ g/cm^3) [416] and the dynamical fragmentation of massive stars ($\bar{\rho} \sim 10^7$ g/cm^3) [393]. Santoro *et al.* [442] found no support for environmental effects in GWTC-1, and found the environmental density would need to be $\sim 20\,\mathrm{g/cm^3}$ to be observable. This likely does not correspond to any of the astrophysical environments mentioned previously. For advanced LIGO design sensitivity, they find that dynamical friction effects are detectable at $\bar{\rho} \gtrsim 10\,\mathrm{g/cm^3}$ for a GW170817-like event, while the effect of collisionless accretion is only visible for densities 10-100 times greater. As there are no proposed environments with such densities, it is unlikely for environmental effects to be visible in advanced LIGO data. They find that XG observatories will be sensitive to environmental densities of $\sim 10^{-3}\,\mathrm{g/cm^3}$, which includes both thin accretion disks and superradiant clouds. It is therefore likely that environmental signatures will only become relevant for GR tests in XG and beyond.

Source Misclassification The problem of source misclassification is ever-present in tests of GR and must be considered when mitigating against false GR violations. For example, while current analyses find no evidence of GW echoes that would provide evidence of black-hole mimickers (see Section 4.3.1), these non-detections only place limits on, e.g., the reflective properties of the ultra-compact objects. As the sensitivity of the GW network improves, we will continue to probe the parameter space of potential black-hole mimickers.

Astrophysical Population Assumptions The problem of fortifying hierarchical tests of GR against population assumptions and modelling systematics will be ever-present. Statistical assumptions on how to combine the information from individual events require care, as they reflect implicit assumptions on the beyond-GR theory that is being tested [241,495]. Incorrect prior assumptions on the astrophysical population can cause biases if the deviation parameters are correlated with individual source parameters. These biases can be mitigated by jointly inferring the astrophysical population when performing hierarchical tests of GR, or in the high-SNR limit of XG detectors if the degeneracies between source parameters and deviation parameters are not perfect [492]. Effects due to the finite size of the catalog [499] or selection effects against large deviations [493] can also lead to biases in population constraints if not properly accounted for. Finally, waveform systematics (both due to missing physics and sub-optimal calibration) can accumulate in a population analysis and lead to infer false GR violations even if the biases are under control at the single-event level [136,490]. This effect will be even more prominent when restricting the test to high-SNR events that can be routinely observed with XG detectors [90].

6 Summary

Since the first detection in 2015, GW observations are now routinely used to test GR in highly dynamical and non-linear gravity regimes. Several tests of GR exist at the moment and the majority of them rely on comparing the GW data with well-motivated, state-of-the-art waveform models. The GW observations from the LIGO-Virgo-KAGRA collaboration have so far not found any deviation from GR, but this may not be the case forever, especially with the increased sensitivity of GW detectors. In the future, all these well-motivated, state-of-the-art waveform models may fall short of explaining all the features in the high-quality data due to the complexity of the physics of GW sources and the detector noise modeling.

In this paper, we listed the possible causes that can lead to an apparent GR deviation using observations from ground-based GW detectors given the current waveform models and data analysis techniques that are available to the community. We grouped these causes into three broad categories: noise systematics, waveform systematics, and astrophysical aspects. Noise systematics include noise being non-stationary and/or non-Gaussian with or without time-overlapping signals present in the data, gaps in data, and errors in instrument calibration. Waveform systematics include cases of missing physics such as eccentricity, tides, kicks, overtones, mirror modes, and non-linear ringdown modes, and sub-optimal modeling and calibration (with NR waveforms) of quasi-circular waveforms. Astrophysical aspects include gravitational lensing, non-vacuum environments, mistaken source classes, and assumptions of astrophysical population.

Our list is admittedly not complete and we might have missed some other important causes of false GR deviation. However, we hope that this paper will serve as a starting point for the community to study, understand, and document the effects of these causes on tests of GR. In a follow-up paper, we will discuss what actions could be taken when a significant GR deviation is detected and propose a possible formulation of a GR violation detection checklist. We hope that these efforts will prepare us for the time when there will be an actual statistically significant GR deviation found in the GW data.

Acknowledgements

Funding information A. Gupta, N.K. Johnson-McDaniel, and P. Narayan are supported by NSF Grants No. AST-2205920 and PHY-2308887. K.G. Arun acknowledges support from the Department of Science and Technology and Science and Engineering Research Board (SERB) of India via the Swarnajayanti Fellowship Grant DST/SJF/PSA-01/2017-18 and the Core Research Grant CRG/2021/004565. K.G. Arun and B.S. Sathyaprakash acknowledge the support of the Indo-US Science and Technology Forum through the Indo-US Centre for Gravitational-Physics and Astronomy, grant IUSSTF/JC-142/2019. E. Barausse acknowledges support from the European Union's H2020 ERC Consolidator Grant "GRavity from Astrophysical to Microscopic Scales" (Grant No. GRAMS-815673), the PRIN 2022 grant "GUVIRP - Gravity tests in the UltraViolet and InfraRed with Pulsar timing", and the EU Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie Grant Agreement No. 101007855. L. Bernard acknowledges financial support from the ANR PRoGRAM project, grant ANR-21-CE31-0003-001 and the EU Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie Grant Agreement no. 101007855. E. Berti and L. Reali are supported by NSF Grants No. AST-2006538, PHY-2207502, PHY-090003 and PHY-20043, by NASA Grants No. 20-LPS20-0011 and 21-ATP21-0010, by the John Templeton Foundation Grant 62840, by the Simons Foundation, and by the Italian Ministry of Foreign Affairs and International Cooperation grant No. PGR01167. V. Cardoso is a Villum Investigator and a DNRF Chair, supported by the VILLUM Foundation (grant no. VIL37766) and the DNRF Chair program (grant no. DNRF162) by the Danish National Research Foundation. V. Cardoso acknowledges financial support provided under the European Union's H2020 ERC Advanced Grant "Black holes: gravitational engines of discovery" grant agreement no. Gravitas–101052587. Views and opinions expressed are however those of the author only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. V. Cardoso has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101007855 and No. 101131233. S.Y. Cheung, T. Clarke, N. Guttman, P.D. Lasky, L. Passenger, and H. Tong are supported by Australian Research Council (ARC) Centre of Excellence for Gravitational-Wave Discovery CE170100004 and CE230100016, Discovery Projects DP220101610 and DP230103088, and LIEF LE210100002. S. Datta acknowledges support from UVA Arts and Sciences Rising Scholars Fellowship. A. Dhani, I. Gupta, R. Kashyap and B.S. Sathyaprakash were supported in part by NSF Grants No. PHY-2207638, AST-2307147, PHY-2308886, and PHYS-2309064. B. Sathyaprakash also thanks the Aspen Center for Physics (ACP) for hospitality during the summer of 2022. J.M. Ezquiaga is supported by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 847523 INTERACTIONS, and by VILLUM FONDEN (grant no. 53101 and 37766). E. Maggio acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG) - project number: 386119226. E. Maggio is supported by the European Union's Horizon Europe research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101107586. A. Maselli acknowledges financial support from MUR PRIN Grants No. 2022-Z9X4XS and 2020KB33TP. S. Tiwari is supported by the Swiss National Science Foundation Ambizione grant no. PZ00P2-202204. K. Yagi is supported by NSF Grant PHY-2207349, PHY-2309066, PHYS-2339969, and the Owens Family Foundation. N. Yunes is supported by the Simons Foundation through Award No. 896696, the NSF Grant No. PHY-2207650, and the NASA Award No. 80NSSC22K0806.

References

[1] B. P. Abbott et al., Tests of general relativity with GW150914, Phys. Rev. Lett. **116**(22), 221101 (2016), doi:10.1103/PhysRevLett.116.221101, [Erratum: Phys.Rev.Lett. 121, 129902 (2018)], 1602.03841.

[2] B. P. Abbott et al., Tests of General Relativity with GW170817, Phys. Rev. Lett. **123**(1), 011102 (2019), doi:10.1103/PhysRevLett.123.011102, 1811.00364.

[3] B. P. Abbott et al., Tests of General Relativity with the Binary Black Hole Signals from the LIGO-Virgo Cata Phys. Rev. D **100**(10), 104036 (2019), doi:10.1103/PhysRevD.100.104036, 1903.04467.

[4] R. Abbott et al., Tests of general relativity with binary black holes from the second LIGO-Virgo gravitation Phys. Rev. D **103**(12), 122002 (2021), doi:10.1103/PhysRevD.103.122002, 2010.14529.

[5] R. Abbott et al., Tests of General Relativity with GWTC-3 (2021), 2112.06861.

[6] N. Yunes, K. Yagi and F. Pretorius, Theoretical Physics Implications of the Binary Black-Hole Mergers GW1 Phys. Rev. D **94**(8), 084002 (2016), doi:10.1103/PhysRevD.94.084002, 1603.08955.

[7] R. Nair, S. Perkins, H. O. Silva and N. Yunes, Fundamental Physics Implications for Higher-Curvature Theories from Binary Black Hole Signals in the L Phys. Rev. Lett. **123**(19), 191101 (2019), doi:10.1103/PhysRevLett.123.191101, 1905.00870.

[8] H. O. Silva, A. M. Holgado, A. Cárdenas-Avendaño and N. Yunes, Astrophysical and theoretical physics implications from multimessenger neutron star observations, Phys. Rev. Lett. **126**(18), 181101 (2021), doi:10.1103/PhysRevLett.126.181101, 2004.01253.

[9] S. E. Perkins, R. Nair, H. O. Silva and N. Yunes, Improved gravitational-wave constraints on higher-order curvature theories of gravity, Phys. Rev. D **104**(2), 024060 (2021), doi:10.1103/PhysRevD.104.024060, 2104.11189.

[10] K. Schumacher, S. E. Perkins, A. Shaw, K. Yagi and N. Yunes, Gravitational wave constraints on Einstein-æther theory with LIGO/Virgo data, Phys. Rev. D **108**(10), 104053 (2023), doi:10.1103/PhysRevD.108.104053, 2304.06801.

[11] T. Callister, L. Jenks, D. Holz and N. Yunes, A New Probe of Gravitational Parity Violation Through (Non-) (2023), 2312.12532.

[12] M. Lagos, L. Jenks, M. Isi, K. Hotokezaka, B. D. Metzger, E. Burns, W. M. Farr, S. Perkins, K. W. K. Wong and N. Yunes, Birefringence tests of gravity with multimessenger binaries, Phys. Rev. D **109**(12), 124003 (2024), doi:10.1103/PhysRevD.109.124003, 2402.05316.

[13] J. Aasi et al., Advanced LIGO, Class. Quant. Grav. **32**, 074001 (2015), doi:10.1088/0264-9381/32/7/074001, 1411.4547.

[14] T. Akutsu et al., Overview of KAGRA: Calibration, detector characterization, physical environmental moni PTEP **2021**(5), 05A102 (2021), doi:10.1093/ptep/ptab018, 2009.09305.

[15] The KAGRA Collaboration, the LIGO Scientific Collaboration, and the Virgo Collaboration, LIGO, Virgo and KAGRA Observing Run Plans, https://observing.docs.ligo.org/plan/ (2022).

[16] B. P. Abbott et al., Prospects for observing and localizing gravitational-wave transients with Advanced LIG Living Rev. Rel. **21**(1), 3 (2018), doi:10.1007/s41114-020-00026-9, 1304.0670.

[17] Iyer et al., LIGO-India, Proposal of the Consortium for Indian Initiative in Gravitational-wave Observation https://dcc.ligo.org/LIGO-T2200287/public (2011).

[18] M. Saleem et al., The science case for LIGO-India, Class. Quant. Grav. **39**(2), 025004 (2022), doi:10.1088/1361-6382/ac3b99, 2105.01716.

[19] Fritschel et al., Report of the LSC Post-O5 Study Group, Tech. Rep. T2200287, https://dcc.ligo.org/LIGO-T2200287/public (2022).

[20] M. Evans et al., A Horizon Study for Cosmic Explorer: Science, Observatories, and Community (2021), 2109.09882.

[21] S. Hild et al., Sensitivity Studies for Third-Generation Gravitational Wave Observatories, Class. Quant. Grav. **28**, 094013 (2011), doi:10.1088/0264-9381/28/9/094013, 1012.0908.

[22] P. Amaro-Seoane et al., Laser Interferometer Space Antenna (2017), 1702.00786.

[23] J. Luo et al., TianQin: a space-borne gravitational wave detector, Class. Quant. Grav. **33**(3), 035010 (2016), doi:10.1088/0264-9381/33/3/035010, 1512.02076.

[24] J. Mei et al., The TianQin project: current progress on science and technology, PTEP **2021**(5), 05A107 (2021), doi:10.1093/ptep/ptaa114, 2008.10332.

[25] Y.-L. Wu et al., China's first step towards probing the expanding universe and the nature of gravity using a Commun. Phys. **4**(1), 34 (2021), doi:10.1038/s42005-021-00529-z.

[26] S. Kawamura et al., The Japanese space gravitational wave antenna DECIGO, Class. Quant. Grav. **23**, S125 (2006), doi:10.1088/0264-9381/23/8/S17.

[27] S. Kawamura et al., Current status of space gravitational wave antenna DECIGO and B-DECIGO, PTEP **2021**(5), 05A105 (2021), doi:10.1093/ptep/ptab019, 2006.13545.

[28] P. Ajith et al., The Lunar Gravitational-wave Antenna: Mission Studies and Science Case (2024), 2404.09181.

[29] G. T. Horowitz, M. Kolanowski, G. N. Remmen and J. E. Santos, Extremal Kerr Black Holes as Amplifiers of New Physics, Phys. Rev. Lett. **131**(9), 091402 (2023), doi:10.1103/PhysRevLett.131.091402, 2303.07358.

[30] R. Jackiw and S. Y. Pi, Chern-Simons modification of general relativity, Phys. Rev. D **68**, 104012 (2003), doi:10.1103/PhysRevD.68.104012, gr-qc/0308071.

[31] S. Alexander and N. Yunes, Chern-Simons Modified General Relativity, Phys. Rept. **480**, 1 (2009), doi:10.1016/j.physrep.2009.07.002, 0907.2562.

[32] E. Barausse, V. Cardoso and P. Pani, Can environmental effects spoil precision gravitational-wave astrophys Phys. Rev. D **89**(10), 104059 (2014), doi:10.1103/PhysRevD.89.104059, 1404.7149.

[33] V. Cardoso, E. Franzin and P. Pani, Is the gravitational-wave ringdown a probe of the event horizon?, Phys. Rev. Lett. **116**(17), 171101 (2016), doi:10.1103/PhysRevLett.116.171101, [Erratum: Phys.Rev.Lett. 117, 089902 (2016)], 1602.07309.

[34] C. de Rham and S. Melville, Gravitational Rainbows: LIGO and Dark Energy at its Cutoff, Phys. Rev. Lett. **121**(22), 221101 (2018), doi:10.1103/PhysRevLett.121.221101, 1806.09417.

[35] L. ter Haar, M. Bezares, M. Crisostomi, E. Barausse and C. Palenzuela, Dynamics of Screening in Modified Gravity, Phys. Rev. Lett. **126**, 091102 (2021), doi:10.1103/PhysRevLett.126.091102, 2009.03354.

[36] M. Bezares, R. Aguilera-Miret, L. ter Haar, M. Crisostomi, C. Palenzuela and E. Barausse, No Evidence of Kinetic Screening in Simulations of Merging Binary Neutron Stars beyond General Relativ Phys. Rev. Lett. **128**(9), 091103 (2022), doi:10.1103/PhysRevLett.128.091103, 2107.05648.

[37] M. Bezares, L. ter Haar, M. Crisostomi, E. Barausse and C. Palenzuela, Kinetic screening in nonlinear stellar oscillations and gravitational collapse, Phys. Rev. D **104**(4), 044022 (2021), doi:10.1103/PhysRevD.104.044022, 2105.13992.

[38] R. Abbott et al., GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo during the Second Phys. Rev. X **13**(4), 041039 (2023), doi:10.1103/PhysRevX.13.041039, 2111.03606.

[39] E. Maggio, H. O. Silva, A. Buonanno and A. Ghosh, Tests of general relativity in the nonlinear regime: A parametrized plunge-merger-ringdown gravitational Phys. Rev. D **108**(2), 024043 (2023), doi:10.1103/PhysRevD.108.024043, 2212.09655.

[40] K. G. Arun et al., New horizons for fundamental physics with LISA, Living Rev. Rel. **25**(1), 4 (2022), doi:10.1007/s41114-022-00036-9, 2205.01597.

[41] B. P. Abbott et al., A guide to LIGO–Virgo detector noise and extraction of transient gravitational-wave sig Class. Quant. Grav. **37**(5), 055002 (2020), doi:10.1088/1361-6382/ab685e, 1908.11170.

[42] B. P. Abbott et al., Characterization of transient noise in Advanced LIGO relevant to gravitational wave sig Class. Quant. Grav. **33**(13), 134001 (2016), doi:10.1088/0264-9381/33/13/134001, 1602.03844.

[43] D. Davis et al., LIGO detector characterization in the second and third observing runs, Class. Quant. Grav. **38**(13), 135014 (2021), doi:10.1088/1361-6382/abfd85, 2101.11673.

[44] E. Capote, L. Dartez and D. Davis, Technical noise, data quality, and calibration requirements for next-gene Class. Quant. Grav. **41**(18), 185001 (2024), doi:10.1088/1361-6382/ad694d, 2404.04761.

[45] A. Biswas, J. McIver and A. Mahabal, New methods to assess and improve LIGO detector duty cycle, Class. Quant. Grav. **37**(17), 175008 (2020), doi:10.1088/1361-6382/ab8650, 1910.12143.

[46] K. Janssens et al., Correlated 1–1000 Hz magnetic field fluctuations from lightning over Earth-scale distar Phys. Rev. D **107**(2), 022004 (2023), doi:10.1103/PhysRevD.107.022004, 2209.00284.

[47] O. Edy, A. Lundgren and L. K. Nuttall, Issues of mismodeling gravitational-wave data for parameter estima Phys. Rev. D **103**(12), 124061 (2021), doi:10.1103/PhysRevD.103.124061, 2101.07743.

[48] S. Kumar, A. H. Nitz and X. J. Forteza, Parameter estimation with non stationary noise in gravitational wav (2022), 2202.12762.

[49] S. A. Usman et al., The PyCBC search for gravitational waves from compact binary coalescence, Class. Quant. Grav. **33**(21), 215004 (2016), doi:10.1088/0264-9381/33/21/215004, 1508.02357.

[50] C. Messick, K. Blackburn, P. Brady, P. Brockill, K. Cannon, R. Cariou, S. Caudill, S. J. Chamberlin, J. D. E. Creighton, R. Everett, C. Hanna, D. Keppel et al., Analysis framework for the prompt discovery of compact binary mergers in gravitational-wave data, Phys. Rev. D **95**, 042001 (2017), doi:10.1103/PhysRevD.95.042001.

[51] T. Venumadhav, B. Zackay, J. Roulet, L. Dai and M. Zaldarriaga, New search pipeline for compact binary mergers: Results for binary black holes in the first observing run Phys. Rev. D **100**(2), 023011 (2019), doi:10.1103/PhysRevD.100.023011, 1902.10341.

[52] S. Mozzon, L. K. Nuttall, A. Lundgren, T. Dent, S. Kumar and A. H. Nitz, Dynamic normalization for compact binary coalescence searches in non-stationary noise, Classical and Quantum Gravity **37**(21), 215014 (2020), doi:10.1088/1361-6382/abac6c, 2002.09407.

[53] B. Zackay, T. Venumadhav, J. Roulet, L. Dai and M. Zaldarriaga, Detecting gravitational waves in data with non-stationary and non-Gaussian noise, Phys. Rev. D **104**(6), 063034 (2021), doi:10.1103/PhysRevD.104.063034, 1908.05644.

[54] K. Chatziioannou, C.-J. Haster, T. B. Littenberg, W. M. Farr, S. Ghonge, M. Millhouse, J. A. Clark and N. Cornish, Noise spectral estimation methods and their impact on gravitational wave measurement of compact bina Phys. Rev. D **100**(10), 104004 (2019), doi:10.1103/PhysRevD.100.104004, 1907.06540.

[55] O. Edy, A. Lundgren and L. K. Nuttall, Issues of mismodeling gravitational-wave data for parameter estima Phys. Rev. D **103**(12), 124061 (2021), doi:10.1103/PhysRevD.103.124061, 2101.07743.

[56] B. P. Abbott et al., Effects of data quality vetoes on a search for compact binary coalescences in Advanced I Class. Quant. Grav. **35**(6), 065010 (2018), doi:10.1088/1361-6382/aaaafa, 1710.02185.

[57] G. Ashton, S. Thiele, Y. Lecoeuche, J. McIver and L. K. Nuttall, Parameterised population models of transient non-Gaussian noise in the LIGO gravitational-wave detecto Class. Quant. Grav. **39**(17), 175004 (2022), doi:10.1088/1361-6382/ac8094, 2110.02689.

[58] M. Cabero et al., Blip glitches in Advanced LIGO data, Class. Quant. Grav. **36**(15), 15 (2019), doi:10.1088/1361-6382/ab2e14, 1901.05093.

[59] S. Soni, C. Austin, A. Effler, R. M. S. Schofield, G. González, V. V. Frolov, J. C. Driggers, A. Pele, A. L. Urban, G. Valdes, R. Abbott, C. Adams et al., Reducing scattered light in LIGO's third observing run, Classical and Quantum Gravity **38**(2), 025016 (2021), doi:10.1088/1361-6382/abc906, 2007.14876.

[60] M. Zevin et al., Gravity Spy: Integrating Advanced LIGO Detector Characterization, Machine Learning, an Class. Quant. Grav. **34**(6), 064003 (2017), doi:10.1088/1361-6382/aa5cea, 1611.04596.

[61] M. Razzano, F. Di Renzo, F. Fidecaro, G. Hemming and S. Katsanevas, GWitchHunters: Machine learning and citizen science to improve the performance of gravitational wave Nucl. Instrum. Meth. A **1048**, 167959 (2023), doi:10.1016/j.nima.2022.167959, 2301.05112.

[62] B. P. Abbott et al., GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral, Phys. Rev. Lett. **119**(16), 161101 (2017), doi:10.1103/PhysRevLett.119.161101, 1710.05832.

[63] R. Macas, J. Pooley, L. K. Nuttall, D. Davis, M. J. Dyer, Y. Lecoeuche, J. D. Lyman, J. McIver and K. Rink, Impact of noise transients on low latency gravitational-wave event localization, Phys. Rev. D **105**(10), 103021 (2022), doi:10.1103/PhysRevD.105.103021, 2202.00344.

[64] D. Davis, T. J. Massinger, A. P. Lundgren, J. C. Driggers, A. L. Urban and L. K. Nuttall, Improving the Sensitivity of Advanced LIGO Using Noise Subtraction, Class. Quant. Grav. **36**(5), 055011 (2019), doi:10.1088/1361-6382/ab01c5, 1809.05348.

[65] J. C. Driggers et al., Improving astrophysical parameter estimation via offline noise subtraction for Advanc Phys. Rev. D **99**(4), 042001 (2019), doi:10.1103/PhysRevD.99.042001, 1806.00532.

[66] G. Vajente, Y. Huang, M. Isi, J. C. Driggers, J. S. Kissel, M. J. Szczepanczyk and S. Vitale, Machine-learning nonstationary noise out of gravitational-wave detectors, Phys. Rev. D **101**(4), 042003 (2020), doi:10.1103/PhysRevD.101.042003, 1911.09083.

[67] B. P. Abbott et al., GW150914: First results from the search for binary black hole coalescence with Advanc Phys. Rev. D **93**(12), 122003 (2016), doi:10.1103/PhysRevD.93.122003, 1602.03839.

[68] N. J. Cornish and T. B. Littenberg, BayesWave: Bayesian Inference for Gravitational Wave Bursts and Instru Class. Quant. Grav. **32**(13), 135012 (2015), doi:10.1088/0264-9381/32/13/135012, 1410.3835.

[69] D. Davis, T. B. Littenberg, I. M. Romero-Shaw, M. Millhouse, J. McIver, F. Di Renzo and G. Ashton, Subtracting glitches from gravitational-wave detector data during the third LIGO-Virgo observ Class. Quant. Grav. **39**(24), 245013 (2022), doi:10.1088/1361-6382/aca238, 2207.03429.

[70] M. Hannam et al., General-relativistic precession in a black-hole binary, Nature **610**(7933), 652 (2022), doi:10.1038/s41586-022-05212-z, 2112.11300.

[71] E. Payne, S. Hourihane, J. Golomb, R. Udall, R. Udall, D. Davis and K. Chatziioannou, Curious case of GW200129: Interplay between spin-precession inference and data-quality issues, Phys. Rev. D **106**(10), 104017 (2022), doi:10.1103/PhysRevD.106.104017, 2206.11932.

[72] A. Bohé et al., Improved effective-one-body model of spinning, nonprecessing binary black holes for the e
Phys. Rev. D **95**(4), 044028 (2017), doi:10.1103/PhysRevD.95.044028, 1611.03703.

[73] R. Cotesta, A. Buonanno, A. Bohé, A. Taracchini, I. Hinder and S. Ossokine, Enriching the Symphony of Gravitational Waves from Binary Black Holes by Tuning Higher Harmonics, Phys. Rev. D **98**(8), 084028 (2018), doi:10.1103/PhysRevD.98.084028, 1803.10701.

[74] D. P. Mihaylov, S. Ossokine, A. Buonanno and A. Ghosh, Fast post-adiabatic waveforms in the time domain: Applications to compact binary coalescences in LIGO
Phys. Rev. D **104**(12), 124087 (2021), doi:10.1103/PhysRevD.104.124087, 2105.06983.

[75] G. Ashton, Gaussian processes for glitch-robust gravitational-wave astronomy, Mon. Not. Roy. Astron. Soc. **520**(2), 2983 (2023), doi:10.1093/mnras/stad341, 2209.15547.

[76] J. Heinzel, C. Talbot, G. Ashton and S. Vitale, Inferring the astrophysical population of gravitational wave sources in the presence of noise transients, Mon. Not. Roy. Astron. Soc. **523**(4), 5972 (2023), doi:10.1093/mnras/stad1823, 2304.02665.

[77] J. Y. L. Kwok, R. K. L. Lo, A. J. Weinstein and T. G. F. Li, Investigation of the effects of non-Gaussian noise transients and their mitigation in parameterized gravit
Phys. Rev. D **105**(2), 024066 (2022), doi:10.1103/PhysRevD.105.024066, 2109.07642.

[78] A. Samajdar, J. Janquart, C. Van Den Broeck and T. Dietrich, Biases in parameter estimation from overlapping gravitational-wave signals in the third-generation detec
Phys. Rev. D **104**(4), 044003 (2021), doi:10.1103/PhysRevD.104.044003, 2102.07544.

[79] T. Regimbau et al., A Mock Data Challenge for the Einstein Gravitational-Wave Telescope, Phys. Rev. D **86**, 122001 (2012), doi:10.1103/PhysRevD.86.122001, 1201.3563.

[80] D. Meacher, K. Cannon, C. Hanna, T. Regimbau and B. S. Sathyaprakash, Second Einstein Telescope Mock Data and Science Challenge: Low Frequency Binary Neutron Star Data A
Phys. Rev. D **93**(2), 024018 (2016), doi:10.1103/PhysRevD.93.024018, 1511.01592.

[81] E. Pizzati, S. Sachdev, A. Gupta and B. Sathyaprakash, Toward inference of overlapping gravitational-wave signals, Phys. Rev. D **105**(10), 104016 (2022), doi:10.1103/PhysRevD.105.104016, 2102.07692.

[82] P. Relton and V. Raymond, Parameter estimation bias from overlapping binary black hole events in second
Phys. Rev. D **104**(8), 084039 (2021), doi:10.1103/PhysRevD.104.084039, 2103.16225.

[83] Y. Himemoto, A. Nishizawa and A. Taruya, Impacts of overlapping gravitational-wave signals on the param
Phys. Rev. D **104**(4), 044010 (2021), doi:10.1103/PhysRevD.104.044010, 2103.14816.

[84] P. Relton, A. Virtuoso, S. Bini, V. Raymond, I. Harry, M. Drago, C. Lazzaro, A. Miani and S. Tiwari, Addressing the challenges of detecting time-overlapping compact binary coalescences, Phys. Rev. D **106**(10), 104045 (2022), doi:10.1103/PhysRevD.106.104045, 2208.00261.

[85] S. Wu and A. H. Nitz, Mock data study for next-generation ground-based detectors: The performance loss Phys. Rev. D **107**(6), 063022 (2023), doi:10.1103/PhysRevD.107.063022, 2209.03135.

[86] A. Antonelli, O. Burke and J. R. Gair, Noisy neighbours: inference biases from overlapping gravitational-wa Mon. Not. Roy. Astron. Soc. **507**(4), 5069 (2021), doi:10.1093/mnras/stab2358, 2104.01897.

[87] J. Janquart, T. Baka, A. Samajdar, T. Dietrich and C. Van Den Broeck, Analyses of overlapping gravitational wave signals using hierarchical subtraction and joint parameter est Mon. Not. Roy. Astron. Soc. **523**(2), 1699 (2023), doi:10.1093/mnras/stad1542, 2211.01304.

[88] J. Langendorff, A. Kolmus, J. Janquart and C. Van Den Broeck, Normalizing Flows as an Avenue to Studying Overlapping Gravitational Wave Signals, Phys. Rev. Lett. **130**(17), 171402 (2023), doi:10.1103/PhysRevLett.130.171402, 2211.15097.

[89] A. D. Johnson, K. Chatziioannou and W. M. Farr, Source confusion from neutron star binaries in ground-based gravitational wave detectors is minimal, Phys. Rev. D **109**(8), 084015 (2024), doi:10.1103/PhysRevD.109.084015, 2402.06836.

[90] Q. Hu and J. Veitch, Accumulating Errors in Tests of General Relativity with Gravitational Waves: Overlap Astrophys. J. **945**(2), 103 (2023), doi:10.3847/1538-4357/acbc18, 2210.04769.

[91] Y. Dang, Z. Wang, D. Liang and L. Shao, Impact of Overlapping Signals on Parameterized Post-Newtonian Astrophys. J. **964**(2), 194 (2024), doi:10.3847/1538-4357/ad2e00, 2311.16184.

[92] R. X. Adhikari et al., A cryogenic silicon interferometer for gravitational-wave detection, Class. Quant. Grav. **37**(16), 165003 (2020), doi:10.1088/1361-6382/ab9143, 2001.11173.

[93] R. Abbott et al., GWTC-2.1: Deep extended catalog of compact binary coalescences observed by LIGO an Phys. Rev. D **109**(2), 022001 (2024), doi:10.1103/PhysRevD.109.022001, 2108.01045.

[94] L. Reali, A. Antonelli, R. Cotesta, S. Borhanian, M. Çalışkan, E. Berti and B. S. Sathyaprakash, The impact of confusion noise on golden binary neutron-star events in next-generation te (2022), 2209.13452.

[95] L. Reali, A. Maselli and E. Berti, The impact of compact binary confusion noise on tests of fundamental phy (2023), 2307.01264.

[96] J. Carre and E. K. Porter, The Effect of Data Gaps on LISA Galactic Binary Parameter Estimation (2010), 1010.1641.

[97] K. Dey, N. Karnesis, A. Toubiana, E. Barausse, N. Korsakova, Q. Baghi and S. Basak, Effect of data gaps on the detectability and parameter estimation of massive black hole binaries with LIS Phys. Rev. D **104**(4), 044035 (2021), doi:10.1103/PhysRevD.104.044035, 2104.12646.

[98] Q. Baghi, I. Thorpe, J. Slutsky, J. Baker, T. Dal Canton, N. Korsakova and N. Karnesis, Gravitational-wave parameter estimation with gaps in LISA: a Bayesian data augmentation method, Phys. Rev. D **100**(2), 022003 (2019), doi:10.1103/PhysRevD.100.022003, 1907.04747.

[99] B. P. Abbott et al., Calibration of the Advanced LIGO detectors for the discovery of the binary black-hole m Phys. Rev. D **95**(6), 062003 (2017), doi:10.1103/PhysRevD.95.062003, 1602.03845.

[100] D. Tuyenbayev et al., Improving LIGO calibration accuracy by tracking and compensating for slow tempor Class. Quant. Grav. **34**(1), 015002 (2017), doi:10.1088/0264-9381/34/1/015002, 1608.05134.

[101] A. Viets et al., Reconstructing the calibrated strain signal in the Advanced LIGO detectors, Class. Quant. Grav. **35**(9), 095015 (2018), doi:10.1088/1361-6382/aab658, 1710.09973.

[102] S. Vitale, W. Del Pozzo, T. G. F. Li, C. Van Den Broeck, I. Mandel, B. Aylott and J. Veitch, Effect of calibration errors on Bayesian parameter estimation for gravitational wave signals from inspiral Phys. Rev. D **85**, 064034 (2012), doi:10.1103/PhysRevD.85.064034, 1111.3044.

[103] E. Payne, C. Talbot, P. D. Lasky, E. Thrane and J. S. Kissel, Gravitational-wave astronomy with a physical calibration model, Phys. Rev. D **102**, 122004 (2020), doi:10.1103/PhysRevD.102.122004, 2009.10193.

[104] E. D. Hall, C. Cahillane, K. Izumi, R. J. E. Smith and R. X. Adhikari, Systematic calibration error requirements for gravitational-wave detectors via the Cramér–Rao bound, Class. Quant. Grav. **36**(20), 205006 (2019), doi:10.1088/1361-6382/ab368c, 1712.09719.

[105] R. Essick, Calibration uncertainty's impact on gravitational-wave observations, Phys. Rev. D **105**(8), 082002 (2022), doi:10.1103/PhysRevD.105.082002, 2202.00823.

[106] E. Goetz et al., Precise calibration of LIGO test mass actuators using photon radiation pressure, Class. Quant. Grav. **26**, 245011 (2009), doi:10.1088/0264-9381/26/24/245011, 0910.5591.

[107] S. Karki et al., The Advanced LIGO Photon Calibrators, Rev. Sci. Instrum. **87**(11), 114503 (2016), doi:10.1063/1.4967303, 1608.05055.

[108] D. Bhattacharjee, Y. Lecoeuche, S. Karki, J. Betzwieser, V. Bossilkov, S. Kandhasamy, E. Payne and R. L. Savage, Fiducial displacements with improved accuracy for the global network of gravitational wave detectors, Class. Quant. Grav. **38**(1), 015009 (2021), doi:10.1088/1361-6382/aba9ed, 2006.00130.

[109] B. P. Abbott et al., Binary Black Hole Mergers in the first Advanced LIGO Observing Run, Phys. Rev. X **6**(4), 041015 (2016), doi:10.1103/PhysRevX.6.041015, [Erratum: Phys.Rev.X 8, 039903 (2018)], 1606.04856.

[110] L. Sun et al., Characterization of systematic error in Advanced LIGO calibration in the second half of O3 (2021), 2107.00129.

[111] W. Farr, B. Farr and T. Littenberg, MODELLING CALIBRATION ERRORS IN CBC WAVEFORMS (2014).

[112] S. Vitale, C.-J. Haster, L. Sun, B. Farr, E. Goetz, J. Kissel and C. Cahillane, Physical approach to the marginalization of LIGO calibration uncertainties, Phys. Rev. D **103**(6), 063016 (2021), doi:10.1103/PhysRevD.103.063016, 2009.10192.

[113] M. Mapelli, Formation Channels of Single and Binary Stellar-Mass Black Holes, Springer Singapore, Singapore, doi:10.1007/978-981-15-4702-7_16-1 (2021), 2106.00699.

[114] P. C. Peters, Gravitational radiation and the motion of two point masses, Phys. Rev. **136**, B1224 (1964), doi:10.1103/PhysRev.136.B1224.

[115] A. Tucker and C. M. Will, Residual eccentricity of inspiralling orbits at the gravitational-wave detection thr Phys. Rev. D **104**(10), 104023 (2021), doi:10.1103/PhysRevD.104.104023, 2108.12210.

[116] L. Wen, On the eccentricity distribution of coalescing black hole binaries driven by the Kozai mechanism i Astrophys. J. **598**, 419 (2003), doi:10.1086/378794, astro-ph/0211492.

[117] R. M. O'Leary, B. Kocsis and A. Loeb, Gravitational waves from scattering of stellar-mass black holes in gala Mon. Not. Roy. Astron. Soc. **395**(4), 2127 (2009), doi:10.1111/j.1365-2966.2009.14653.x, 0807.2638.

[118] Y.-B. Bae, C. Kim and H. M. Lee, Compact binaries ejected from globular clusters as gravitational wave sou Mon. Not. Roy. Astron. Soc. **440**(3), 2714 (2014), doi:10.1093/mnras/stu381, 1308.1641.

[119] F. Antonini, S. Chatterjee, C. L. Rodriguez, M. Morscher, B. Pattabiraman, V. Kalogera and F. A. Rasio, Black hole mergers and blue stragglers from hierarchical triples formed in globular cluster Astrophys. J. **816**(2), 65 (2016), doi:10.3847/0004-637X/816/2/65, 1509.05080.

[120] K. Silsbee and S. Tremaine, Lidov-Kozai Cycles with Gravitational Radiation: Merging Black Holes in Isola Astrophys. J. **836**(1), 39 (2017), doi:10.3847/1538-4357/aa5729, 1608.07642.

[121] J. Samsing, Eccentric Black Hole Mergers Forming in Globular Clusters, Phys. Rev. D **97**(10), 103014 (2018), doi:10.1103/PhysRevD.97.103014, 1711.07452.

[122] C. L. Rodriguez, P. Amaro-Seoane, S. Chatterjee and F. A. Rasio, Post-newtonian dynamics in dense star clusters: Highly eccentric, highly spinning, and repeated binary black hole mergers, Phys. Rev. Lett. **120**, 151101 (2018), doi:10.1103/PhysRevLett.120.151101.

[123] M. Zevin, I. M. Romero-Shaw, K. Kremer, E. Thrane and P. D. Lasky, Implications of Eccentric Observations on Binary Black Hole Formation Channels, Astrophys. J. Lett. **921**(2), L43 (2021), doi:10.3847/2041-8213/ac32dc, 2106.09042.

[124] M. Dall'Amico, M. Mapelli, S. Torniamenti and M. A. Sedda, Eccentric black hole mergers via three-body interactions in young, globular and nuclear star clusters (2023), 2303.07421.

[125] V. Cardoso, C. F. B. Macedo and R. Vicente, Eccentricity evolution of compact binaries and applications to g Phys. Rev. D **103**(2), 023015 (2021), doi:10.1103/PhysRevD.103.023015, 2010.15151.

[126] A. Ramos-Buades, A. Buonanno and J. Gair, Bayesian inference of binary black holes with inspiral-merger- Phys. Rev. D **108**(12), 124063 (2023), doi:10.1103/PhysRevD.108.124063, 2309.15528.

[127] I. M. Romero-Shaw, D. Gerosa and N. Loutrel, Eccentricity or spin precession? Distinguishing subdominant Mon. Not. R. Astron. Soc. **519**(4), 5352 (2023), doi:10.1093/mnras/stad031, 2211.07528.

[128] I. M. Romero-Shaw, P. D. Lasky, E. Thrane and J. C. Bustillo, GW190521: orbital eccentricity and signatures of dynamical formation in a binary black hole merger sig Astrophys. J. Lett. **903**(1), L5 (2020), doi:10.3847/2041-8213/abbe26, 2009.04771.

[129] E. O'Shea and P. Kumar, Correlations in gravitational-wave reconstructions from eccentric binaries: A case Phys. Rev. D **108**(10), 104018 (2023), doi:10.1103/PhysRevD.108.104018, 2107.07981.

[130] M. Favata, C. Kim, K. G. Arun, J. Kim and H. W. Lee, Constraining the orbital eccentricity of inspiralling compact binary systems with Advanced LIGO, Phys. Rev. D **105**(2), 023003 (2022), doi:10.1103/PhysRevD.105.023003, 2108.05861.

[131] Divyajyoti, S. Kumar, S. Tibrewal, I. M. Romero-Shaw and C. K. Mishra, Blind spots and biases: The dangers of ignoring eccentricity in gravitational-wave signals from binary bla Phys. Rev. D **109**(4), 043037 (2024), doi:10.1103/PhysRevD.109.043037, 2309.16638.

[132] P. Dutta Roy and P. Saini, Impact of unmodeled eccentricity on the tidal deformability measurement and in Phys. Rev. D **110**(2), 024002 (2024), doi:10.1103/PhysRevD.110.024002, 2403.02404.

[133] P. Saini, M. Favata and K. G. Arun, Systematic bias on parametrized tests of general relativity due to neglect of orbital eccentricity, Phys. Rev. D **106**, 084031 (2022), doi:10.1103/PhysRevD.106.084031.

[134] S. A. Bhat, P. Saini, M. Favata and K. G. Arun, Systematic bias on the inspiral-merger-ringdown consistency test due to neglect of orbital eccentricity, Phys. Rev. D **107**, 024009 (2023), doi:10.1103/PhysRevD.107.024009.

[135] P. Narayan, N. K. Johnson-McDaniel and A. Gupta, Effect of ignoring eccentricity in testing general relativity with gravitational waves, Phys. Rev. D **108**(6), 064003 (2023), doi:10.1103/PhysRevD.108.064003, 2306.04068.

[136] P. Saini, S. A. Bhat, M. Favata and K. G. Arun, Eccentricity-induced systematic error on parametrized tests of general relativity: Hierarchical Bayesian in Phys. Rev. D **109**(8), 084056 (2024), doi:10.1103/PhysRevD.109.084056, 2311.08033.

[137] M. A. Shaikh, S. A. Bhat and S. J. Kapadia, A study of the inspiral-merger-ringdown consistency test with g Phys. Rev. D **110**(2), 024030 (2024), doi:10.1103/PhysRevD.110.024030, 2402.15110.

[138] V. Gayathri, J. Healy, J. Lange, B. O'Brien, M. Szczepanczyk, I. Bartos, M. Campanelli, S. Klimenko, C. O. Lousto and R. O'Shaughnessy, Eccentricity estimate for black hole mergers with numerical relativity simulations, Nature Astron. **6**(3), 344 (2022), doi:10.1038/s41550-021-01568-w, 2009.05461.

[139] I. M. Romero-Shaw, P. D. Lasky and E. Thrane, Signs of Eccentricity in Two Gravitational-wave Signals May Indicate a Subpopulation of Dynamically Ass Astrophys. J. Lett. **921**(2), L31 (2021), doi:10.3847/2041-8213/ac3138, 2108.01284.

[140] N. Gupte et al., Evidence for eccentricity in the population of binary black holes observed by LIGO-Virgo-K (2024), 2404.14286.

[141] R. Abbott et al., GW190521: A Binary Black Hole Merger with a Total Mass of $150M_\odot$, Phys. Rev. Lett. **125**(10), 101102 (2020), doi:10.1103/PhysRevLett.125.101102, 2009.01075.

[142] E. E. Flanagan and T. Hinderer, Constraining neutron star tidal Love numbers with gravitational wave dete Phys. Rev. D **77**, 021502 (2008), doi:10.1103/PhysRevD.77.021502, 0709.1915.

[143] J. Vines, E. E. Flanagan and T. Hinderer, Post-1-Newtonian tidal effects in the gravitational waveform from Phys. Rev. D **83**, 084051 (2011), doi:10.1103/PhysRevD.83.084051, 1101.1673.

[144] Q. Henry, G. Faye and L. Blanchet, Tidal effects in the gravitational-wave phase evolution of compact binar Phys. Rev. D **102**(4), 044033 (2020), doi:10.1103/PhysRevD.102.044033, [Erratum: Phys.Rev.D 108, 089901 (2023)], 2005.13367.

[145] S. Bernuzzi, A. Nagar, M. Thierfelder and B. Brugmann, Tidal effects in binary neutron star coalescence, Phys. Rev. D **86**, 044030 (2012), doi:10.1103/PhysRevD.86.044030, 1205.3403.

[146] M. Favata, Systematic parameter errors in inspiraling neutron star binaries, Phys. Rev. Lett. **112**, 101101 (2014), doi:10.1103/PhysRevLett.112.101101, 1310.8288.

[147] L. Wade, J. D. E. Creighton, E. Ochsner, B. D. Lackey, B. F. Farr, T. B. Littenberg and V. Raymond, Systematic and statistical errors in a bayesian approach to the estimation of the neutron-star equation of Phys. Rev. D **89**(10), 103012 (2014), doi:10.1103/PhysRevD.89.103012, 1402.5156.

[148] T. Dietrich et al., Matter imprints in waveform models for neutron star binaries: Tidal and self-spin effects Phys. Rev. D **99**(2), 024029 (2019), doi:10.1103/PhysRevD.99.024029, 1804.02235.

[149] D. Bini, T. Damour and G. Faye, Effective action approach to higher-order relativistic tidal interactions in b Phys. Rev. D **85**, 124034 (2012), doi:10.1103/PhysRevD.85.124034, 1202.3565.

[150] D. Bini and T. Damour, Gravitational self-force corrections to two-body tidal interactions and the effective Phys. Rev. D **90**(12), 124037 (2014), doi:10.1103/PhysRevD.90.124037, 1409.6933.

[151] S. Bernuzzi, A. Nagar, T. Dietrich and T. Damour, Modeling the Dynamics of Tidally Interacting Binary Neutron Stars up to the Merger, Phys. Rev. Lett. **114**(16), 161103 (2015), doi:10.1103/PhysRevLett.114.161103, 1412.4553.

[152] S. Akcay, S. Bernuzzi, F. Messina, A. Nagar, N. Ortiz and P. Rettegno, Effective-one-body multipolar waveform for tidally interacting binary neutron stars up to merger, Phys. Rev. D **99**(4), 044051 (2019), doi:10.1103/PhysRevD.99.044051, 1812.02744.

[153] J. Steinhoff, T. Hinderer, A. Buonanno and A. Taracchini, Dynamical Tides in General Relativity: Effective Action and Effective-One-Body Hamiltonian, Phys. Rev. D **94**(10), 104028 (2016), doi:10.1103/PhysRevD.94.104028, 1608.01907.

[154] T. Hinderer et al., Effects of neutron-star dynamic tides on gravitational waveforms within the effective-or
Phys. Rev. Lett. **116**(18), 181101 (2016), doi:10.1103/PhysRevLett.116.181101, 1602.00599.

[155] T. Dietrich, S. Bernuzzi and W. Tichy, Closed-form tidal approximants for binary neutron star gravitational
Phys. Rev. D **96**(12), 121501 (2017), doi:10.1103/PhysRevD.96.121501, 1706.02969.

[156] A. Abac, T. Dietrich, A. Buonanno, J. Steinhoff and M. Ujevic, New and robust gravitational-waveform model for high-mass-ratio binary neutron star systems with dyna
Phys. Rev. D **109**(2), 024062 (2024), doi:10.1103/PhysRevD.109.024062, 2311.07456.

[157] T. Hinderer, Tidal Love numbers of neutron stars, Astrophys. J. **677**, 1216 (2008), doi:10.1086/533487, 0711.2420.

[158] C. Pacilio, A. Maselli, M. Fasano and P. Pani, Ranking Love Numbers for the Neutron Star Equation of State
Phys. Rev. Lett. **128**(10), 101101 (2022), doi:10.1103/PhysRevLett.128.101101, 2104.10035.

[159] R. Kashyap, A. Dhani and B. Sathyaprakash, Systematic errors due to quasiuniversal relations in binary ne
Phys. Rev. D **106**(12), 123001 (2022), doi:10.1103/PhysRevD.106.123001, 2209.02757.

[160] M. Evans et al., Cosmic Explorer: A Submission to the NSF MPSAC ngGW Subcommittee (2023), 2306.13745.

[161] R. Huxford, R. Kashyap, S. Borhanian, A. Dhani, I. Gupta and B. S. Sathyaprakash, Accuracy of neutron star radius measurement with the next generation of terrestrial gravitational-wave o
Phys. Rev. D **109**(10), 103035 (2024), doi:10.1103/PhysRevD.109.103035, 2307.05376.

[162] A. Samajdar and T. Dietrich, Waveform systematics for binary neutron star gravitational wave signals: effe
Phys. Rev. D **98**(12), 124030 (2018), doi:10.1103/PhysRevD.98.124030, 1810.03936.

[163] R. Gamba, M. Breschi, S. Bernuzzi, M. Agathos and A. Nagar, Waveform systematics in the gravitational-wave inference of tidal parameters and equation of state from
Phys. Rev. D **103**(12), 124015 (2021), doi:10.1103/PhysRevD.103.124015, 2009.08467.

[164] J. S. Read, Waveform uncertainty quantification and interpretation for gravitational-wave astronomy, Class. Quant. Grav. **40**(13), 135002 (2023), doi:10.1088/1361-6382/acd29b, 2301.06630.

[165] G. Pratten, P. Schmidt and N. Williams, Impact of Dynamical Tides on the Reconstruction of the Neutron S
Phys. Rev. Lett. **129**(8), 081102 (2022), doi:10.1103/PhysRevLett.129.081102, 2109.07566.

[166] N. Williams, G. Pratten and P. Schmidt, Prospects for distinguishing dynamical tides in inspiralling binary r
Phys. Rev. D **105**(12), 123032 (2022), doi:10.1103/PhysRevD.105.123032, 2203.00623.

[167] W. C. G. Ho and D. Lai, Resonant tidal excitations of rotating neutron stars in coalescing binaries, Mon. Not. Roy. Astron. Soc. **308**, 153 (1999), doi:10.1046/j.1365-8711.1999.02703.x, astro-ph/9812116.

[168] S. Ma, H. Yu and Y. Chen, Excitation of f-modes during mergers of spinning binary neutron star, Phys. Rev. D **101**(12), 123020 (2020), doi:10.1103/PhysRevD.101.123020, 2003.02373.

[169] D. Lai, Secular instability of g modes in rotating neutron stars, Mon. Not. Roy. Astron. Soc. **307**, 1001 (1999), doi:10.1046/j.1365-8711.1999.02723.x, astro-ph/9806378.

[170] J. Steinhoff, T. Hinderer, T. Dietrich and F. Foucart, Spin effects on neutron star fundamental-mode dynamical tides: Phenomenology and comparison to nu
Phys. Rev. Res. **3**(3), 033129 (2021), doi:10.1103/PhysRevResearch.3.033129, 2103.06100.

[171] T. Abdelsalhin, L. Gualtieri and P. Pani, Post-Newtonian spin-tidal couplings for compact binaries, Phys. Rev. D **98**(10), 104046 (2018), doi:10.1103/PhysRevD.98.104046, 1805.01487.

[172] E. Poisson, Gravitational waves from inspiraling compact binaries: The Quadrupole moment term, Phys. Rev. D **57**, 5287 (1998), doi:10.1103/PhysRevD.57.5287, gr-qc/9709032.

[173] N. V. Krishnendu, K. G. Arun and C. K. Mishra, Testing the binary black hole nature of a compact binary coalescence, Phys. Rev. Lett. **119**(9), 091101 (2017), doi:10.1103/PhysRevLett.119.091101, 1701.06318.

[174] A. Nagar et al., Time-domain effective-one-body gravitational waveforms for coalescing compact binaries
Phys. Rev. D **98**(10), 104052 (2018), doi:10.1103/PhysRevD.98.104052, 1806.01772.

[175] A. Nagar, F. Messina, P. Rettegno, D. Bini, T. Damour, A. Geralico, S. Akcay and S. Bernuzzi, Nonlinear-in-spin effects in effective-one-body waveform models of spin-aligned, inspiralling
Phys. Rev. D **99**(4), 044007 (2019), doi:10.1103/PhysRevD.99.044007, 1812.07923.

[176] H. Yu, N. N. Weinberg, P. Arras, J. Kwon and T. Venumadhav, Beyond the linear tide: impact of the non-linear tidal response of neutron stars on gravitational wavefor
Mon. Not. Roy. Astron. Soc. **519**(3), 4325 (2023), doi:10.1093/mnras/stac3614, 2211.07002.

[177] K. Kawaguchi, K. Kyutoku, H. Nakano and M. Shibata, Extracting the cutoff frequency in the gravitational-wave spectrum of black hole-neutron star mergers (2017), 1709.02754.

[178] X. Jiménez Forteza, T. Abdelsalhin, P. Pani and L. Gualtieri, Impact of high-order tidal terms on binary neutron-star waveforms, Phys. Rev. D **98**(12), 124014 (2018), doi:10.1103/PhysRevD.98.124014, 1807.08016.

[179] D. Lai, Resonant oscillations and tidal heating in coalescing binary neutron stars, Mon. Not. Roy. Astron. Soc. **270**, 611 (1994), doi:10.1093/mnras/270.3.611, astro-ph/9404062.

[180] H. Yu and N. N. Weinberg, Resonant tidal excitation of superfluid neutron stars in coalescing binaries, Mon. Not. Roy. Astron. Soc. **464**(3), 2622 (2017), doi:10.1093/mnras/stw2552, 1610.00745.

[181] H. Yu and N. N. Weinberg, Dynamical tides in coalescing superfluid neutron star binaries with hyperon co
Mon. Not. Roy. Astron. Soc. **470**(1), 350 (2017), doi:10.1093/mnras/stx1188, 1705.04700.

[182] D. Tsang, J. S. Read, T. Hinderer, A. L. Piro and R. Bondarescu, Resonant Shattering of Neutron Star Crusts, Phys. Rev. Lett. **108**, 011102 (2012), doi:10.1103/PhysRevLett.108.011102, 1110.0467.

[183] Z. Pan, Z. Lyu, B. Bonga, N. Ortiz and H. Yang, Probing Crust Meltdown in Inspiraling Binary Neutron Stars, Phys. Rev. Lett. **125**(20), 201102 (2020), doi:10.1103/PhysRevLett.125.201102, 2003.03330.

[184] S. Y. Lau and K. Yagi, Probing hybrid stars with gravitational waves via interfacial modes, Phys. Rev. D **103**(6), 063015 (2021), doi:10.1103/PhysRevD.103.063015, 2012.13000.

[185] E. Poisson, Gravitomagnetic tidal resonance in neutron-star binary inspirals, Phys. Rev. D **101**(10), 104028 (2020), doi:10.1103/PhysRevD.101.104028, 2003.10427.

[186] E. E. Flanagan and E. Racine, Gravitomagnetic resonant excitation of Rossby modes in coalescing neutron Phys. Rev. D **75**, 044001 (2007), doi:10.1103/PhysRevD.75.044001, gr-qc/0601029.

[187] S. Ma, H. Yu and Y. Chen, Detecting resonant tidal excitations of Rossby modes in coalescing neutron-star Phys. Rev. D **103**(6), 063020 (2021), doi:10.1103/PhysRevD.103.063020, 2010.03066.

[188] P. K. Gupta, J. Steinhoff and T. Hinderer, Effect of dynamical gravitomagnetic tides on measurability of tida Phys. Rev. D **108**(12), 124040 (2023), doi:10.1103/PhysRevD.108.124040, 2302.11274.

[189] C. Cutler and L. Lindblom, The Effect of Viscosity on Neutron Star Oscillations, Astrophysical Journal **314**, 234 (1987), doi:10.1086/165052.

[190] P. B. Jones, Bulk viscosity of neutron-star matter, Phys. Rev. D **64**(8), 084003 (2001), doi:10.1103/PhysRevD.64.084003.

[191] M. D. Duez, Y. T. Liu, S. L. Shapiro and B. C. Stephens, General relativistic hydrodynamics with viscosity: Contraction, catastrophic collapse, and disk formation Phys. Rev. D **69**, 104030 (2004), doi:10.1103/PhysRevD.69.104030, astro-ph/0402502.

[192] G. Camelio, L. Gavassino, M. Antonelli, S. Bernuzzi and B. Haskell, Simulating bulk viscosity in neutron stars. I. Formalism, Phys. Rev. D **107**(10), 103031 (2023), doi:10.1103/PhysRevD.107.103031, 2204.11809.

[193] A. Bauswein, H. T. Janka and R. Oechslin, Testing Approximations of Thermal Effects in Neutron Star Merg Phys. Rev. D **82**, 084043 (2010), doi:10.1103/PhysRevD.82.084043, 1006.3315.

[194] C. Constantinou, B. Muccioli, M. Prakash and J. M. Lattimer, Degenerate limit thermodynamics beyond leading order for models of dense matter, Annals Phys. **363**, 533 (2015), doi:10.1016/j.aop.2015.10.003, 1507.07874.

[195] A. Carbone and A. Schwenk, Ab initio constraints on thermal effects of the nuclear equation of state, Phys. Rev. C **100**(2), 025805 (2019), doi:10.1103/PhysRevC.100.025805, 1904.00924.

[196] C. Raithel, V. Paschalidis and F. Özel, Realistic finite-temperature effects in neutron star merger simulation Phys. Rev. D **104**(6), 063016 (2021), doi:10.1103/PhysRevD.104.063016, 2104.07226.

[197] J. Fields, A. Prakash, M. Breschi, D. Radice, S. Bernuzzi and A. da Silva Schneider, Thermal Effects in Binary Neutron Star Mergers, Astrophys. J. Lett. **952**(2), L36 (2023), doi:10.3847/2041-8213/ace5b2, 2302.11359.

[198] J. M. Lattimer, The nuclear equation of state and neutron star masses, Ann. Rev. Nucl. Part. Sci. **62**, 485 (2012), doi:10.1146/annurev-nucl-102711-095018, 1305.3510.

[199] G. Baym, T. Hatsuda, T. Kojo, P. D. Powell, Y. Song and T. Takatsuka, From hadrons to quarks in neutron stars: a review, Rep. Prog. Phys. **81**(5), 056902 (2018), doi:10.1088/1361-6633/aaae14.

[200] M. Deliyergiyev, A. Del Popolo, L. Tolos, M. Le Delliou, X. Lee and F. Burgio, Dark compact objects: an extensive overview, Phys. Rev. D **99**(6), 063015 (2019), doi:10.1103/PhysRevD.99.063015, 1903.01183.

[201] B. K. K. Lee, M.-c. Chu and L.-M. Lin, Could the GW190814 Secondary Component Be a Bosonic Dark Matt
Astrophys. J. **922**(2), 242 (2021), doi:10.3847/1538-4357/ac2735, 2110.05538.

[202] B. Kain, Dark matter admixed neutron stars, Phys. Rev. D **103**(4), 043009 (2021), doi:10.1103/PhysRevD.103.043009, 2102.08257.

[203] K.-L. Leung, M.-c. Chu and L.-M. Lin, Tidal deformability of dark matter admixed neutron stars, Phys. Rev. D **105**(12), 123010 (2022), doi:10.1103/PhysRevD.105.123010, 2207.02433.

[204] J. L. Ripley, A. Hegade K. R. and N. Yunes, Probing internal dissipative processes of neutron stars with grav
Phys. Rev. D **108**(10), 103037 (2023), doi:10.1103/PhysRevD.108.103037, 2306.15633.

[205] J. L. Ripley, A. Hegade K. R., R. S. Chandramouli and N. Yunes, A constraint on the dissipative tidal deformability of neutron stars, Nature Astron. **8**(10), 1277 (2024), doi:10.1038/s41550-024-02323-7, 2312.11659.

[206] A. Hegade K. R., J. L. Ripley and N. Yunes, Dynamical tidal response of nonrotating relativistic stars, Phys. Rev. D **109**(10), 104064 (2024), doi:10.1103/PhysRevD.109.104064, 2403.03254.

[207] M. Breschi, S. Bernuzzi, K. Chakravarti, A. Camilletti, A. Prakash and A. Perego, Kilohertz gravitational waves from binary neutron star mergers: Numerical-relativity informed postmerge
Phys. Rev. D **109**(6), 064009 (2024), doi:10.1103/PhysRevD.109.064009, 2205.09112.

[208] M. Breschi, S. Bernuzzi, F. Zappa, M. Agathos, A. Perego, D. Radice and A. Nagar, kilohertz gravitational waves from binary neutron star remnants: time-domain model and constraints on extreme matter (10), 104029 (2019), doi:10.1103/PhysRevD.100.104029, 1908.11418.

[209] T. Soultanis, A. Bauswein and N. Stergioulas, Analytic models of the spectral properties of gravitational wa
Phys. Rev. D **105**(4), 043020 (2022), doi:10.1103/PhysRevD.105.043020, 2111.08353.

[210] A. Prakash, D. Radice, D. Logoteta, A. Perego, V. Nedora, I. Bombaci, R. Kashyap, S. Bernuzzi and A. Endrizzi, Signatures of deconfined quark phases in binary neutron star mergers, Physical Review D **104**(8), 083029 (2021), doi:10.1103/PhysRevD.104.083029.

[211] A. Prakash, I. Gupta, M. Breschi, R. Kashyap, D. Radice, S. Bernuzzi, D. Logoteta and B. S. Sathyaprakash, Detectability of QCD phase transitions in binary neutron star mergers: Bayesian inference with the next g Phys. Rev. D **109**(10), 103008 (2024), doi:10.1103/PhysRevD.109.103008, 2310.06025.

[212] K. Yagi and N. Yunes, I-Love-Q, Science **341**, 365 (2013), doi:10.1126/science.1236462, 1302.4499.

[213] K. Yagi and N. Yunes, I-Love-Q Relations in Neutron Stars and their Applications to Astrophysics, Gravitati Phys. Rev. D **88**(2), 023009 (2013), doi:10.1103/PhysRevD.88.023009, 1303.1528.

[214] K. Yagi and N. Yunes, Approximate Universal Relations for Neutron Stars and Quark Stars, Phys. Rept. **681**, 1 (2017), doi:10.1016/j.physrep.2017.03.002, 1608.02582.

[215] K. Yagi, Multipole Love Relations, Phys. Rev. D **89**(4), 043011 (2014), doi:10.1103/PhysRevD.89.043011, [Erratum: Phys.Rev.D 96, 129904 (2017), Erratum: Phys.Rev.D 97, 129901 (2018)], 1311.0872.

[216] K. Yagi and N. Yunes, Binary Love Relations, Class. Quant. Grav. **33**(13), 13LT01 (2016), doi:10.1088/0264-9381/33/13/13LT01, 1512.02639.

[217] K. Yagi and N. Yunes, Approximate Universal Relations among Tidal Parameters for Neutron Star Binaries, Class. Quant. Grav. **34**(1), 015006 (2017), doi:10.1088/1361-6382/34/1/015006, 1608.06187.

[218] K. Chatziioannou, C.-J. Haster and A. Zimmerman, Measuring the neutron star tidal deformability with equation-of-state-independent relations and gravitati Phys. Rev. D **97**(10), 104036 (2018), doi:10.1103/PhysRevD.97.104036, 1804.03221.

[219] B. P. Abbott et al., GW170817: Measurements of neutron star radii and equation of state, Phys. Rev. Lett. **121**(16), 161101 (2018), doi:10.1103/PhysRevLett.121.161101, 1805.11581.

[220] Z. Carson, K. Chatziioannou, C.-J. Haster, K. Yagi and N. Yunes, Equation-of-state insensitive relations after GW170817, Phys. Rev. D **99**(8), 083016 (2019), doi:10.1103/PhysRevD.99.083016, 1903.03909.

[221] T. Gupta, B. Majumder, K. Yagi and N. Yunes, I-Love-Q Relations for Neutron Stars in dynamical Chern Simons Gravity, Class. Quant. Grav. **35**(2), 025009 (2018), doi:10.1088/1361-6382/aa9c68, 1710.07862.

[222] A. Maselli, P. Pnigouras, N. G. Nielsen, C. Kouvaris and K. D. Kokkotas, Dark stars: gravitational and electromagnetic observables, Phys. Rev. D **96**(2), 023005 (2017), doi:10.1103/PhysRevD.96.023005, 1704.07286.

[223] E. Berti, V. De Luca, L. Del Grosso and P. Pani, Tidal Love numbers and approximate universal relations for fermion soliton stars (2024), 2404.06979.

[224] K. Nordtvedt, Equivalence principle for massive bodies. ii. theory, Phys. Rev. **169**, 1017 (1968), doi:10.1103/PhysRev.169.1017.

[225] D. M. Eardley, Observable effects of a scalar gravitational field in a binary pulsar, Astrophysical Journal **196**, L59 (1975), doi:10.1086/181744.

[226] C. M. Will and H. W. Zaglauer, Gravitational Radiation, Close Binary Systems, and the Brans-dicke Theory Astrophys. J. **346**, 366 (1989), doi:10.1086/168016.

[227] C. M. Will, Gravitational Radiation from Binary Systems in Alternative Metric Theories of Gravity: Dipole Astrophys. J. **214**, 826 (1977), doi:10.1086/155313.

[228] T. Damour and G. Esposito-Farese, Tensor multiscalar theories of gravitation, Class. Quant. Grav. **9**, 2093 (1992), doi:10.1088/0264-9381/9/9/015.

[229] C. M. Will, Theory and Experiment in Gravitational Physics, ISBN 978-0-511-56424-6, 978-0-521-43973-2, doi:10.1017/CBO9780511564246 (1993).

[230] M. J. Fitchett, The influence of gravitational wave momentum losses on the centre of mass motion of a Ne Mon. Not. R. Astron. Soc. **203**, 1049 (1983).

[231] M. Favata, S. A. Hughes and D. E. Holz, How black holes get their kicks: Gravitational radiation recoil revi Astrophys. J. Lett. **607**, L5 (2004), doi:10.1086/421552, astro-ph/0402056.

[232] M. Favata, Gravitational-wave memory revisited: memory from the merger and recoil of binary black hol J. Phys. Conf. Ser. **154**, 012043 (2009), doi:10.1088/1742-6596/154/1/012043, 0811.3451.

[233] A. Torres-Orjuela, X. Chen and P. Amaro-Seoane, Phase shift of gravitational waves induced by aberration, Phys. Rev. D **101**, 083028 (2020), doi:10.1103/PhysRevD.101.083028, 2001.00721.

[234] P. Mahapatra, M. Favata and K. G. Arun, Testing general relativity via direct measurement of black hole kic Phys. Rev. D **110**(8), 084041 (2024), doi:10.1103/PhysRevD.110.084041, 2308.08319.

[235] D. Gerosa and C. J. Moore, Black hole kicks as new gravitational wave observables, Phys. Rev. Lett. **117**(1), 011101 (2016), doi:10.1103/PhysRevLett.117.011101, 1606.04226.

[236] M. Boyle, Transformations of asymptotic gravitational-wave data, Phys. Rev. D **93**(8), 084031 (2016), doi:10.1103/PhysRevD.93.084031, 1509.00862.

[237] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, D. Gerosa, L. C. Stein, L. E. Kidder and H. P. Pfeiffer, Surrogate models for precessing binary black hole simulations with unequal masses, Phys. Rev. Research. **1**, 033015 (2019), doi:10.1103/PhysRevResearch.1.033015, 1905.09300.

[238] V. Varma, M. Isi and S. Biscoveanu, Extracting the Gravitational Recoil from Black Hole Merger Signals, Phys. Rev. Lett. **124**(10), 101104 (2020), doi:10.1103/PhysRevLett.124.101104, 2002.00296.

[239] S. A. Hughes and K. Menou, Golden binaries for LISA: Robust probes of strong-field gravity, Astrophys. J. **623**, 689 (2005), doi:10.1086/428826, astro-ph/0410148.

[240] A. Ghosh et al., Testing general relativity using golden black-hole binaries, Phys. Rev. D **94**(2), 021101 (2016), doi:10.1103/PhysRevD.94.021101, 1602.02453.

[241] M. Isi, M. Giesler, W. M. Farr, M. A. Scheel and S. A. Teukolsky, Testing the no-hair theorem with GW150914, Phys. Rev. Lett. **123**(11), 111102 (2019), doi:10.1103/PhysRevLett.123.111102, 1905.00869.

[242] G. Carullo, W. Del Pozzo and J. Veitch, Observational Black Hole Spectroscopy: A time-domain multimode Phys. Rev. D **99**(12), 123029 (2019), doi:10.1103/PhysRevD.99.123029, [Erratum: Phys.Rev.D 100, 089903 (2019)], 1902.07527.

[243] C. V. Vishveshwara, Scattering of Gravitational Radiation by a Schwarzschild Black-hole, Nature **227**, 936 (1970), doi:10.1038/227936a0.

[244] W. H. Press, Long Wave Trains of Gravitational Waves from a Vibrating Black Hole, Astrophys. J. Lett. **170**, L105 (1971), doi:10.1086/180849.

[245] S. L. Detweiler, BLACK HOLES AND GRAVITATIONAL WAVES. III. THE RESONANT FREQUENCIES OF RO Astrophys. J. **239**, 292 (1980), doi:10.1086/158109.

[246] O. Dreyer, B. J. Kelly, B. Krishnan, L. S. Finn, D. Garrison and R. Lopez-Aleman, Black hole spectroscopy: Testing general relativity through gravitational wave observations, Class. Quant. Grav. **21**, 787 (2004), doi:10.1088/0264-9381/21/4/003, gr-qc/0309007.

[247] E. Berti, V. Cardoso and C. M. Will, On gravitational-wave spectroscopy of massive black holes with the spa Phys. Rev. D **73**, 064030 (2006), doi:10.1103/PhysRevD.73.064030, gr-qc/0512160.

[248] A. Dhani, S. Völkel, A. Buonanno, H. Estelles, J. Gair, H. P. Pfeiffer, L. Pompili and A. Toubiana, Systematic Biases in Estimating the Properties of Black Holes Due to Inaccurate Gravitationa (2024), 2404.05811.

[249] S. Gossan, J. Veitch and B. S. Sathyaprakash, Bayesian model selection for testing the no-hair theorem with Phys. Rev. D **85**, 124056 (2012), doi:10.1103/PhysRevD.85.124056, 1111.5819.

[250] V. Baibhav, M. H.-Y. Cheung, E. Berti, V. Cardoso, G. Carullo, R. Cotesta, W. Del Pozzo and F. Duque, Agnostic black hole spectroscopy: Quasinormal mode content of numerical relativity waveforms and limi Phys. Rev. D **108**(10), 104020 (2023), doi:10.1103/PhysRevD.108.104020, 2302.03050.

[251] V. Baibhav, E. Berti, V. Cardoso and G. Khanna, Black Hole Spectroscopy: Systematic Errors and Ringdown Energy Estimates, Phys. Rev. D **97**(4), 044048 (2018), doi:10.1103/PhysRevD.97.044048, 1710.02156.

[252] V. Baibhav and E. Berti, Multimode black hole spectroscopy, Phys. Rev. D **99**(2), 024005 (2019), doi:10.1103/PhysRevD.99.024005, 1809.03500.

[253] M. Giesler, M. Isi, M. A. Scheel and S. Teukolsky, Black Hole Ringdown: The Importance of Overtones, Phys. Rev. X **9**(4), 041060 (2019), doi:10.1103/PhysRevX.9.041060, 1903.08284.

[254] A. Dhani, Importance of mirror modes in binary black hole ringdown waveform, Phys. Rev. D **103**(10), 104048 (2021), doi:10.1103/PhysRevD.103.104048, 2010.08602.

[255] A. Dhani and B. S. Sathyaprakash, Overtones, mirror modes, and mode-mixing in binary black hole merge (2021), 2107.14195.

[256] L. Magaña Zertuche et al., High precision ringdown modeling: Multimode fits and BMS frames, Phys. Rev. D **105**(10), 104015 (2022), doi:10.1103/PhysRevD.105.104015, 2110.15922.

[257] S. Ma, K. Mitman, L. Sun, N. Deppe, F. Hébert, L. E. Kidder, J. Moxon, W. Throwe, N. L. Vu and Y. Chen, Quasinormal-mode filters: A new approach to analyze the gravitational-wave ringdown of binary black-h Phys. Rev. D **106**(8), 084036 (2022), doi:10.1103/PhysRevD.106.084036, 2207.10870.

[258] S. Ma, L. Sun and Y. Chen, Black Hole Spectroscopy by Mode Cleaning, Phys. Rev. Lett. **130**(14), 141401 (2023), doi:10.1103/PhysRevLett.130.141401, 2301.06705.

[259] M. H.-Y. Cheung, E. Berti, V. Baibhav and R. Cotesta, Extracting linear and nonlinear quasinormal modes from black hole merger simulations, Phys. Rev. D **109**(4), 044069 (2024), doi:10.1103/PhysRevD.109.044069, 2310.04489.

[260] K. Takahashi and H. Motohashi, Iterative extraction of overtones from black hole ringdown, Class. Quant. Grav. **41**(19), 195023 (2024), doi:10.1088/1361-6382/ad72c9, 2311.12762.

[261] Y. Qiu, X. J. Forteza and P. Mourier, Linear versus nonlinear modeling of black hole ringdowns, Phys. Rev. D **109**(6), 064075 (2024), doi:10.1103/PhysRevD.109.064075, 2312.15904.

[262] T. A. Clarke et al., Toward a self-consistent framework for measuring black hole ringdowns, Phys. Rev. D **109**(12), 124030 (2024), doi:10.1103/PhysRevD.109.124030, 2402.02819.

[263] M. Isi and W. M. Farr, Analyzing black-hole ringdowns (2021), 2107.05609.

[264] S. R. Coleman, J. Preskill and F. Wilczek, Quantum hair on black holes, Nucl. Phys. B **378**, 175 (1992), doi:10.1016/0550-3213(92)90008-Y, hep-th/9201059.

[265] M. Isi and W. M. Farr, Revisiting the ringdown of GW150914 (2022), 2202.02941.

[266] E. Finch and C. J. Moore, Searching for a ringdown overtone in GW150914, Phys. Rev. D **106**(4), 043005 (2022), doi:10.1103/PhysRevD.106.043005, 2205.07809.

[267] S. Ma, L. Sun and Y. Chen, Using rational filters to uncover the first ringdown overtone in GW150914, Phys. Rev. D **107**(8), 084010 (2023), doi:10.1103/PhysRevD.107.084010, 2301.06639.

[268] R. Cotesta, G. Carullo, E. Berti and V. Cardoso, Analysis of Ringdown Overtones in GW150914, Phys. Rev. Lett. **129**(11), 111102 (2022), doi:10.1103/PhysRevLett.129.111102, 2201.00822.

[269] M. Crisostomi, K. Dey, E. Barausse and R. Trotta, Neural posterior estimation with guaranteed exact coverage: The ringdown of GW150914, Phys. Rev. D **108**(4), 044029 (2023), doi:10.1103/PhysRevD.108.044029, 2305.18528.

[270] V. Gennari, G. Carullo and W. Del Pozzo, Searching for ringdown higher modes with a numerical relativity- Eur. Phys. J. C **84**(3), 233 (2024), doi:10.1140/epjc/s10052-024-12550-x, 2312.12515.

[271] A. Correia, Y.-F. Wang, J. Westerweck and C. D. Capano, Low evidence for ringdown overtone in GW150914 when marginalizing over time and sky location uncer Phys. Rev. D **110**(4), L041501 (2024), doi:10.1103/PhysRevD.110.L041501, 2312.14118.

[272] L. Sberna, P. Bosch, W. E. East, S. R. Green and L. Lehner, Nonlinear effects in the black hole ringdown: Absorption-induced mode excitation, Phys. Rev. D **105**(6), 064046 (2022), doi:10.1103/PhysRevD.105.064046, 2112.11168.

[273] M. Lagos and L. Hui, Generation and propagation of nonlinear quasinormal modes of a Schwarzschild bla Phys. Rev. D **107**(4), 044040 (2023), doi:10.1103/PhysRevD.107.044040, 2208.07379.

[274] P. J. Nee, S. H. Völkel and H. P. Pfeiffer, Role of black hole quasinormal mode overtones for ringdown analy Phys. Rev. D **108**(4), 044032 (2023), doi:10.1103/PhysRevD.108.044032, 2302.06634.

[275] H. Zhu, J. L. Ripley, A. Cárdenas-Avendaño and F. Pretorius, Challenges in quasinormal mode extraction: Perspectives from numerical solutions to the Teukolsky equa Phys. Rev. D **109**(4), 044010 (2024), doi:10.1103/PhysRevD.109.044010, 2309.13204.

[276] S. Bhagwat, X. J. Forteza, P. Pani and V. Ferrari, Ringdown overtones, black hole spectroscopy, and no-hair theorem tests, Phys. Rev. D **101**(4), 044033 (2020), doi:10.1103/PhysRevD.101.044033, 1910.08708.

[277] E. Berti, J. Cardoso, V. Cardoso and M. Cavaglia, Matched-filtering and parameter estimation of ringdown waveforms, Phys. Rev. D **76**, 104044 (2007), doi:10.1103/PhysRevD.76.104044, 0707.1202.

[278] E. Berti, A. Sesana, E. Barausse, V. Cardoso and K. Belczynski, Spectroscopy of Kerr black holes with Earth- and space-based interferometers, Phys. Rev. Lett. **117**(10), 101102 (2016), doi:10.1103/PhysRevLett.117.101102, 1605.09286.

[279] R. Brito, A. Buonanno and V. Raymond, Black-hole Spectroscopy by Making Full Use of Gravitational-Wave Phys. Rev. D **98**(8), 084038 (2018), doi:10.1103/PhysRevD.98.084038, 1805.00293.

[280] M. Cabero, J. Westerweck, C. D. Capano, S. Kumar, A. B. Nielsen and B. Krishnan, Black hole spectroscopy in the next decade, Phys. Rev. D **101**(6), 064044 (2020), doi:10.1103/PhysRevD.101.064044, 1911.01361.

[281] X. Jiménez Forteza, S. Bhagwat, P. Pani and V. Ferrari, Spectroscopy of binary black hole ringdown using overtones and angular modes, Phys. Rev. D **102**(4), 044053 (2020), doi:10.1103/PhysRevD.102.044053, 2005.03260.

[282] A. Ghosh, R. Brito and A. Buonanno, Constraints on quasinormal-mode frequencies with LIGO-Virgo binar Phys. Rev. D **103**(12), 124041 (2021), doi:10.1103/PhysRevD.103.124041, 2104.01906.

[283] S. Bhagwat, C. Pacilio, P. Pani and M. Mapelli, Landscape of stellar-mass black-hole spectroscopy with third-generation gravitational-wave detectors, Phys. Rev. D **108**(4), 043019 (2023), doi:10.1103/PhysRevD.108.043019, 2304.02283.

[284] I. Ota and C. Chirenti, Overtones or higher harmonics? Prospects for testing the no-hair theorem with gra Phys. Rev. D **101**(10), 104005 (2020), doi:10.1103/PhysRevD.101.104005, 1911.00440.

[285] I. Ota and C. Chirenti, Black hole spectroscopy horizons for current and future gravitational wave detector Phys. Rev. D **105**(4), 044015 (2022), doi:10.1103/PhysRevD.105.044015, 2108.01774.

[286] E. Thrane, P. D. Lasky and Y. Levin, Challenges testing the no-hair theorem with gravitational waves, Phys. Rev. D **96**(10), 102004 (2017), doi:10.1103/PhysRevD.96.102004, 1706.05152.

[287] J. C. Bustillo, P. D. Lasky and E. Thrane, Black-hole spectroscopy, the no-hair theorem, and GW150914: Ker Phys. Rev. D **103**(2), 024041 (2021), doi:10.1103/PhysRevD.103.024041.

[288] S. Bhagwat, X. J. Forteza, P. Pani and V. Ferrari, Ringdown overtones, black hole spectroscopy, and no-hair theorem tests, Phys. Rev. D **101**(4), 044033 (2020), doi:10.1103/PhysRevD.101.044033, 1910.08708.

[289] K. Mitman et al., Nonlinearities in Black Hole Ringdowns, Phys. Rev. Lett. **130**(8), 081402 (2023), doi:10.1103/PhysRevLett.130.081402, 2208.07380.

[290] M. H.-Y. Cheung et al., Nonlinear Effects in Black Hole Ringdown, Phys. Rev. Lett. **130**(8), 081401 (2023), doi:10.1103/PhysRevLett.130.081401, 2208.07374.

[291] N. Khera, A. Ribes Metidieri, B. Bonga, X. Jiménez Forteza, B. Krishnan, E. Poisson, D. Pook-Kolb, E. Schnetter and H. Yang, Nonlinear Ringdown at the Black Hole Horizon, Phys. Rev. Lett. **131**(23), 231401 (2023), doi:10.1103/PhysRevLett.131.231401, 2306.11142.

[292] H. Zhu et al., Nonlinear effects in black hole ringdown from scattering experiments: Spin and initial data Phys. Rev. D **109**(10), 104050 (2024), doi:10.1103/PhysRevD.109.104050, 2401.00805.

[293] J. Redondo-Yuste, G. Carullo, J. L. Ripley, E. Berti and V. Cardoso, Spin dependence of black hole ringdown nonlinearities, Phys. Rev. D **109**(10), L101503 (2024), doi:10.1103/PhysRevD.109.L101503, 2308.14796.

[294] S. Ma and H. Yang, Excitation of quadratic quasinormal modes for Kerr black holes, Phys. Rev. D **109**(10), 104070 (2024), doi:10.1103/PhysRevD.109.104070, 2401.15516.

[295] S. Yi, A. Kuntz, E. Barausse, E. Berti, M. H.-Y. Cheung, K. Kritos and A. Maselli, Nonlinear quasinormal mode detectability with next-generation gravitational wave detectors, Phys. Rev. D **109**(12), 124029 (2024), doi:10.1103/PhysRevD.109.124029, 2403.09767.

[296] X. J. Forteza, S. Bhagwat, S. Kumar and P. Pani, Novel Ringdown Amplitude-Phase Consistency Test, Phys. Rev. Lett. **130**(2), 021001 (2023), doi:10.1103/PhysRevLett.130.021001, 2205.14910.

[297] M. Isi and W. M. Farr, Comment on "Analysis of Ringdown Overtones in GW150914", Phys. Rev. Lett. **131**(16), 169001 (2023), doi:10.1103/PhysRevLett.131.169001, 2310.13869.

[298] G. Carullo, R. Cotesta, E. Berti and V. Cardoso, Reply to Comment on "Analysis of Ringdown Overtones in GW150914", Phys. Rev. Lett. **131**, 169002 (2023), doi:10.1103/PhysRevLett.131.169002, 2310.20625.

[299] Y.-F. Wang, C. D. Capano, J. Abedi, S. Kastha, B. Krishnan, A. B. Nielsen, A. H. Nitz and J. Westerweck, A frequency-domain perspective on GW150914 ringdown overtone (2023), 2310.19645.

[300] C. D. Capano, M. Cabero, J. Westerweck, J. Abedi, S. Kastha, A. H. Nitz, Y.-F. Wang, A. B. Nielsen and B. Krishnan, Multimode Quasinormal Spectrum from a Perturbed Black Hole, Phys. Rev. Lett. **131**(22), 221402 (2023), doi:10.1103/PhysRevLett.131.221402, 2105.05238.

[301] E. Finch and C. J. Moore, Frequency-domain analysis of black-hole ringdowns, Phys. Rev. D **104**(12), 123034 (2021), doi:10.1103/PhysRevD.104.123034, 2108.09344.

[302] H. Siegel, M. Isi and W. M. Farr, Ringdown of GW190521: Hints of multiple quasinormal modes with a pre
Phys. Rev. D **108**(6), 064008 (2023), doi:10.1103/PhysRevD.108.064008, 2307.11975.

[303] E. Berti, V. Cardoso, J. A. Gonzalez, U. Sperhake, M. Hannam, S. Husa and B. Bruegmann, Inspiral, merger and ringdown of unequal mass black hole binaries: A Multipolar analysis, Phys. Rev. D **76**, 064034 (2007), doi:10.1103/PhysRevD.76.064034, gr-qc/0703053.

[304] L. E. Kidder, Using full information when computing modes of post-Newtonian waveforms from inspiralli
Phys. Rev. D **77**, 044016 (2008), doi:10.1103/PhysRevD.77.044016, 0710.0614.

[305] K. G. Arun, A. Buonanno, G. Faye and E. Ochsner, Higher-order spin effects in the amplitude and phase of gravitational waveforms emitted by inspiraling c
Phys. Rev. D **79**, 104023 (2009), doi:10.1103/PhysRevD.79.104023, [Erratum: Phys.Rev.D 84, 049901 (2011)], 0810.5336.

[306] Y. Pan, A. Buonanno, M. Boyle, L. T. Buchman, L. E. Kidder, H. P. Pfeiffer and M. A. Scheel, Inspiral-merger-ringdown multipolar waveforms of nonspinning black-hole binaries using the effe
Phys. Rev. D **84**, 124052 (2011), doi:10.1103/PhysRevD.84.124052, 1106.1021.

[307] L. Blanchet, Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries, Living Rev. Rel. **17**, 2 (2014), doi:10.12942/lrr-2014-2, 1310.1528.

[308] A. Buonanno, Y.-b. Chen and M. Vallisneri, Detecting gravitational waves from precessing binaries of spinn
Phys. Rev. D **67**, 104025 (2003), doi:10.1103/PhysRevD.67.104025, [Erratum: Phys.Rev.D 74, 029904 (2006)], gr-qc/0211087.

[309] P. Schmidt, M. Hannam and S. Husa, Towards models of gravitational waveforms from generic binaries: A
Phys. Rev. D **86**, 104063 (2012), doi:10.1103/PhysRevD.86.104063, 1207.3088.

[310] R. O'Shaughnessy, L. London, J. Healy and D. Shoemaker, Precession during merger: Strong polarization changes are observationally accessible features of strong-f
Phys. Rev. D **87**(4), 044038 (2013), doi:10.1103/PhysRevD.87.044038, 1209.3712.

[311] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten and M. Pürrer, Simple Model of Complete Precessing Black-Hole-Binary Gravitational Waveforms, Phys. Rev. Lett. **113**(15), 151101 (2014), doi:10.1103/PhysRevLett.113.151101, 1308.3271.

[312] T. Islam, S. E. Field, C.-J. Haster and R. Smith, Improved analysis of GW190412 with a precessing numerical relativity surrogate waveform model, Phys. Rev. D **103**(10), 104027 (2021), doi:10.1103/PhysRevD.103.104027, 2010.04848.

[313] M. Boyle, L. E. Kidder, S. Ossokine and H. P. Pfeiffer, Gravitational-wave modes from precessing black-hole binaries (2014), 1409.4431.

[314] A. Ramos-Buades, P. Schmidt, G. Pratten and S. Husa, Validity of common modeling approximations for precessing binary black holes with higher-order modes, Phys. Rev. D **101**(10), 103014 (2020), doi:10.1103/PhysRevD.101.103014, 2001.10936.

[315] J. E. Thompson, E. Hamilton, L. London, S. Ghosh, P. Kolitsidou, C. Hoy and M. Hannam, PhenomXO4a: a phenomenological gravitational-wave model for precessing black-hole binaries with hig Phys. Rev. D **109**(6), 063012 (2024), doi:10.1103/PhysRevD.109.063012, 2312.10025.

[316] S. Ghosh, P. Kolitsidou and M. Hannam, First frequency-domain phenomenological model of the multipole Phys. Rev. D **109**(2), 024061 (2024), doi:10.1103/PhysRevD.109.024061, 2310.16980.

[317] C. Kalaghatgi and M. Hannam, Investigating the effect of in-plane spin directions for precessing binary bla Phys. Rev. D **103**(2), 024024 (2021), doi:10.1103/PhysRevD.103.024024, 2008.09957.

[318] P. Kolitsidou, J. E. Thompson and M. Hannam, Impact of anti-symmetric contributions to signal multipoles in the measurement of black-hole spins (2024), 2402.00813.

[319] C. García-Quirós, M. Colleoni, S. Husa, H. Estellés, G. Pratten, A. Ramos-Buades, M. Mateu-Lucena and R. Jaume, Multimode frequency-domain model for the gravitational wave signal from nonprecessing black-hole bina Phys. Rev. D **102**(6), 064002 (2020), doi:10.1103/PhysRevD.102.064002, 2001.10914.

[320] L. Pompili et al., Laying the foundation of the effective-one-body waveform models SEOBNRv5: Improve Phys. Rev. D **108**(12), 124035 (2023), doi:10.1103/PhysRevD.108.124035, 2303.18039.

[321] G. Pratten et al., Computationally efficient models for the dominant and subdominant harmonic modes of Phys. Rev. D **103**(10), 104056 (2021), doi:10.1103/PhysRevD.103.104056, 2004.06503.

[322] A. Ramos-Buades, A. Buonanno, H. Estellés, M. Khalil, D. P. Mihaylov, S. Ossokine, L. Pompili and M. Shiferaw, Next generation of accurate and efficient multipolar precessing-spin effective-one-body waveforms for bin Phys. Rev. D **108**(12), 124037 (2023), doi:10.1103/PhysRevD.108.124037, 2303.18046.

[323] A. Klein, P. Jetzer and M. Sereno, Parameter estimation for coalescing massive binary black holes with LIS Phys. Rev. D **80**, 064027 (2009), doi:10.1103/PhysRevD.80.064027, 0907.3318.

[324] F. H. Shaik, J. Lange, S. E. Field, R. O'Shaughnessy, V. Varma, L. E. Kidder, H. P. Pfeiffer and D. Wysocki, Impact of subdominant modes on the interpretation of gravitational-wave signals from heavy binary blac Phys. Rev. D **101**(12), 124054 (2020), doi:10.1103/PhysRevD.101.124054, 1911.02693.

[325] N. V. Krishnendu and F. Ohme, Interplay of spin-precession and higher harmonics in the parameter estimat Phys. Rev. D **105**(6), 064012 (2022), doi:10.1103/PhysRevD.105.064012, 2110.00766.

[326] N. Loutrel, R. Brito, A. Maselli and P. Pani, Relevance of precession for tests of the black hole no hair theore Phys. Rev. D **110**(4), 044003 (2024), doi:10.1103/PhysRevD.110.044003, 2309.17404.

[327] M. Favata, Post-Newtonian corrections to the gravitational-wave memory for quasi-circular, inspiralling co Phys. Rev. D **80**, 024002 (2009), doi:10.1103/PhysRevD.80.024002, 0812.0069.

[328] D. A. Nichols, Spin memory effect for compact binaries in the post-Newtonian approximation, Phys. Rev. D **95**(8), 084048 (2017), doi:10.1103/PhysRevD.95.084048, 1702.03300.

[329] K. Mitman, J. Moxon, M. A. Scheel, S. A. Teukolsky, M. Boyle, N. Deppe, L. E. Kidder and W. Throwe, Computation of displacement and spin gravitational memory in numerical relativity, Phys. Rev. D **102**(10), 104007 (2020), doi:10.1103/PhysRevD.102.104007, 2007.11562.

[330] J. Yoo et al., Numerical relativity surrogate model with memory effects and post-Newtonian hybridization Phys. Rev. D **108**(6), 064027 (2023), doi:10.1103/PhysRevD.108.064027, 2306.03148.

[331] Y. Xu, M. Rosselló-Sastre, S. Tiwari, M. Ebersold, E. Z. Hamilton, C. García-Quirós, H. Estellés and S. Husa, Enhancing gravitational wave parameter estimation with nonlinear memory: Breaking the distance-inclir Phys. Rev. D **109**(12), 123034 (2024), doi:10.1103/PhysRevD.109.123034, 2403.00441.

[332] S. Tiwari, M. Ebersold and E. Z. Hamilton, Leveraging gravitational-wave memory to distinguish neutron s Phys. Rev. D **104**(12), 123024 (2021), doi:10.1103/PhysRevD.104.123024, 2110.11171.

[333] D. Lopez, S. Tiwari and M. Ebersold, Gravitational wave memory of compact binary coalescence in the pre Phys. Rev. D **109**(4), 043039 (2024), doi:10.1103/PhysRevD.109.043039, 2305.04761.

[334] A. M. Grant and D. A. Nichols, Outlook for detecting the gravitational-wave displacement and spin memor Phys. Rev. D **107**(6), 064056 (2023), doi:10.1103/PhysRevD.107.064056, [Erratum: Phys.Rev.D 108, 029901 (2023)], 2210.16266.

[335] A. Nagar, P. Rettegno, R. Gamba, S. Albanesi, A. Albertini and S. Bernuzzi, Analytic systematics in next generation of effective-one-body gravitational waveform models for future ol Phys. Rev. D **108**(12), 124018 (2023), doi:10.1103/PhysRevD.108.124018, 2304.09662.

[336] J. Mac Uilliam, S. Akcay and J. E. Thompson, Survey of four precessing waveform models for binary black hole systems, Phys. Rev. D **109**(8), 084077 (2024), doi:10.1103/PhysRevD.109.084077, 2402.06781.

[337] Q. Hu and J. Veitch, Assessing the model waveform accuracy of gravitational waves, Phys. Rev. D **106**(4), 044042 (2022), doi:10.1103/PhysRevD.106.044042, 2205.08448.

[338] L. Pompili, A. Buonanno and M. Pürrer, Accounting for Numerical-Relativity Calibration Uncertainty in Gra (2024), 2410.16859.

[339] C. J. Moore and J. R. Gair, Novel Method for Incorporating Model Uncertainties into Gravitational Wave Pa Phys. Rev. Lett. **113**, 251101 (2014), doi:10.1103/PhysRevLett.113.251101, 1412.3657.

[340] Z. Doctor, B. Farr, D. E. Holz and M. Pürrer, Statistical Gravitational Waveform Models: What to Simulate Next?, Phys. Rev. D **96**(12), 123011 (2017), doi:10.1103/PhysRevD.96.123011, 1706.05408.

[341] D. Williams, I. S. Heng, J. Gair, J. A. Clark and B. Khamesra, Precessing numerical relativity waveform surrogate model for binary black holes: A Gaussian process reg Phys. Rev. D **101**(6), 063011 (2020), doi:10.1103/PhysRevD.101.063011, 1903.09204.

[342] S. Khan, Probabilistic model for the gravitational wave signal from merging black holes, Phys. Rev. D **109**(10), 104045 (2024), doi:10.1103/PhysRevD.109.104045, 2403.11534.

[343] A. Nagar and P. Rettegno, Next generation: Impact of high-order analytical information on effective one bo Phys. Rev. D **104**(10), 104004 (2021), doi:10.1103/PhysRevD.104.104004, 2108.02043.

[344] C. B. Owen, C.-J. Haster, S. Perkins, N. J. Cornish and N. Yunes, Waveform accuracy and systematic uncertainties in current gravitational wave observations, Phys. Rev. D **108**(4), 044018 (2023), doi:10.1103/PhysRevD.108.044018, 2301.11941.

[345] D. A. Kosower, R. Monteiro and D. O'Connell, The SAGEX review on scattering amplitudes Chapter 14: Cla J. Phys. A **55**(44), 443015 (2022), doi:10.1088/1751-8121/ac8846, 2203.13025.

[346] N. E. J. Bjerrum-Bohr, P. H. Damgaard, L. Plante and P. Vanhove, The SAGEX review on scattering amplitudes Chapter 13: Post-Minkowskian expansion from scattering an J. Phys. A **55**(44), 443014 (2022), doi:10.1088/1751-8121/ac7a78, 2203.13024.

[347] A. Antonelli, A. Buonanno, J. Steinhoff, M. van de Meent and J. Vines, Energetics of two-body Hamiltonians in post-Minkowskian gravity, Phys. Rev. D **99**(10), 104004 (2019), doi:10.1103/PhysRevD.99.104004, 1901.07102.

[348] P. Di Vecchia, C. Heissenberg, R. Russo and G. Veneziano, The gravitational eikonal: From particle, string and brane collisions to black-hole encounters, Phys. Rept. **1083**, 1 (2024), doi:10.1016/j.physrep.2024.06.002, 2306.16488.

[349] M. Khalil, A. Buonanno, J. Steinhoff and J. Vines, Energetics and scattering of gravitational two-body systems at fourth post-Minkowskian order, Phys. Rev. D **106**(2), 024042 (2022), doi:10.1103/PhysRevD.106.024042, 2204.05047.

[350] T. Damour and P. Rettegno, Strong-field scattering of two black holes: Numerical relativity meets post-Min Phys. Rev. D **107**(6), 064051 (2023), doi:10.1103/PhysRevD.107.064051, 2211.01399.

[351] P. Rettegno, G. Pratten, L. M. Thomas, P. Schmidt and T. Damour, Strong-field scattering of two spinning black holes: Numerical relativity versus post-Minkowskian gravity Phys. Rev. D **108**(12), 124016 (2023), doi:10.1103/PhysRevD.108.124016, 2307.06999.

[352] A. Buonanno, G. U. Jakobsen and G. Mogull, Post-Minkowskian theory meets the spinning effective-one-bo Phys. Rev. D **110**(4), 044038 (2024), doi:10.1103/PhysRevD.110.044038, 2402.12342.

[353] E. E. Flanagan and S. A. Hughes, Measuring gravitational waves from binary black hole coalescences: 2. T Phys. Rev. D **57**, 4566 (1998), doi:10.1103/PhysRevD.57.4566, gr-qc/9710129.

[354] L. Lindblom, B. J. Owen and D. A. Brown, Model Waveform Accuracy Standards for Gravitational Wave Da Phys. Rev. D **78**, 124020 (2008), doi:10.1103/PhysRevD.78.124020, 0809.3844.

[355] S. T. McWilliams, B. J. Kelly and J. G. Baker, Observing mergers of non-spinning black-hole binaries, Phys. Rev. D **82**, 024014 (2010), doi:10.1103/PhysRevD.82.024014, 1004.0961.

[356] K. Chatziioannou, A. Klein, N. Yunes and N. Cornish, Constructing Gravitational Waves from Generic Spin-Precessing Compact Binary Inspirals, Phys. Rev. D **95**(10), 104004 (2017), doi:10.1103/PhysRevD.95.104004, 1703.03967.

[357] M. Pürrer and C.-J. Haster, Gravitational waveform accuracy requirements for future ground-based detect Phys. Rev. Res. **2**(2), 023151 (2020), doi:10.1103/PhysRevResearch.2.023151, 1912.10055.

[358] S. Ossokine et al., Multipolar Effective-One-Body Waveforms for Precessing Binary Black Holes: Construct Phys. Rev. D **102**(4), 044055 (2020), doi:10.1103/PhysRevD.102.044055, 2004.09442.

[359] A. Toubiana and J. R. Gair, Indistinguishability criterion and estimating the presence of biases (2024), 2401.06845.

[360] A. Toubiana, L. Pompili, A. Buonanno, J. R. Gair and M. L. Katz, Measuring source properties and quasi-normal-mode frequencies of heavy massive black-hole binaries wi (2023), 2307.15086.

[361] L. Barsotti, P. Fritschel, M. Evans and S. Gras, Updated advanced ligo sensitivity design curve, LIGO Document T1800044-v5 (2018).

[362] V. Kapil, L. Reali, R. Cotesta and E. Berti, Systematic bias from waveform modeling for binary black hole po Phys. Rev. D **109**(10), 104043 (2024), doi:10.1103/PhysRevD.109.104043, 2404.00090.

[363] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jiménez Forteza and A. Bohé, Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological m Phys. Rev. D **93**(4), 044007 (2016), doi:10.1103/PhysRevD.93.044007, 1508.07253.

[364] G. Pratten, S. Husa, C. Garcia-Quiros, M. Colleoni, A. Ramos-Buades, H. Estelles and R. Jaume, Setting the cornerstone for a family of models for gravitational waves from compact binaries: Phys. Rev. D **102**(6), 064001 (2020), doi:10.1103/PhysRevD.102.064001, 2001.11412.

[365] D. Ferguson, K. Jani, P. Laguna and D. Shoemaker, Assessing the readiness of numerical relativity for LISA and 3G detectors, Phys. Rev. D **104**(4), 044037 (2021), doi:10.1103/PhysRevD.104.044037, 2006.04272.

[366] A. Jan, D. Ferguson, J. Lange, D. Shoemaker and A. Zimmerman, Accuracy limitations of existing numerical relativity waveforms on the data analysis of current and future (2023), 2312.10241.

[367] R. Takahashi and T. Nakamura, Wave effects in gravitational lensing of gravitational waves from chirping Astrophys. J. **595**, 1039 (2003), doi:10.1086/377430, astro-ph/0305055.

[368] L. Dai and T. Venumadhav, On the waveforms of gravitationally lensed gravitational waves (2017), 1702.04724.

[369] P. Schneider, J. Ehlers and E. Falco, Gravitational lenses, Springer-Verlag (1992), 10.1007/978-1-4612-2756-4.

[370] A. Liu, I. C. F. Wong, S. H. W. Leong, A. More, O. A. Hannuksela and T. G. F. Li, Exploring the hidden Universe: a novel phenomenological approach for recovering arbitrary gravitationa Mon. Not. Roy. Astron. Soc. **525**(3), 4149 (2023), doi:10.1093/mnras/stad1302, 2302.09870.

[371] M. Wright and M. Hendry, Gravelamps: Gravitational wave lensing mass profile model selection, The Astrophysical Journal **935**(2), 68 (2022), doi:10.3847/1538-4357/ac7ec2.

[372] E. Seo, O. A. Hannuksela and T. G. F. Li, Improving Detection of Gravitational-wave Microlensing Using Rep Astrophys. J. **932**(1), 50 (2022), doi:10.3847/1538-4357/ac6dea, 2110.03308.

[373] A. Mishra, A. K. Meena, A. More, S. Bose and J. S. Bagla, Gravitational lensing of gravitational waves: effect of microlens population in lensing galaxies, Mon. Not. Roy. Astron. Soc. **508**(4), 4869 (2021), doi:10.1093/mnras/stab2875, 2102.03946.

[374] A. K. Meena, A. Mishra, A. More, S. Bose and J. S. Bagla, Gravitational lensing of gravitational waves: Probability of microlensing in galaxy-scale lens population, Mon. Not. Roy. Astron. Soc. **517**(1), 872 (2022), doi:10.1093/mnras/stac2721, 2205.05409.

[375] J. M. Ezquiaga, D. E. Holz, W. Hu, M. Lagos and R. M. Wald, Phase effects from strong gravitational lensing of gravitational waves, Phys. Rev. D **103**, 064047 (2021), doi:10.1103/PhysRevD.103.064047.

[376] Y. Wang, R. K. L. Lo, A. K. Y. Li and Y. Chen, Identifying Type II Strongly Lensed Gravitational-Wave Images in Third-Generation Gravitational-Wave D Phys. Rev. D **103**(10), 104055 (2021), doi:10.1103/PhysRevD.103.104055, 2101.08264.

[377] J. Janquart, E. Seo, O. A. Hannuksela, T. G. F. Li and C. V. D. Broeck, On the Identification of Individual Gravitational-wave Image Types of a Lensed System Using Higher-ord Astrophys. J. Lett. **923**(1), L1 (2021), doi:10.3847/2041-8213/ac3bcf, 2110.06873.

[378] A. Vijaykumar, A. K. Mehta and A. Ganguly, Detection and parameter estimation challenges of type-II lense Phys. Rev. D **108**(4), 043036 (2023), doi:10.1103/PhysRevD.108.043036, 2202.06334.

[379] J. M. Ezquiaga, W. Hu, M. Lagos, M.-X. Lin and F. Xu, Modified gravitational wave propagation with higher modes and its degeneracies with lensing, JCAP **08**(08), 016 (2022), doi:10.1088/1475-7516/2022/08/016, 2203.13252.

[380] M. Wright, J. Janquart and N. Johnson-McDaniel, Effect of Deviations from General Relativity on Searches for Gravitational Wave Microlensing and Type II (2024), 2403.08957.

[381] J. Janquart et al., Follow-up analyses to the O3 LIGO–Virgo–KAGRA lensing searches, Mon. Not. Roy. Astron. Soc. **526**(3), 3832 (2023), doi:10.1093/mnras/stad2909, 2306.03827.

[382] A. Mishra, N. V. Krishnendu and A. Ganguly, Unveiling microlensing biases in testing general relativity wit Phys. Rev. D **110**(8), 084009 (2024), doi:10.1103/PhysRevD.110.084009, 2311.08446.

[383] J. M. Ezquiaga, W. Hu and M. Lagos, Apparent Superluminality of Lensed Gravitational Waves, Phys. Rev. D **102**(2), 023531 (2020), doi:10.1103/PhysRevD.102.023531, 2005.10702.

[384] T. Suyama, On arrival time difference between lensed gravitational waves and light, Astrophys. J. **896**(1), 46 (2020), doi:10.3847/1538-4357/ab8d3f, 2003.11748.

[385] A. I. Harte, Gravitational lensing beyond geometric optics: I. Formalism and observables, Gen. Rel. Grav. **51**(1), 14 (2019), doi:10.1007/s10714-018-2494-x, 1808.06203.

[386] G. Cusin and M. Lagos, Gravitational wave propagation beyond geometric optics, Phys. Rev. D **101**(4), 044041 (2020), doi:10.1103/PhysRevD.101.044041, 1910.13326.

[387] C. Yin and J.-h. He, Detectability of Single Spinless Stellar-Mass Black Holes through Gravitational Lensin Phys. Rev. Lett. **132**(1), 011401 (2024), doi:10.1103/PhysRevLett.132.011401, 2312.12451.

[388] E. C. Ostriker, Dynamical friction in a gaseous medium, Astrophys. J. **513**, 252 (1999), doi:10.1086/306858, astro-ph/9810324.

[389] H. Kim and W.-T. Kim, Dynamical Friction of a Circular-Orbit Perturber in a Gaseous Medium, Astrophys. J. **665**, 432 (2007), doi:10.1086/519302, 0705.0084.

[390] H. Kim, W.-T. Kim and F. J. Sanchez-Salcedo, Dynamical Friction of Double Perturbers in a Gaseous Mediu Astrophys. J. Lett. **679**, L33 (2008), doi:10.1086/589149, 0804.2010.

[391] E. Barausse, Relativistic dynamical friction in a collisional fluid, Mon. Not. Roy. Astron. Soc. **382**, 826 (2007), doi:10.1111/j.1365-2966.2007.12408.x, 0709.0211.

[392] A. Loeb, Electromagnetic Counterparts to Black Hole Mergers Detected by LIGO, Astrophys. J. Lett. **819**(2), L21 (2016), doi:10.3847/2041-8205/819/2/L21, 1602.04735.

[393] J. M. Fedrow, C. D. Ott, U. Sperhake, J. Blackman, R. Haas, C. Reisswig and A. De Felice, Gravitational Waves from Binary Black Hole Mergers inside Stars, Phys. Rev. Lett. **119**(17), 171103 (2017), doi:10.1103/PhysRevLett.119.171103, 1704.07383.

[394] D. J. D'Orazio and A. Loeb, Single progenitor model for GW150914 and GW170104, Phys. Rev. D **97**(8), 083008 (2018), doi:10.1103/PhysRevD.97.083008, 1706.04211.

[395] B. McKernan, K. E. S. Ford, W. Lyra and H. B. Perets, Intermediate mass black holes in AGN discs - I. Production and growth, Mon. Not. Roy. Astron. Soc. **425**(1), 460 (2012), doi:10.1111/j.1365-2966.2012.21486.x, 1206.2309.

[396] I. Bartos, B. Kocsis, Z. Haiman and S. Márka, Rapid and Bright Stellar-mass Binary Black Hole Mergers in Active Galactic Nuclei, Astrophys. J. **835**(2), 165 (2017), doi:10.3847/1538-4357/835/2/165, 1602.03831.

[397] N. C. Stone, B. D. Metzger and Z. Haiman, Assisted inspirals of stellar mass black holes embedded in AGN Mon. Not. Roy. Astron. Soc. **464**(1), 946 (2017), doi:10.1093/mnras/stw2260, 1602.04226.

[398] A. Caputo, L. Sberna, A. Toubiana, S. Babak, E. Barausse, S. Marsat and P. Pani, Gravitational-wave detection and parameter estimation for accreting black-hole binaries and their electro Astrophys. J. **892**(2), 90 (2020), doi:10.3847/1538-4357/ab7b66, 2001.03620.

[399] A. Toubiana et al., Detectable environmental effects in GW190521-like black-hole binaries with LISA, Phys. Rev. Lett. **126**(10), 101105 (2021), doi:10.1103/PhysRevLett.126.101105, 2010.06056.

[400] L. Sberna et al., Observing GW190521-like binary black holes and their environment with LISA, Phys. Rev. D **106**(6), 064056 (2022), doi:10.1103/PhysRevD.106.064056, 2205.08550.

[401] I. Mandel and F. S. Broekgaarden, Rates of compact object coalescences, Living Rev. Rel. **25**(1), 1 (2022), doi:10.1007/s41114-021-00034-3, 2107.14239.

[402] K. Eda, Y. Itoh, S. Kuroyanagi and J. Silk, New Probe of Dark-Matter Properties: Gravitational Waves from Phys. Rev. Lett. **110**(22), 221101 (2013), doi:10.1103/PhysRevLett.110.221101, 1301.5971.

[403] A. Coogan, G. Bertone, D. Gaggero, B. J. Kavanagh and D. A. Nichols, Measuring the dark matter environments of black hole binaries with gravitational waves, Phys. Rev. D **105**(4), 043009 (2022), doi:10.1103/PhysRevD.105.043009, 2108.04154.

[404] P. Gondolo and J. Silk, Dark matter annihilation at the galactic center, Phys. Rev. Lett. **83**, 1719 (1999), doi:10.1103/PhysRevLett.83.1719, astro-ph/9906391.

[405] G. Bertone and D. Merritt, Time-dependent models for dark matter at the Galactic Center, Phys. Rev. D **72**, 103502 (2005), doi:10.1103/PhysRevD.72.103502, astro-ph/0501555.

[406] Y. N. Eroshenko, Dark matter density spikes around primordial black holes, Astron. Lett. **42**(6), 347 (2016), doi:10.1134/S1063773716060013, 1607.00612.

[407] S. M. Boucenna, F. Kuhnel, T. Ohlsson and L. Visinelli, Novel Constraints on Mixed Dark-Matter Scenarios of Primordial Black Holes and WIMPs, JCAP **07**, 003 (2018), doi:10.1088/1475-7516/2018/07/003, 1712.06383.

[408] B. Punsly, High-Energy Gamma-Ray Emission from Galactic Kerr-Newman Black Holes. I. The Central Eng Astrophysical Journal **498**(2), 640 (1998), doi:10.1086/305561.

[409] B. Punsly, High-Energy Gamma-Ray Emission from Galactic Kerr-Newman Black Holes. II. The Radiative J Astrophysical Journal **498**(2), 660 (1998), doi:10.1086/305591.

[410] C. Bonvin, C. Caprini, R. Sturani and N. Tamanini, Effect of matter structure on the gravitational waveform, Phys. Rev. D **95**(4), 044029 (2017), doi:10.1103/PhysRevD.95.044029, 1609.08093.

[411] N. Tamanini, A. Klein, C. Bonvin, E. Barausse and C. Caprini, Peculiar acceleration of stellar-origin black hole binaries: Measurement and biases with LISA, Phys. Rev. D **101**(6), 063002 (2020), doi:10.1103/PhysRevD.101.063002, 1907.02018.

[412] K. Inayoshi, N. Tamanini, C. Caprini and Z. Haiman, Probing stellar binary black hole formation in galactic nuclei via the imprint of their center of mass accele

Phys. Rev. D **96**(6), 063014 (2017), doi:10.1103/PhysRevD.96.063014, 1702.06529.

[413] X. Chen, S. Li and Z. Cao, Mass–redshift degeneracy for the gravitational-wave sources in the vicinity of su

Mon. Not. Roy. Astron. Soc. **485**(1), L141 (2019), doi:10.1093/mnrasl/slz046, 1703.10543.

[414] N. Seto, S. Kawamura and T. Nakamura, Possibility of direct measurement of the acceleration of the univer

Phys. Rev. Lett. **87**, 221103 (2001), doi:10.1103/PhysRevLett.87.221103, astro-ph/0108011.

[415] A. Nishizawa, K. Yagi, A. Taruya and T. Tanaka, Cosmology with space-based gravitational-wave detectors — dark energy and primordial gravitational wa

Phys. Rev. D **85**, 044047 (2012), doi:10.1103/PhysRevD.85.044047, 1110.2865.

[416] R. Brito, V. Cardoso and P. Pani, Superradiance: New Frontiers in Black Hole Physics, Lect. Notes Phys. **971**, 1 (2020), doi:10.1007/978-3-030-46622-0, 1501.06570.

[417] A. Arvanitaki, M. Baryakhtar, S. Dimopoulos, S. Dubovsky and R. Lasenby, Black Hole Mergers and the QCD Axion at Advanced LIGO, Phys. Rev. D **95**(4), 043001 (2017), doi:10.1103/PhysRevD.95.043001, 1604.03958.

[418] V. Cardoso, F. Duque and T. Ikeda, Tidal effects and disruption in superradiant clouds: a numerical investig

Phys. Rev. D **101**(6), 064054 (2020), doi:10.1103/PhysRevD.101.064054, 2001.01729.

[419] D. Baumann, G. Bertone, J. Stout and G. M. Tomaselli, Ionization of gravitational atoms, Phys. Rev. D **105**(11), 115036 (2022), doi:10.1103/PhysRevD.105.115036, 2112.14777.

[420] V. De Luca, A. Maselli and P. Pani, Modeling frequency-dependent tidal deformability for environmental bla

Phys. Rev. D **107**(4), 044058 (2023), doi:10.1103/PhysRevD.107.044058, 2212.03343.

[421] R. Brito, S. Ghosh, E. Barausse, E. Berti, V. Cardoso, I. Dvorkin, A. Klein and P. Pani, Stochastic and resolvable gravitational waves from ultralight bosons, Phys. Rev. Lett. **119**(13), 131101 (2017), doi:10.1103/PhysRevLett.119.131101, 1706.05097.

[422] R. Brito, S. Ghosh, E. Barausse, E. Berti, V. Cardoso, I. Dvorkin, A. Klein and P. Pani, Gravitational wave searches for ultralight bosons with LIGO and LISA, Phys. Rev. D **96**(6), 064050 (2017), doi:10.1103/PhysRevD.96.064050, 1706.06311.

[423] C. Palomba et al., Direct constraints on ultra-light boson mass from searches for continuous gravitational

Phys. Rev. Lett. **123**, 171101 (2019), doi:10.1103/PhysRevLett.123.171101, 1909.08854.

[424] L. Sun, R. Brito and M. Isi, Search for ultralight bosons in Cygnus X-1 with Advanced LIGO, Phys. Rev. D **101**(6), 063020 (2020), doi:10.1103/PhysRevD.101.063020, [Erratum: Phys.Rev.D 102, 089902 (2020)], 1909.11267.

[425] V. Dergachev and M. A. Papa, Results from high-frequency all-sky search for continuous gravitational wave Phys. Rev. D **103**(6), 063019 (2021), doi:10.1103/PhysRevD.103.063019, 2012.04232.

[426] R. Abbott et al., All-sky search for gravitational wave emission from scalar boson clouds around spinning Phys. Rev. D **105**(10), 102001 (2022), doi:10.1103/PhysRevD.105.102001, 2111.15507.

[427] A. S. Eddington, The Internal Constitution of the Stars, Cambridge Science Classics. Cambridge University Press, doi:10.1017/CBO9780511600005 (1988).

[428] M. C. Begelman, Nearly collisionless spherical accretion, Monthly Notices of the Royal Astronomical Society **181**(2), 347 (1977), doi:10.1093/mnras/181.2.347, https://academic.oup.com/mnras/article-pdf/181/2/347/9333877/mnras181-0347.pdf.

[429] V. Cardoso and A. Maselli, Constraints on the astrophysical environment of binaries with gravitational-wav Astron. Astrophys. **644**, A147 (2020), doi:10.1051/0004-6361/202037654, 1909.05870.

[430] A. Boudon, P. Brax, P. Valageas and L. K. Wong, Gravitational waves from binary black holes in a self-interacting scalar dark matter cloud, Phys. Rev. D **109**(4), 043504 (2024), doi:10.1103/PhysRevD.109.043504, 2305.18540.

[431] Q. Henry, F. Larrouturou and C. Le Poncin-Lafitte, Electromagnetic fields in compact binaries: A post-Newtonian approach, Phys. Rev. D **108**(2), 024020 (2023), doi:10.1103/PhysRevD.108.024020, 2303.17536.

[432] Q. Henry, F. Larrouturou and C. Le Poncin-Lafitte, Electromagnetic fields in compact binaries: Post-Newtonian wave generation and application to double w Phys. Rev. D **109**(8), 084048 (2024), doi:10.1103/PhysRevD.109.084048, 2310.03785.

[433] A. Vijaykumar, A. Tiwari, S. J. Kapadia, K. G. Arun and P. Ajith, Waltzing Binaries: Probing the Line-of-sight Acceleration of Merging Compact Objects with Gravitational Astrophys. J. **954**(1), 105 (2023), doi:10.3847/1538-4357/acd77d, 2302.09651.

[434] X. Chen and Z.-F. Shen, Retrieving the True Masses of Gravitational-wave Sources, MDPI Proc. **17**(1), 4 (2019), doi:10.3390/proceedings2019017004, 1906.11055.

[435] J. Samsing and T. Ilan, Topology of black hole binary–single interactions, Mon. Not. Roy. Astron. Soc. **476**(2), 1548 (2018), doi:10.1093/mnras/sty197, 1706.04672.

[436] J. Samsing and T. Ilan, Double Gravitational Wave Mergers, Mon. Not. Roy. Astron. Soc. **482**(1), 30 (2019), doi:10.1093/mnras/sty2249, 1709.01660.

[437] M. Tory, E. Grishin and I. Mandel, Empirical stability boundary for hierarchical triples, Publications of the Astronomical Society of Australia **39**, e062 (2022), doi:10.1017/pasa.2022.57, 2208.14005.

[438] Y. Kozai, Secular perturbations of asteroids with high inclination and eccentricity, Astronomical Journal **67**, 591 (1962), doi:10.1086/108790.

[439] M. L. Lidov, The evolution of orbits of artificial satellites of planets under the action of gravitational pertu Planetary and Space Science **9**(10), 719 (1962), doi:10.1016/0032-0633(62)90129-0.

[440] P. Gupta, H. Suzuki, H. Okawa and K.-i. Maeda, Gravitational Waves from Hierarchical Triple Systems with Kozai-Lidov Oscillation, Phys. Rev. D **101**(10), 104053 (2020), doi:10.1103/PhysRevD.101.104053, 1911.11318.

[441] R. S. Chandramouli and N. Yunes, Ready-to-use analytic model for gravitational waves from a hierarchical Phys. Rev. D **105**(6), 064009 (2022), doi:10.1103/PhysRevD.105.064009, 2107.00741.

[442] G. Caneva Santoro, S. Roy, R. Vicente, M. Haney, O. J. Piccinni, W. Del Pozzo and M. Martinez, First Constraints on Compact Binary Environments from LIGO-Virgo Data, Phys. Rev. Lett. **132**(25), 251401 (2024), doi:10.1103/PhysRevLett.132.251401, 2309.05061.

[443] J. L. Jaramillo, R. Panosso Macedo and L. A. Sheikh, Gravitational Wave Signatures of Black Hole Quasinormal Mode Instability, Phys. Rev. Lett. **128**(21), 211102 (2022), doi:10.1103/PhysRevLett.128.211102, 2105.03451.

[444] M. H.-Y. Cheung, K. Destounis, R. P. Macedo, E. Berti and V. Cardoso, Destabilizing the Fundamental Mode of Black Holes: The Elephant and the Flea, Phys. Rev. Lett. **128**(11), 111103 (2022), doi:10.1103/PhysRevLett.128.111103, 2111.05415.

[445] V. Cardoso, S. Kastha and R. Panosso Macedo, Physical significance of the black hole quasinormal mode sp Phys. Rev. D **110**(2), 024016 (2024), doi:10.1103/PhysRevD.110.024016, 2404.01374.

[446] W. Junker and G. Schäfer, Binary systems - higher order gravitational radiation damping and wave emission, Mon. Not. R. Astron. Soc **254**, 146 (1992).

[447] M. Davis, R. Ruffini, W. H. Press and R. H. Price, Gravitational radiation from a particle falling radially into a schwarzschild black hole, Phys. Rev. Lett. **27**, 1466 (1971), doi:10.1103/PhysRevLett.27.1466.

[448] F. J. Zerilli, Gravitational field of a particle falling in a schwarzschild geometry analyzed in tensor harmonics, Phys. Rev. D **2**, 2141 (1970), doi:10.1103/PhysRevD.2.2141.

[449] L. E. Simone, E. Poisson and C. M. Will, Headon collision of compact objects in general relativity: Compari Phys. Rev. D **52**, 4481 (1995), doi:10.1103/PhysRevD.52.4481, gr-qc/9506080.

[450] J. Calderón Bustillo, N. Sanchis-Gual, A. Torres-Forné and J. A. Font, Confusing Head-On Collisions with Precessing Intermediate-Mass Binary Black Hole Mergers, Phys. Rev. Lett. **126**(20), 201101 (2021), doi:10.1103/PhysRevLett.126.201101, 2009.01066.

[451] R. Gamba, M. Breschi, G. Carullo, S. Albanesi, P. Rettegno, S. Bernuzzi and A. Nagar, GW190521 as a dynamical capture of two nonspinning black holes, Nature Astron. **7**(1), 11 (2023), doi:10.1038/s41550-022-01813-w, 2106.05575.

[452] V. Cardoso and P. Pani, Testing the nature of dark compact objects: a status report, Living Rev. Rel. **22**(1), 4 (2019), doi:10.1007/s41114-019-0020-4, 1904.05363.

[453] E. Maggio, P. Pani and G. Raposo, Testing the nature of dark compact objects with gravitational waves (2021), doi:10.1007/978-981-15-4702-7_29-1, 2105.06410.

[454] E. G. M. Ferreira, Ultra-light dark matter, Astron. Astrophys. Rev. **29**(1), 7 (2021), doi:10.1007/s00159-021-00135-6, 2005.03254.

[455] S. L. Liebling and C. Palenzuela, Dynamical boson stars, Living Rev. Relativity **26**, 1 (2023), doi:10.1007/s41114-023-00043-4, 1202.5809.

[456] N. Sennett, T. Hinderer, J. Steinhoff, A. Buonanno and S. Ossokine, Distinguishing Boson Stars from Black Holes and Neutron Stars from Tidal Interactions in Inspiraling Bin Phys. Rev. D **96**(2), 024002 (2017), doi:10.1103/PhysRevD.96.024002, 1704.08651.

[457] P. O. Mazur and E. Mottola, Gravitational vacuum condensate stars, Proc. Natl. Acad. Sci. U.S.A. **101**, 9545 (2004), doi:10.1073/pnas.0402717101, gr-qc/0407075.

[458] E. Mottola, The effective theory of gravity and dynamical vacuum energy, JHEP **11**, 037 (2022), doi:10.1007/JHEP11(2022)037, 2205.04703.

[459] N. Uchikata, S. Yoshida and P. Pani, Tidal deformability and I-Love-Q relations for gravastars with polytrop Phys. Rev. D **94**(6), 064015 (2016), doi:10.1103/PhysRevD.94.064015, 1607.03593.

[460] V. Cardoso, E. Franzin, A. Maselli, P. Pani and G. Raposo, Testing strong-field gravity with tidal Love numbers, Phys. Rev. D **95**(8), 084014 (2017), doi:10.1103/PhysRevD.95.084014, [Addendum: Phys.Rev.D 95, 089901 (2017)], 1701.01116.

[461] S. Chakraborty, E. Maggio, M. Silvestrini and P. Pani, Dynamical tidal Love numbers of Kerr-like compact objects (2023), 2310.06023.

[462] T. Dietrich, M. Ujevic, W. Tichy, S. Bernuzzi and B. Bruegmann, Gravitational waves and mass ejecta from binary neutron star mergers: Effect of the mass-ratio, Phys. Rev. D **95**(2), 024029 (2017), doi:10.1103/PhysRevD.95.024029, 1607.06636.

[463] Y. Asali, P. T. H. Pang, A. Samajdar and C. Van Den Broeck, Probing resonant excitations in exotic compact objects via gravitational waves, Phys. Rev. D **102**(2), 024016 (2020), doi:10.1103/PhysRevD.102.024016, 2004.05128.

[464] N. Loutrel, R. Brito, A. Maselli and P. Pani, Inspiraling compact objects with generic deformations, Phys. Rev. D **105**(12), 124050 (2022), doi:10.1103/PhysRevD.105.124050, 2203.01725.

[465] S. Mukherjee, S. Datta, S. Tiwari, K. S. Phukon and S. Bose, Toward establishing the presence or absence of horizons in coalescing binaries of compact objects by usir Phys. Rev. D **106**(10), 104032 (2022), doi:10.1103/PhysRevD.106.104032, 2202.08661.

[466] M. V. S. Saketh, Z. Zhou, S. Ghosh, J. Steinhoff and D. Chatterjee, Investigating tidal heating in neutron stars via gravitational Raman scattering (2024), 2407.08327.

[467] N. Sanchis-Gual, C. Herdeiro, J. A. Font, E. Radu and F. Di Giovanni, Head-on collisions and orbital mergers of Proca stars, Phys. Rev. D **99**(2), 024017 (2019), doi:10.1103/PhysRevD.99.024017, 1806.07779.

[468] M. Bezares, M. Bošković, S. Liebling, C. Palenzuela, P. Pani and E. Barausse, Gravitational waves and kicks from the merger of unequal mass, highly compact boson stars, Phys. Rev. D **105**(6), 064067 (2022), doi:10.1103/PhysRevD.105.064067, 2201.06113.

[469] N. Siemonsen and W. E. East, Binary boson stars: Merger dynamics and formation of rotating remnant sta Phys. Rev. D **107**(12), 124018 (2023), doi:10.1103/PhysRevD.107.124018, 2302.06627.

[470] T. Evstafyeva, U. Sperhake, I. M. Romero-Shaw and M. Agathos, Gravitational-Wave Data Analysis with High-Precision Numerical Relativity Simulations of Boson Star Me Phys. Rev. Lett. **133**(13), 131401 (2024), doi:10.1103/PhysRevLett.133.131401, 2406.02715.

[471] V. Cardoso, S. Hopper, C. F. B. Macedo, C. Palenzuela and P. Pani, Gravitational-wave signatures of exotic compact objects and of quantum corrections at the horizon scale, Phys. Rev. D **94**(8), 084031 (2016), doi:10.1103/PhysRevD.94.084031, 1608.08637.

[472] J. Westerweck, A. Nielsen, O. Fischer-Birnholtz, M. Cabero, C. Capano, T. Dent, B. Krishnan, G. Meadors and A. H. Nitz, Low significance of evidence for black hole echoes in gravitational wave data, Phys. Rev. D **97**(12), 124037 (2018), doi:10.1103/PhysRevD.97.124037, 1712.09966.

[473] A. B. Nielsen, C. D. Capano, O. Birnholtz and J. Westerweck, Parameter estimation and statistical significance of echoes following black hole signals in the first Advanc Phys. Rev. D **99**(10), 104012 (2019), doi:10.1103/PhysRevD.99.104012, 1811.04904.

[474] R. K. L. Lo, T. G. F. Li and A. J. Weinstein, Template-based Gravitational-Wave Echoes Search Using Bayesian Model Selection, Phys. Rev. D **99**(8), 084052 (2019), doi:10.1103/PhysRevD.99.084052, 1811.07431.

[475] N. Uchikata, H. Nakano, T. Narikawa, N. Sago, H. Tagoshi and T. Tanaka, Searching for black hole echoes from the LIGO-Virgo Catalog GWTC-1, Phys. Rev. D **100**(6), 062006 (2019), doi:10.1103/PhysRevD.100.062006, 1906.00838.

[476] K. W. Tsang, A. Ghosh, A. Samajdar, K. Chatziioannou, S. Mastrogiovanni, M. Agathos and C. Van Den Broeck, A morphology-independent search for gravitational wave echoes in data from the first and second observ Phys. Rev. D **101**(6), 064012 (2020), doi:10.1103/PhysRevD.101.064012, 1906.11168.

[477] A. Miani, C. Lazzaro, G. A. Prodi, S. Tiwari, M. Drago, E. Milotti and G. Vedovato, Constraints on the amplitude of gravitational wave echoes from black hole ringdown using minimal assu Phys. Rev. D **108**(6), 064018 (2023), doi:10.1103/PhysRevD.108.064018, 2302.12158.

[478] J. Abedi, H. Dykaar and N. Afshordi, Echoes from the Abyss: Tentative evidence for Planck-scale structure Phys. Rev. D **96**(8), 082004 (2017), doi:10.1103/PhysRevD.96.082004, 1612.00266.

[479] R. S. Conklin, B. Holdom and J. Ren, Gravitational wave echoes through new windows, Phys. Rev. D **98**(4), 044021 (2018), doi:10.1103/PhysRevD.98.044021, 1712.06517.

[480] J. Abedi and N. Afshordi, Echoes from the Abyss: A highly spinning black hole remnant for the binary neu JCAP **11**, 010 (2019), doi:10.1088/1475-7516/2019/11/010, 1803.10454.

[481] J. Abedi, L. F. Longo Micchi and N. Afshordi, GW190521: Search for echoes due to stimulated Hawking rad Phys. Rev. D **108**(4), 044047 (2023), doi:10.1103/PhysRevD.108.044047, 2201.00047.

[482] E. Barausse, R. Brito, V. Cardoso, I. Dvorkin and P. Pani, The stochastic gravitational-wave background in the absence of horizons, Class. Quant. Grav. **35**(20), 20LT01 (2018), doi:10.1088/1361-6382/aae1de, 1805.08229.

[483] N. K. Johnson-McDaniel, A. Mukherjee, R. Kashyap, P. Ajith, W. Del Pozzo and S. Vitale, Constraining black hole mimickers with gravitational wave observations, Phys. Rev. D **102**, 123010 (2020), doi:10.1103/PhysRevD.102.123010, 1804.08026.

[484] V. Cardoso, M. Kimura, A. Maselli and L. Senatore, Black Holes in an Effective Field Theory Extension of General Relativity, Phys. Rev. Lett. **121**(25), 251105 (2018), doi:10.1103/PhysRevLett.121.251105, 1808.08962.

[485] P. A. Cano, B. Ganchev, D. R. Mayerson and A. Ruipérez, Black hole multipoles in higher-derivative gravity, JHEP **12**, 120 (2022), doi:10.1007/JHEP12(2022)120, 2208.01044.

[486] M. Vaglio, C. Pacilio, A. Maselli and P. Pani, Multipolar structure of rotating boson stars, Phys. Rev. D **105**(12), 124020 (2022), doi:10.1103/PhysRevD.105.124020, 2203.07442.

[487] R. P. Bhatt, S. Chakraborty and S. Bose, Addressing issues in defining the Love numbers for black holes, Phys. Rev. D **108**(8), 084013 (2023), doi:10.1103/PhysRevD.108.084013, 2306.13627.

[488] N. Sennett, R. Brito, A. Buonanno, V. Gorbenko and L. Senatore, Gravitational-Wave Constraints on an Effective Field-Theory Extension of General Relativity, Phys. Rev. D **102**(4), 044056 (2020), doi:10.1103/PhysRevD.102.044056, 1912.09917.

[489] J. R. Gair and C. J. Moore, Quantifying and mitigating bias in inference on gravitational wave source popu Phys. Rev. D **91**(12), 124062 (2015), doi:10.1103/PhysRevD.91.124062, 1504.02767.

[490] C. J. Moore, E. Finch, R. Buscicchio and D. Gerosa, Testing general relativity with gravitational-wave catalogs: the insidious nature of waveform systematics (2021), doi:10.1016/j.isci.2021.102577, 2103.16486.

[491] D. Psaltis, C. Talbot, E. Payne and I. Mandel, Probing the Black Hole Metric. I. Black Hole Shadows and Bir Phys. Rev. D **103**, 104036 (2021), doi:10.1103/PhysRevD.103.104036, 2012.02117.

[492] E. Payne, M. Isi, K. Chatziioannou and W. M. Farr, Fortifying gravitational-wave tests of general relativity against astrophysical assumptions, Phys. Rev. D **108**(12), 124060 (2023), doi:10.1103/PhysRevD.108.124060, 2309.04528.

[493] R. Magee, M. Isi, E. Payne, K. Chatziioannou, W. M. Farr, G. Pratten and S. Vitale, Impact of selection biases on tests of general relativity with gravitational-wave inspirals, Phys. Rev. D **109**(2), 023014 (2024), doi:10.1103/PhysRevD.109.023014, 2311.03656.

[494] S. E. Woosley and A. Heger, The Pair-Instability Mass Gap for Black Holes, Astrophys. J. Lett. **912**(2), L31 (2021), doi:10.3847/2041-8213/abf2c4, 2103.07933.

[495] A. Zimmerman, C.-J. Haster and K. Chatziioannou, On combining information from multiple gravitational wave sources, Phys. Rev. D **99**(12), 124044 (2019), doi:10.1103/PhysRevD.99.124044, 1903.11008.

[496] S. E. Perkins, N. Yunes and E. Berti, Probing Fundamental Physics with Gravitational Waves: The Next Gen Phys. Rev. D **103**(4), 044024 (2021), doi:10.1103/PhysRevD.103.044024, 2010.09010.

[497] M. Isi, K. Chatziioannou and W. M. Farr, Hierarchical test of general relativity with gravitational waves, Phys. Rev. Lett. **123**(12), 121101 (2019), doi:10.1103/PhysRevLett.123.121101, 1904.08011.

[498] M. Isi, W. M. Farr and K. Chatziioannou, Comparing Bayes factors and hierarchical inference for testing ger Phys. Rev. D **106**(2), 024048 (2022), doi:10.1103/PhysRevD.106.024048, 2204.10742.

[499] C. Pacilio, D. Gerosa and S. Bhagwat, Catalog variance of testing general relativity with gravitational-wave Phys. Rev. D **109**(8), L081302 (2024), doi:10.1103/PhysRevD.109.L081302, 2310.03811.

[500] M. Pürrer and C.-J. Haster, Gravitational waveform accuracy requirements for future ground-based detect Physical Review Research **2**(2), 023151 (2020), doi:10.1103/PhysRevResearch.2.023151, 1912.10055.

[501] B. P. Abbott et al., GW190425: Observation of a Compact Binary Coalescence with Total Mass $\sim 3.4 M_{\odot}$, Astrophys. J. Lett. **892**(1), L3 (2020), doi:10.3847/2041-8213/ab75f5, 2001.01761.

[502] G. Pratten, P. Schmidt and N. Williams, Impact of Dynamical Tides on the Reconstruction of the Neutron S Phys. Rev. Lett. **129**(8), 081102 (2022), doi:10.1103/PhysRevLett.129.081102, 2109.07566.

[503] T. Hinderer, B. D. Lackey, R. N. Lang and J. S. Read, Tidal deformability of neutron stars with realistic equations of state and their gravitational wave signatur Phys. Rev. D **81**(12), 123016 (2010), doi:10.1103/PhysRevD.81.123016, 0911.3535.

[504] J. P. Pereira, M. Bejger, J. L. Zdunik and P. Haensel, Differentiating between sharp and smoother phase transitions in neutron stars, Phys. Rev. D **105**(12), 123015 (2022), doi:10.1103/PhysRevD.105.123015, 2201.01217.

[505] R. Abbott et al., GW190412: Observation of a Binary-Black-Hole Coalescence with Asymmetric Masses, Phys. Rev. D **102**(4), 043015 (2020), doi:10.1103/PhysRevD.102.043015, 2004.08342.

[506] R. Abbott et al., GW190814: Gravitational Waves from the Coalescence of a 23 Solar Mass Black Hole wit Astrophys. J. Lett. **896**(2), L44 (2020), doi:10.3847/2041-8213/ab960f, 2006.12611.

[507] S. J. Miller, M. Isi, K. Chatziioannou, V. Varma and I. Mandel, GW190521: Tracing imprints of spin-precession on the most massive black hole binary, Phys. Rev. D **109**(2), 024024 (2024), doi:10.1103/PhysRevD.109.024024, 2310.01544.

[508] R. Macas, A. Lundgren and G. Ashton, Revisiting the evidence for precession in GW200129 with machine l Phys. Rev. D **109**(6), 062006 (2024), doi:10.1103/PhysRevD.109.062006, 2311.09921.

[509] P. D. Lasky, E. Thrane, Y. Levin, J. Blackman and Y. Chen, Detecting gravitational-wave memory with LIGO: implications of GW150914, Phys. Rev. Lett. **117**(6), 061102 (2016), doi:10.1103/PhysRevLett.117.061102, 1605.01415.

[510] M. Hübner, C. Talbot, P. D. Lasky and E. Thrane, Measuring gravitational-wave memory in the first LIGO/Virgo gravitational-wave transient catalog, Phys. Rev. D **101**(2), 023011 (2020), doi:10.1103/PhysRevD.101.023011, 1911.12496.

[511] M. Hübner, P. Lasky and E. Thrane, Memory remains undetected: Updates from the second LIGO/Virgo gra Phys. Rev. D **104**(2), 023004 (2021), doi:10.1103/PhysRevD.104.023004, 2105.02879.

[512] G. P. Smith, A. Robertson, G. Mahler, M. Nicholl, D. Ryczanowski, M. Bianconi, K. Sharon, R. Massey, J. Richard and M. Jauzac, Discovering gravitationally lensed gravitational waves: predicted rates, candidate selection, and localizat Mon. Not. Roy. Astron. Soc. **520**(1), 702 (2023), doi:10.1093/mnras/stad140, 2204.12977.

[513] K. K. Y. Ng, K. W. K. Wong, T. Broadhurst and T. G. F. Li, Precise LIGO Lensing Rate Predictions for Binary Black Holes, Phys. Rev. D **97**(2), 023012 (2018), doi:10.1103/PhysRevD.97.023012, 1703.06319.

[514] F. Xu, J. M. Ezquiaga and D. E. Holz, Please Repeat: Strong Lensing of Gravitational Waves as a Probe of C Astrophys. J. **929**(1), 9 (2022), doi:10.3847/1538-4357/ac58f8, 2105.14390.