



Hearing gravity from the cosmos: GWTC-2 probes general relativity at cosmological scales



Jose María Ezquiaga

Kavli Institute for Cosmological Physics and Enrico Fermi Institute, The University of Chicago, Chicago, IL 60637, USA

ARTICLE INFO

Article history:

Received 22 July 2021

Received in revised form 2 September 2021

Accepted 18 September 2021

Available online 29 September 2021

Editor: M. Trodden

ABSTRACT

Gravitational-wave (GW) catalogs are rapidly increasing in number, allowing for statistical analyses of the population of compact binaries. Nonetheless, GW inference of cosmology has typically relied on additional electromagnetic counterparts or galaxy catalogs. I present a new probe of cosmological modifications of general relativity with GW data only. I focus on deviations of the GW luminosity distance constrained with the astrophysical population of binary black holes (BBHs). The three key observables are 1) the number of events as a function of luminosity distance, 2) the stochastic GW background of unresolved binaries and 3) the location of any feature in the source mass distribution, such as the pair instability supernova (PISN) gap. Modifications of the GW luminosity distance are a priori degenerate with the unknown evolution of the merger rate and source masses. However, a large damping of the GW amplitude predicts a reduction of the events and lowering of the edges of the PISN gap with redshift that is against standard astrophysical expectations. Applying a hierarchical Bayesian analysis to the current LIGO–Virgo catalog (GWTC-2), the strongest constraints to date are placed on deviations from the GW luminosity distance, finding $c_M = -3.2^{+3.4}_{-2.0}$ at 68% C.L., which is ~ 10 times better than multi-messenger GW170817 bounds. These modifications also affect the determination of the BBH masses, which is crucial to accommodate the high-mass binary GW190521 away from the PISN gap. In this analysis it is found that the maximum mass of 99% of the population shifts to lower masses with increased uncertainty, $m_{99\%} = 46.2^{+11.4}_{-9.1} M_\odot$ at 68% C.L. Testing gravity at large scales with the population of BBHs will become increasingly relevant with future catalogs, providing an independent and self-contained test of the standard cosmological model.

© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>). Funded by SCOAP³.

1. Introduction

The first three observing runs of advanced LIGO [1] and Virgo [2] have seen a rapid growth in the number of gravitational wave (GW) detections [3,4] from 3, to 8, to 39, indicating that the field will soon transition to the era of population analysis - where outliers will flag new phenomena, but the core science will arise from statistical analyses of many events. The current catalog of the LIGO–Virgo Collaboration (LVC) is known as GWTC-2 and contains 50 events. Future observing runs will accumulate hundreds (fourth run) to thousands of events (fifth run).

Preparing in advance, population studies are already central to the LVC astrophysical program [5,6]. Among many interesting findings, GWTC-2 has shown support for the theory of pair instability supernova (PISN), which predicts a mass gap in the mass distribution of black holes between $\sim 50 - 120 M_\odot$ [5–7]. In their analysis only $2^{+3.4}_{-1.7}\%$ of binary black holes (BBH) have primary masses

above $45 M_\odot$ [6]. Another key observable is the merger rate history [8], which according to GWTC-2 is probably growing with redshift, but not faster than the star formation rate [6]. If the PISN mass gap is indeed present in nature, next observing runs will constrain also the upper edge [9].

Although present astrophysical uncertainties play a crucial role in the interpretation of GW catalogs, population studies are not limited to modeling the source population. A good example are constraints on the cosmic expansion from the location of the lower and upper edge of the PISN gap [9,10]. In general, mass distribution information allows to probe different background cosmologies [11]. Beyond testing cosmological parameters, I will show that astrophysical population analyses can probe one of the pillars of the standard model of cosmology, namely, the validity of general relativity (GR) at large scales.

Gravity can be tested with GW number counts [12], looking for deviations to the universal signal-to-noise ratio (SNR) distribution [13,14]. Nonetheless, such universal relation is only valid if the merger rate does not evolve with redshift. Therefore, this test only applies to the low-redshift universe, $z \lesssim 0.5$ [14]. However, cos-

E-mail address: ezquiaga@uchicago.edu.

mological modifications of the GW propagation are most relevant at high-redshifts since they accumulate over long travel distances. Some of these theories have been proposed to solve the H_0 tension [15,16]. GW catalogs alone also probe waveform distortions [17,18], GW lensing beyond GR [19] and birefringence [20].

Tests of gravity have seen a proliferation in light of multi-messenger GW astronomy [21]. If a prompt counterpart is detected, the speed of GWs can be constrained [22,23], as beautifully exemplified with GW170817 [24–27]. Moreover, by directly measuring the source redshift one can infer its electromagnetic (EM) luminosity distance d_L^{em} (assuming a cosmology) and test differences w.r.t. the GW luminosity distance d_L^{gw} [28–31]. After GW170817 [32,33], constraints on additional friction terms in the cosmological propagation proportional to c_M were set to $c_M = -9_{-28}^{+21}$ at 68.3% C.L [33], where $c_M = 0$ defines GR.¹ Alternatively, one can use the GW localization volume to statistically infer the redshift with galaxy catalogs [36]. Recent analyses of GWTC-2 find $\Xi_0 = 1.88_{-1.10}^{+3.83}$ at 68.3% C.L for B-band and completeness threshold $P_{\text{th}} = 0.2$ [37], where in this parametrization $d_L^{\text{gw}}/d_L^{\text{em}}(z \gg 1) = \Xi_0$ and thus GR corresponds to $\Xi_0 = 1$. Despite the great promise of multi-messenger tests, their applicability heavily relies on the number of bright multi-messenger mergers and the completeness of galaxy catalogs [37]. In contrast, this proposal only relies on GW data and can be considered as a guaranteed test.

2. BBH population and merger rates

BBHs merge along the history of the universe following a comoving merger rate $\mathcal{R}(z)$. This quantity is highly model dependent and to present day mostly unknown. As a working hypothesis all BBHs will be assumed to be remnants of stars, opposed to a primordial origin [38]. Thus $\mathcal{R}(z)$ should be negligible at high-redshift, before the peak of star formation z_p . To accommodate this astrophysical prior, I follow the parametrization [39]

$$\mathcal{R}(z) = \mathcal{R}_0 C_0 \frac{(1+z)^\alpha}{1 + \left(\frac{1+z}{1+z_p}\right)^{\alpha+\beta}}, \quad (1)$$

which peaks around z_p and has a slope towards z_p and after controlled by α and β respectively. $C_0(z_p, \alpha, \beta) = 1 + (1+z_p)^{-\alpha-\beta}$ sets $\mathcal{R}(0) = \mathcal{R}_0$. All $\{\mathcal{R}_0, z_p, \alpha, \beta\}$ are free parameters to be constrained by observations. In concrete astrophysical scenarios they can be matched, for example, to the peak of star formation and the delay time distribution [39]. Analyses of multiple populations together [40] are left for future work.

To compute the number of detections one needs to include selection effects – how probable is to detect a binary with given intrinsic parameters. Following [6], I take a broken power-law model for the primary mass $p(m_1) \propto m_1^{-\kappa_1}$ with sharp cutoffs at m_{\min} and m_{\max} . The transition between the two slopes κ_1 and κ_2 occurs at a breakpoint m_{break} . For the secondary mass I assume a uniform distribution between m_{\min} and m_1 . Then, the selection effects are encapsulated in the probability of detection p_{det} , which depends on the redshift and masses together with the detector network sensitivity assuming that the inclination angle, sky position and orientation are randomly distributed (see [41] for more details). Altogether, assuming that the mass distribution is constant throughout cosmic time, the detection rate per redshift and component masses is

$$\frac{d^3 \dot{N}_{\text{det}}}{dz dm_1 dm_2} = \frac{\mathcal{R}(z)}{1+z} \frac{dV_c}{dz} p(m_1, m_2) p_{\text{det}}(z, m_1, m_2), \quad (2)$$

¹ Multi-messenger constraints can be tighten if assumed that GW190521 had a counterpart [34], although current observations seem insufficient [35].

where V_c is the comoving volume. Spin priors are uniformly distributed in the aligned component ($-1 \leq s_{1,2,z} \leq 1$) following [6].

Noticeably, if the source population and background cosmology are fixed, any modification in the number of events has to arise from the selection bias p_{det} . Precisely, modifications of gravity will change the SNR affecting the probability of detecting binaries at different redshifts. Note that although some astrophysical formation channels predict some redshift evolution of the source masses [42,43], the cosmological modifications of gravity consider next will have much larger effects.²

3. Probing cosmological modifications of gravity

Assuming that the GW emission and detection follows GR and that there are no additional tensor fields or chirality, beyond GR corrections can be encapsulated in the propagation equation

$$h'' + (2 + \nu)\mathcal{H}h' + (c_g^2 k^2 + \Delta\omega^2)h = 0, \quad (3)$$

for both polarizations $h_{+,\times}$. Three possible modifications can occur: an anomalous propagation speed $c_g \neq c$, a modified dispersion relation $\Delta\omega^2 \neq 0$ and a change in the GW amplitude when $\nu \neq 0$. Relevantly, c_g has been strongly constrained by GW170817 [44] and the modified dispersion relation can be probed directly searching for waveform distortions [18].³ For these reasons I fix $c_g = c$ and $\Delta\omega^2 = 0$ and concentrate on ν which uniquely determines the relation between the GW luminosity distance and the EM luminosity distance [30]:

$$\frac{d_L^{\text{gw}}(z)}{d_L^{\text{em}}(z)} = \exp \left[\frac{1}{2} \int_0^z \frac{\nu(z')}{1+z'} dz' \right]. \quad (4)$$

In GR, $d_L^{\text{gw}}|_{\text{GR}} = d_L^{\text{em}} = (1+z) \int_0^z \frac{c}{H(z)} dz$, where the last equality assumes flat cosmologies. The cosmology is fixed to Planck2018 [46] since the effect of ν dominates over background modifications [28].

Motivated by cosmological modifications of gravity which aim at explaining the present accelerated expansion, I will assume that the additional friction scales with the fractional dark energy density

$$\nu(z) = c_M \frac{\Omega_{\text{DE}}(z)}{\Omega_{\text{DE},0}}, \quad (5)$$

where c_M is a constant and $\Omega_{\text{DE}}(z) = \Omega_{\text{DE},0} H_0^2 / H(z)^2$. This evolution is common in scalar-tensor theories [47–49]. Since the SNR scales inversely with the luminosity distance, $\rho/\rho_{\text{gr}} = d_L^{\text{em}}/d_L^{\text{gw}}$, this leads to

$$\frac{\rho}{\rho_{\text{gr}}} = \exp \left[\frac{-1}{2} \frac{c_M}{\Omega_{\text{DE},0}} \log \left[\frac{1+z}{(\Omega_{\text{M},0}(1+z)^3 + \Omega_{\text{DE},0})^{1/3}} \right] \right], \quad (6)$$

which modifies p_{det} in (2). Other parameterizations are possible [49], being interesting to parametrize directly the ratio of distances $d_L^{\text{gw}}/d_L^{\text{em}} = \Xi_0 + (1 - \Xi_0)/(1+z)^n$ [29], but this (Ξ_0, n) -model is left for future work.

Under these assumptions, the modifications of d_L^{gw} are described in very simple terms. If c_M is positive, d_L^{gw} will be larger

² Note that this is because the model independent modifications considered here are at present largely unconstrained. For concrete modified gravity theories bounded by other data, e.g. by EM surveys, the effect might be equally important or subdominant to the astrophysical effects.

³ A modified dispersion could bias parameter estimation, but given current constraints [18] and number of events [45] it is reasonable to assume that they will not systematically affect the population inference dominantly.

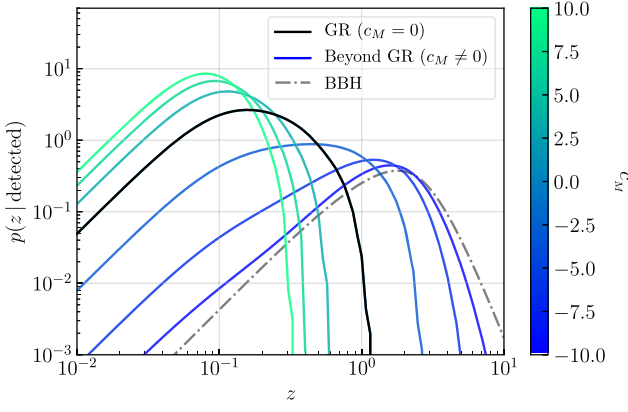


Fig. 1. Observed redshift distribution for a BBH population within GR ($c_M = 0$) and beyond ($c_M \neq 0$) for LIGO/Virgo O3 sensitivity. The merger rate follows Eq. (1) with $\alpha = 1.9$, $\beta = 3.4$ and $z_p = 2.4$; and the mass distribution a broken power law with $m_{\min} = 5M_\odot$, $m_{\max} = 87M_\odot$, $\kappa_1 = 1.6$, $\kappa_2 = 5.6$ and $b = 0.43$. The dashed gray line indicates the sources' redshift distribution $p(z) \propto \mathcal{R}(z)dV_c/dz/(1+z)$.

and the overall signal will be quieter. On the other hand, a negative c_M reduces d_L^{gw} amplifying the GW. Similarly, the higher the redshift, the more important any of these effects become. Although modifications of d_L^{gw} are *a priori* degenerate with arbitrary $\mathcal{R}(z)$ and $p(m_1, m_2)$, e.g. a louder signal could be interpreted as a closer or heavier source, beyond GR effects can lead to signatures that are against standard astrophysical expectations as I discuss next.

3.1. Detection rates

If the cosmological propagation systematically changes the GW amplitude, it is easy to understand that this will affect the number of detections and how far one can hear them. This is explicit in Fig. 1 where the detected redshift distribution is plotted for present sensitivities. If $c_M \gg 1$, then only a fraction of the expected GR events are observed. On the opposite end, if $c_M \ll -1$ one detects events much further, eventually observing the entire population. This plot suggests that the shape of $p(z|\text{detected})$ can constrain modifications of d_L^{gw} . Of course, any meaningful bound has to be placed allowing to vary the other unknown parameters of the model, as it will be done with the Bayesian inference.

Despite the observed redshift distribution being correlated with the merger rate evolution, modification of d_L^{gw} can produce unexpected results by astrophysical priors breaking some of these degeneracies. For example, a decreasing rate of events with redshift before $z \sim 1$ would conflict with BBHs following the star formation rate [50] and with typical delay times below ~ 10 Gyr [51]. This could serve to constrain $c_M \gg 1$. Similarly, a modulation of the number of events with redshift would be astrophysically highly unexpected for a single dominant population [40,51], but possible if GWs mix with other tensor fields introducing an oscillatory pattern in d_L^{gw} [52,53]. The latter nonetheless goes beyond the parametrization (4).

3.2. Stochastic GW background of unresolved binaries

Even though present detectors are only sensitive to relatively low-redshift events, the (non)-observation of the stochastic GW background (SGWB) produced by unresolved binaries provides valuable information. In fact, as shown in Fig. 1, unless $c_M \ll -1$, only a small fraction of all mergers are being detected. The SGWB has the advantage that their sources are at higher redshift and thus more sensitive to modifications of d_L^{gw} .

The energy density of the SGWB can be computed summing over the energy flux emitted by all non-detected events deter-

mined by $1 - p_{\text{det}}$. The dimensionless energy density Ω_{gw} scaling, including the inspiral phase only, is

$$\Omega_{\text{gw}}(f) \sim f^{2/3} \iint \mathcal{M}_c^{5/6} \frac{\mathcal{R}(z)}{H(z)(1+z)^{1/3}} \left(\frac{d_L^{\text{em}}}{d_L^{\text{gw}}} \right)^2 \times p(\mathcal{M}_c)(1 - p_{\text{det}}) d\mathcal{M}_c dz, \quad (7)$$

where the detailed derivation is deferred to the supplemental material since it follows closely the classical result [54]. Interestingly, the luminosity distance ratio appears quadratically in Ω_{gw} . Although not included here, Ω_{gw} could constrain modifications of the GW emission [55–57].

Deviations in d_L^{gw} do not alter the typical $f^{2/3}$ spectral shape of Ω_{gw} . However, they shift the turnaround point of the spectrum at $f > 100$ Hz. A positive c_M moves the maximum to higher frequencies because the quieter sources w.r.t. GR behave as lighter ones reducing the effective minimum mass of the population. The peak of Ω_{gw} is, unfortunately, beyond ground-based detector sensitivities preventing the detection of this possible signature of modified gravity.

3.3. Source mass distribution

Modifying d_L^{gw} will bias the inferred source masses. This is particularly relevant when the distribution of masses $p(m_1, m_2)$ has a distinct mass scale, since this will break the degeneracy with the modified d_L^{gw} . For BBHs, PISN theory sets two reference scales: the edges of the gap. A d_L^{gw} beyond GR will change the inferred location of the gap as exemplified in Fig. 2. Negative c_M moves the PISN gap to higher values and vice-versa. This is because $c_M < 0$ allows to expand the horizon redshift and apparently massive events could be just at higher redshift. In particular, the primary, source mass posterior of GW190521, the most massive event so far [58], shifts according to the sign of c_M as displayed in the right side of Fig. 2. If $c_M < 0$ is assumed, GW190521's inferred mass moves below the gap and vice-versa.

These results are complementary to [59] where local modifications of gravity changing the PISN mass gap in the source population were studied. Here the modified propagation changes the inferred mass gap location, but not $p(m_1, m_2)$. In addition, this idea could be extended to other source populations. For instance, binary neutron stars (BNSs) have a narrow mass distribution that could constrain d_L^{gw} when tidal effects identify the compact object as a neutron star [60]. Similarly, if the BNS merger rate was measured with EM observations, next-generation GW detectors could tightly constrain c_M since $\mathcal{R}(z)$ would be fixed [61]. Note that the location of the PISN mass gap can also shift up to $\mathcal{O}(10M_\odot)$ due to unknown astrophysics such as metallicity, stellar winds or nuclear reaction rates [62].

Throughout this analysis the mass distribution is assumed to be constant in time, but modifications of d_L^{gw} would be roughly equivalent to change the source masses at different redshifts by $\tilde{m}_i(z) \approx (d_L^{\text{em}}/d_L^{\text{gw}})^{6/5} m_i$. Thus, a $c_M > 0$ emulates decreasing the maximum mass with redshift. This contradicts astrophysical expectations of m_{\max} increasing with redshift due to the decrease in metallicity [63]. Therefore, the evolution of PISN gap's edges could be a key determinant to test modifications of gravity. Recently, [64] have shown evidence of an increase of m_{\max} for a power-law model.

4. Constraints from GWTC-2

To test cosmological modifications of gravity with current data, I develop a hierarchical Bayesian pipeline. Since this statistical framework is widely used in the GW community, details are presented in the supplemental material. The key differences with the standard analysis are:

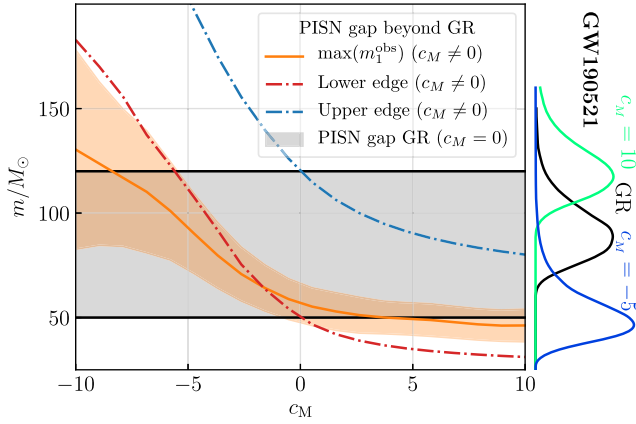


Fig. 2. Impact of the modified luminosity distance on the inference of the PISN mass gap, fixed at $50 - 120 M_{\odot}$. In the main panel the solid orange line represents the maximum inferred source-frame mass with 39 detections averaged over 100 mock simulations including selection bias for O3 sensitivity. The orange band indicates the 1σ dispersion. Dashed-dotted lines correspond to the displacement of the edges of the mass gap computed from the horizon distance in modified gravity. On the right, the posteriors for the primary mass of GW190521 [58] are presented for different values of c_M .

- i) source masses and redshifts inferred values depend on c_M ; $m_i(m_{iz}, d_L^{gw}, c_M)$ and $z(d_L^{gw}, c_M)$,
- ii) the probability of detection is also a function of c_M ; $p_{\text{det}}(z, m_1, m_2, c_M)$.

With these considerations at hand, I analyze GWTC-2, using the same detection threshold as the LVC [6]. The parameters are: c_M for modified gravity, $\{\mathcal{R}_0, \alpha, \beta, z_p\}$ for the merger rate history and $\{\kappa_1, \kappa_2, m_{\text{max}}, b\}$ for the broken power-law mass distribution. m_{min} is fixed to $5 M_{\odot}$.

The main results are summarized in Fig. 3 where the posterior distribution for c_M , α and $m_{99\%}$ are presented. $m_{99\%}$ corresponds to the maximum mass of 99% of the events derived from the posterior mass distribution. There are several important results. First, the modification of gravity can be tightly constrained with BBH data only to $c_M = -3.2^{+3.4}_{-2.0}$ at 68% C.L. This is ~ 10 times better than current multi-messenger constraints [33]. Second, the merger rate slope is still likely positive ($\alpha > 0$ at 65% probability) but its uncertainty increases w.r.t to the LVC results. This is due to the degeneracy between c_M and α . Finally, $m_{99\%}$ shifts to smaller masses with larger errors, $m_{99\%} = 46.2^{+20.4}_{-13.0} M_{\odot}$ at 90% C.L., when compared to the LVC uniform-in-comoving-volume ($\alpha = \beta = 0$) results ($m_{99\%} = 57.8^{+12.5}_{-8.7}$ at 90% C.L. [6]). Therefore, allowing for a modified GW propagation (without dispersion effects) makes GWTC-2 lean towards the theory of PISN.

In addition, it is found that the current non-detection of the SGWB does not impose stronger constraints on c_M than individual events. Moreover, GWTC-2 has not enough high-redshift sources to constrain β and z_p in the parametrization of $\mathcal{R}(z)$. Constraints on the other mass distribution parameters do not change significantly w.r.t. LVC results [6]. For completeness, full posterior samples are presented in the supplemental material. It has been verified that the inferred model is consistent with the observed data performing a posterior predictive check analogous to [6].

Although the bounds on c_M are subject to BBH population modeling, the parametrization chosen is flexible enough so that, under the astrophysical origin assumption, these results are robust. In other words, for other parameterizations consistent with the data, the limits on c_M will be of the same order of magnitude. The fact that there is a preference for $c_M < 0$ is consistent with astrophysical expectations that the PISN gap might increase with redshift [64], thus requiring additional information to disentangle if this effect is due to unmodeled astrophysics or new physics.

Bounds on $c_M > 0$ are driven in part by the PISN-motivated prior on $m_{\text{max}} \leq 100 M_{\odot}$. An exploration of different population models/priors will be addressed in the future.

GW data has the unique potential to constrain deviations from GR that predict $d_L^{gw} \neq d_L^{\text{em}}$, but cosmological modifications of gravity does not restrict to GW data. For scalar-tensor theories, the GR modifications originating the friction term ν in (3) will affect other cosmological observables. Present CMB/LSS bounds [65] are comparable to the constraints on this letter, although future surveys expect to improve them by several orders of magnitude [48]. Despite EM surveys possibly having more constraining power than GW catalogs for some specific modified gravity theories, a multi-probe analysis will still be interesting since the GW data can break degeneracies in the parameter space. For example, it has been shown that LISA standard sirens could complement LSS bounds [66].

5. Future prospects

GW observations contain a wealth of information about our universe. In this letter I have proposed a new probe of gravity at cosmological scales using the population of BBHs. This test requires GW data only and is thus a guaranteed output of any present or future GW catalog.

Applying a hierarchical Bayesian analysis to GWTC-2, I find that, under the assumption of astrophysical origin, BBH observations constrain modifications in the GW luminosity distance more strongly than multi-messenger observations from GW170817, with overall results being consistent with GR. This is because changes in d_L^{gw} alter the inferred redshift and source mass distributions, shifting e.g. the expected PISN mass gap. The GW amplitude damping with redshift ($c_M > 0$) is particularly falsifiable since it leads to rates and PISN mass gap edges that decrease with redshift, which is against standard astrophysical predictions.

This analysis can be extended to incorporate other parameterizations of d_L^{gw} , being particularly interesting to test GW oscillations modulating the observed redshift distribution [52,53]. Similarly, including waveform distortions due to a modified dispersion relation could provide additional constraints beyond GR theories [67]. Although the background cosmology has been fixed throughout the analysis, BBHs observations can also be used to constrain H_0 and Ω_m [11]. A background and perturbation analysis would in principle be possible, but probably only for the scope of next-generation detectors.

This d_L^{gw} test can also be applied to other BBH populations, as those LISA will hear from space [68]. From all LISA sources, the ones at high- z such as extreme-mass-ratio inspirals and super-massive black holes would be more interesting. Lensing effects could be incorporated in a similar way, with the probability of detection modified by the optical depth.

BBH observations have proven to be a powerful test of gravity at cosmological scales. Future GW observations will only improve our understanding of the cosmological model, providing a complementary and independent probe to EM surveys.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

I am grateful to the past and present University of Chicago LIGO group (Reed Essick, Amanda Farah, Maya Fishbach, Daniel Holz

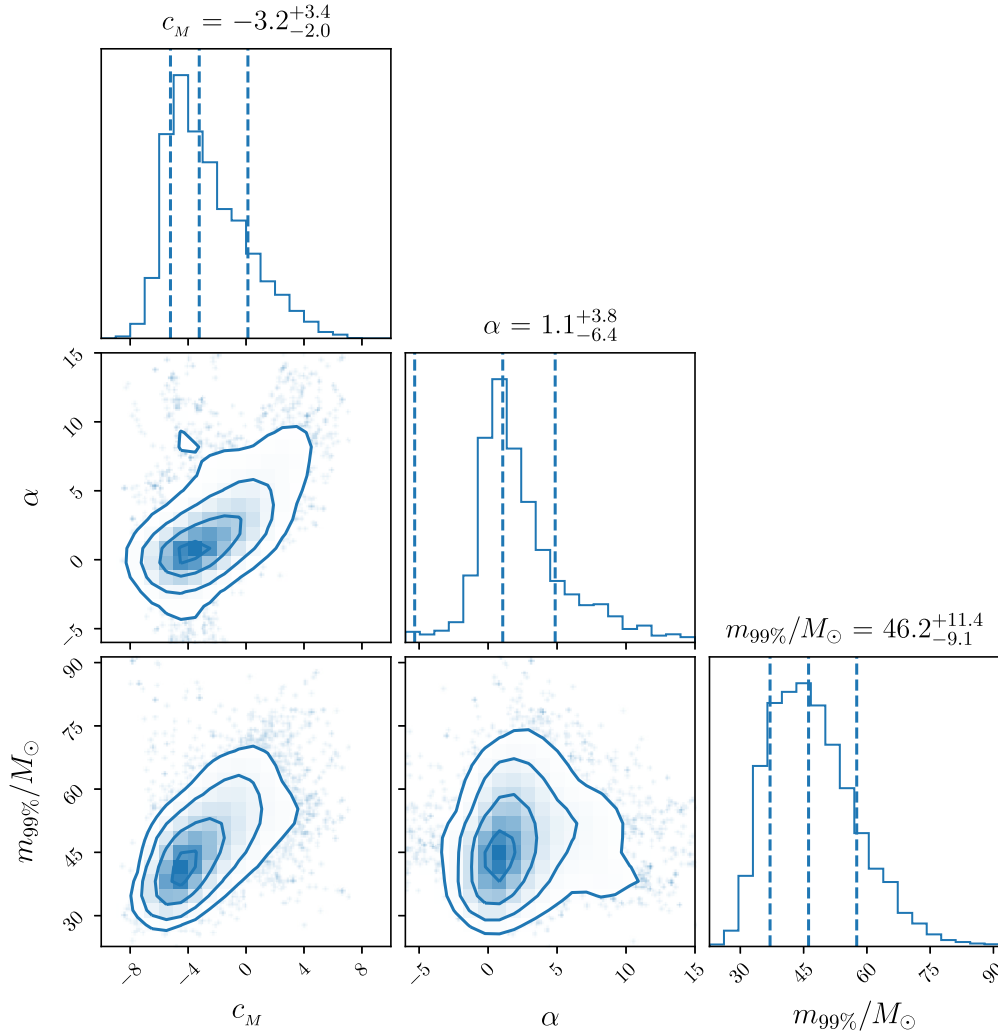


Fig. 3. Posterior distributions for the modification of the GW luminosity distance c_M , the slope of the BBH merger rate α and the maximum mass of 99% of events $m_{99\%}$ inferred from GWTC-2. Vertical lines correspond to means and 68% confidence intervals. Constraints on c_M are ~ 10 times better than current multi-messenger bounds from GW170817 [33]. Both α and $m_{99\%}$ are correlated with c_M , with negative values of c_M leaning towards the predictions of the theory of PISN.

and Mike Zevin) for insightful conversations about GW population analyses, as well as the LIGO–Virgo R&P and Cosmo groups for feedback on the results. I also acknowledge feedback on the manuscript from Maya Fishbach, Max Isi, Macarena Lagos, Simone Mastrogiovanni, Suvodip Mukherjee and Miguel Zumalacárregui. This analysis has used emcee [69] for the MCMC and corner [70] to present the posteriors. I am supported by NASA through the NASA Hubble Fellowship grant HST-HF2-51435.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. I am also supported by the Kavli Institute for Cosmological Physics through an endowment from the Kavli Foundation and its founder Fred Kavli. This research has made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center (<https://www.gw-openscience.org/>), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration.

Appendix A. Statistical analysis

In this supplemental material I summarize the hierarchical Bayesian pipeline developed in this analysis (see e.g. [71] for a general discussion of this statistical framework). The first step is, of

course, Bayes theorem. The posterior distribution of a given set of parameters Λ describing a given population of BBHs follows from

$$p(\Lambda|\{d_i\}) \propto p(\{d_i|\Lambda\})\pi(\Lambda), \quad (\text{A.1})$$

where $p(\{d_i|\Lambda\})$ is the likelihood of obtaining N_{obs} GW events with data $\{d_i\}$, while $\pi(\Lambda)$ are the prior expectations on Λ . Information about the stochastic background can also be included by the likelihood product

$$p(\{d_i\}, \Omega_{\text{gw}}|\Lambda) = p_{\text{bbh}}(\{d_i|\Lambda\}) \times p_{\text{sgwb}}(\Omega_{\text{gw}}|\Lambda), \quad (\text{A.2})$$

setting for example that the SNR of Ω_{gw} should be less than 2 during O3a. However this will not be included in the final results since it is found that it does not constrain more than individual events.

The likelihood of resolvable events can be described by a Poissonian process

$$p_{\text{bbh}}(\{d_i|\Lambda) \propto N_{\text{det}}(\Lambda)^{N_{\text{obs}}} e^{-N_{\text{det}}(\Lambda)} \times \prod_{i=1}^{N_{\text{obs}}} \frac{1}{\xi(\Lambda)} \left\langle \frac{p(\phi_i|\Lambda)}{\pi_{\text{pe}}(\phi_i)} \right\rangle_{\text{samples}}, \quad (\text{A.3})$$

where $\xi = N_{\text{det}}/N_{\text{bbh}}$ is the ratio between the expected detected mergers N_{det} and the actual merger N_{bbh} . Note that the data like-

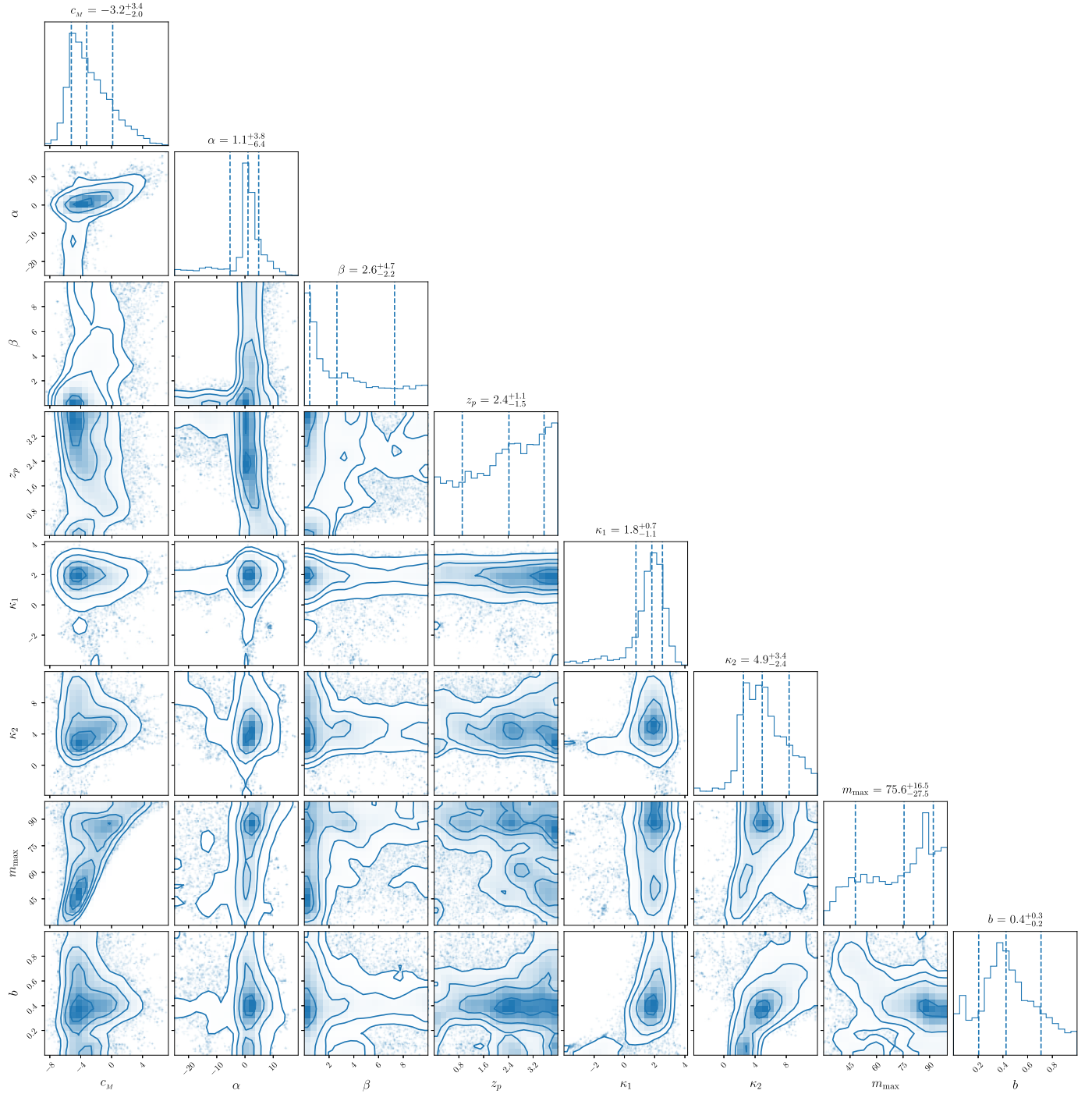


Fig. 4. Posterior distributions from the analysis of modifications of the GW luminosity distance in the BBH population from GWTC-2. The parameters of the analysis are c_M for the modification of gravity, $\{\alpha, \beta, z_p\}$ for the merger rate history and $\{\kappa_1, \kappa_2, m_{\max}, b\}$ for the broken power-law mass distribution. The local merger rate \mathcal{R}_0 has been marginalized using a uniform in log prior.

likelihood, $p(d_i|\phi)$, given the GW parameters ϕ , is not directly accessible. Instead there are only the event posteriors samples $p(\phi|d_i)$ to which it is necessary to factor out the prior used in the parameter estimation $\pi_{\text{pe}}(\phi)$.

The main observables are the inferred redshifts and source masses, thus $\phi = \{z, m_1, m_2\}$. Remember that, as explained in the main text, these three quantities depend on c_M and are derived from the observed data of $\{m_{1z}, m_{2z}, d_L^{\text{gw}}\}$. For these parameters the only relevant parameter estimation prior is $\pi(d_L^{\text{gw}}) \propto (d_L^{\text{gw}})^2$

since for the masses the LVC uses a uniform prior. The prior $\pi_{\text{pe}}(z, m_1, m_2)$ is directly obtained including the Jacobian [5]:

$$\pi_{\text{pe}}(z, m_1, m_2) \propto (d_L^{\text{gw}})^2 (1+z)^2 \frac{\partial d_L^{\text{gw}}}{\partial z}, \quad (\text{A.4})$$

where in this case

$$\frac{\partial d_L^{\text{gw}}}{\partial z} = \frac{d_L^{\text{gw}}}{1+z} + \frac{(1+z)c}{H(z)} \frac{d_L^{\text{gw}}}{d_L^{\text{em}}} + \frac{v}{2(1+z)} d_L^{\text{gw}}, \quad (\text{A.5})$$

following Eq. (4) in the main text. Altogether, the BBH likelihood can be written as

$$p_{\text{bbh}}(\{d_i\}|\Lambda) \propto e^{-N_{\text{det}}(\Lambda)} \times \prod_{i=1}^{N_{\text{obs}}} \left\langle \frac{dN(\phi|\Lambda)/d\phi}{\pi_{\text{pe}}(\phi)} \right\rangle_{\text{samples}}. \quad (\text{A.6})$$

This result can be simplified further if the local merger rate \mathcal{R}_0 is marginalized using a uniform in log prior to obtain

$$p_{\text{bbh}}(\{d_i\}|\Lambda) \propto \xi^{-N_{\text{obs}}} \times \prod_{i=1}^{N_{\text{obs}}} \left\langle \frac{p(\phi|\Lambda)}{\pi_{\text{pe}}(\phi)} \right\rangle_{\text{samples}}, \quad (\text{A.7})$$

which does not depend on \mathcal{R}_0 .

The BBH population is modeled with a merger rate history following Eq. (1) in the main text. For the primary mass a broken power-law distribution is used:

$$p(m_1) \propto \begin{cases} m_1^{\kappa_1}, & m_{\min} < m_1 < m_{\text{break}} \\ m_1^{\kappa_2}, & m_{\text{break}} < m_1 < m_{\max} \\ 0, & \text{elsewhere} \end{cases}, \quad (\text{A.8})$$

where $m_{\text{break}} = m_{\min} + b(m_{\max} - m_{\min})$ and $b \in (0, 1]$. In the limit of $b \rightarrow 1$ one finds $m_{\text{break}} \rightarrow m_{\max}$. On the other hand, the secondary source mass is uniformly sampled below m_1 and above m_{\min} . In the analysis the minimum mass is fixed to $5M_{\odot}$. Therefore, in total the BBH population is modeled by 8 parameters: $\Lambda_{\text{bbh}} = \{\mathcal{R}_0, \alpha, \beta, z_p, \kappa_1, \kappa_2, m_{\max}, b\}$. The modification of gravity is modeled with 1 parameter: $\Lambda_{\text{gravity}} = \{c_M\}$. The priors are chosen to be uniform distributions in the ranges: $\log_{10} \mathcal{R}_0 \in [-3, 3]$, $\alpha \in [-25, 25]$, $\beta \in [0, 10]$, $z_p \in [0, 4]$, $\kappa_{1,2} \in [-4, 12]$, $m_{\max} \in [30, 100]$, $b \in [0, 1]$ and $c_M \in [-12, 12]$. The probability of detection during O3a is computed using the public sensitivity of matched filter searches.⁴

Appendix B. Stochastic background of GWs with modified propagation

Next in the supplemental material I provide a derivation of how the modified GW propagation affects the stochastic background of unresolved binaries. I focus in particular on modification in the GW luminosity distance. This derivation extends the classical result of [54] beyond GR, allowing for $d_L^{\text{gw}} \neq d_L^{\text{em}}$.

The dimensionless stochastic GW background is defined as

$$\Omega_{\text{gw}}(f) = \frac{1}{\rho_c} \frac{d\rho_{\text{gw}}}{d \ln f} = \frac{f}{c\rho_c} F(f), \quad (\text{B.1})$$

where $\rho_c = 3c^2 H_0^2 / 8\pi G$ is the critical energy density and frequencies are in the detector frame. In the second equality, the total energy flux $F(f) = cd\rho_{\text{gw}}/df$ is introduced. The total flux is nothing but the energy emitted by all binaries per unit area:

$$F(f) = c\dot{N} \frac{dE_{\text{gw}}(f)}{df} \frac{(1+z)^2}{4\pi(d_L^{\text{gw}})^2}. \quad (\text{B.2})$$

The number of events per detector frame time has already been defined in Eq. (2) of the main text. There, in order to account for all the binaries which cannot be detected individually one simply needs to substitute $p_{\text{det}} \rightarrow (1 - p_{\text{det}})$. One obtains

$$\frac{d^3 \dot{N}_{\text{unresolv}}}{dz dm_1 dm_2} = \frac{\mathcal{R}(z)}{(1+z)} \frac{dV_c}{dz} p(m_1, m_2)(1 - p_{\text{det}}). \quad (\text{B.3})$$

The energy emitted per frequency is given by (recall only modification in the GW propagation is being considered)

$$\frac{dE_{\text{gw}}}{df} = (1+z) \frac{(G\pi)^{2/3}}{3} \mathcal{M}_c^{5/3} f_s^{-1/3} \quad (\text{B.4})$$

during the inspiral of a circular binary. A more general expression can be obtained simply noting that $dE_{\text{gw}}/df \sim d_L^2 f^2 \langle |\tilde{h}(f)|^2 \rangle_{\vec{\Omega}}$ where $\tilde{h}(f)$ is the Fourier transform of the time domain strain (during the inspiral $\tilde{h}(f) \sim f^{-7/3}$) which has been averaged over all possible sky locations and orientations $\vec{\Omega}$. Noticeably, the number of events scales with the differential comoving volume

$$\frac{dV_c}{dz} = \frac{4\pi r(z)^2}{H(z)} = \frac{4\pi (d_L^{\text{em}})^2}{(1+z)^2 H(z)}. \quad (\text{B.5})$$

Importantly, the comoving rate scales with $(d_L^{\text{em}})^2$ while the GW energy flux scales with $1/(d_L^{\text{gw}})^2$. In GR these two quantities are equal and cancel each other. However, beyond GR they do not and Ω_{gw} depends on their ratio square $(d_L^{\text{em}}/d_L^{\text{gw}})^2$. The final result for the inspiral signal is given in Eq. (7) of the main text. If one wants to include the full emitted signal this can be generalized to

$$\Omega_{\text{gw}}(f) = \frac{4\pi^2 f^3}{3H_0^2} \int |\tilde{h}_{\text{gr}}(f, \vec{\phi})|^2 \frac{\mathcal{R}(z)}{(1+z)} \frac{dV_c}{dz} \left(\frac{d_L^{\text{em}}}{d_L^{\text{gw}}} \right)^2 \times p(m_1, m_2)(1 - p_{\text{det}}(\vec{\phi})) d\vec{\phi}, \quad (\text{B.6})$$

where $\vec{\phi} = \{m_1, m_2, z\}$ and $d\vec{\phi} = dz dm_1 dm_2$. Here \tilde{h}_{gr} is the GR emitted signal (which is inversely proportional to d_L^{em} and thus this distance factor cancels with the one from dV_c/dz).

Appendix C. Full posterior samples

For completeness I present the full posterior distributions for all the parameters in the analysis. The results are displayed in Fig. 4. It is to be noted that in Fig. 3 of the main text the range in α was cut below -6 . This is because as shown in this figure, the parametrization used in this analysis saturates at $\alpha \lesssim -5$ and the inference is the same. In any case, 85% of the posterior is above this value.

As it can be seen in Fig. 4, several of the model parameters are essentially unconstrained with present data. This is the case for example of the peak of the merger rate z_p and its decreasing slope β . This is because current detectors are not sensitive to high redshift where these parameters become relevant.

References

- [1] J. Aasi, et al., LIGO Scientific, Class. Quantum Gravity 32 (2015) 074001, arXiv:1411.4547.
- [2] F. Acernese, et al., VIRGO, Class. Quantum Gravity 32 (2015) 024001, arXiv:1408.3978.
- [3] B.P. Abbott, et al., LIGO Scientific, Virgo, Phys. Rev. X 9 (2019) 031040, arXiv:1811.12907.
- [4] R. Abbott, et al., LIGO Scientific, Virgo, arXiv:2010.14527, 2020.
- [5] B.P. Abbott, et al., LIGO Scientific, Virgo, Astrophys. J. Lett. 882 (2019) L24, arXiv:1811.12940.
- [6] R. Abbott, et al., LIGO Scientific, Virgo, arXiv:2010.14533, 2020.
- [7] M. Fishbach, D.E. Holz, Astrophys. J. Lett. 851 (2017) L25, arXiv:1709.08584.
- [8] M. Fishbach, D.E. Holz, W.M. Farr, Astrophys. J. Lett. 863 (2018) L41, arXiv:1805.10270.
- [9] J.M. Ezquiaga, D.E. Holz, Astrophys. J. Lett. 909 (2021) L23, arXiv:2006.02211.
- [10] W.M. Farr, M. Fishbach, J. Ye, D. Holz, Astrophys. J. Lett. 883 (2019) L42, arXiv:1908.09084.
- [11] S. Mastrogiovanni, K. Leyde, C. Karathanasis, E. Chassande-Mottin, D.A. Steer, J. Gair, A. Ghosh, R. Gray, S. Mukherjee, S. Rinaldi, arXiv:2103.14663, 2021.
- [12] E. Calabrese, N. Battaglia, D.N. Spergel, Class. Quantum Gravity 33 (2016) 165004, arXiv:1602.03883.
- [13] B.F. Schutz, Class. Quantum Gravity 28 (2011) 125023, arXiv:1102.5421.
- [14] H.-Y. Chen, D.E. Holz, arXiv:1409.0522, 2014.
- [15] M. Zumalacargui, Phys. Rev. D 102 (2020) 023523, arXiv:2003.06396.
- [16] T. Abadi, E.D. Kovetz, Phys. Rev. D 103 (2021) 023530, arXiv:2011.13853.
- [17] M. Isi, K. Chatziioannou, W.M. Farr, Phys. Rev. Lett. 123 (2019) 121101, arXiv:1904.08011.
- [18] R. Abbott, et al., LIGO Scientific, Virgo, arXiv:2010.14529, 2020.

⁴ <https://dcc.ligo.org/LIGO-P2000217/public>.

- [19] J.M. Ezquiaga, M. Zumalacárregui, *Phys. Rev. D* 102 (2020) 124048, arXiv:2009.12187.
- [20] M. Okounkova, W.M. Farr, M. Isi, L.C. Stein, arXiv:2101.11153, 2021.
- [21] J.M. Ezquiaga, M. Zumalacárregui, *Front. Astron. Space Sci.* 5 (2018) 44, arXiv:1807.09241.
- [22] L. Lombriser, A. Taylor, *J. Cosmol. Astropart. Phys.* 1603 (2016) 031, arXiv:1509.08458.
- [23] D. Bettoni, J.M. Ezquiaga, K. Hinterbichler, M. Zumalacárregui, *Phys. Rev. D* 95 (2017) 084029, arXiv:1608.01982.
- [24] J.M. Ezquiaga, M. Zumalacárregui, *Phys. Rev. Lett.* 119 (2017) 251304, arXiv:1710.05901.
- [25] P. Creminelli, F. Vernizzi, *Phys. Rev. Lett.* 119 (2017) 251302, arXiv:1710.05877.
- [26] T. Baker, E. Bellini, P.G. Ferreira, M. Lagos, J. Noller, I. Sawicki, *Phys. Rev. Lett.* 119 (2017) 251301, arXiv:1710.06394.
- [27] J. Sakstein, B. Jain, *Phys. Rev. Lett.* 119 (2017) 251303, arXiv:1710.05893.
- [28] E. Belgacem, Y. Dirian, S. Foffa, M. Maggiore, *Phys. Rev. D* 97 (2018) 104066, arXiv:1712.08108.
- [29] E. Belgacem, Y. Dirian, S. Foffa, M. Maggiore, *Phys. Rev. D* 98 (2018) 023510, arXiv:1805.08731.
- [30] E. Belgacem, et al., LISA Cosmology Working Group, *J. Cosmol. Astropart. Phys.* 07 (2019) 024, arXiv:1906.01593.
- [31] S. Mukherjee, B.D. Wandelt, J. Silk, *Mon. Not. R. Astron. Soc.* 502 (2021) 1136, arXiv:2012.15316.
- [32] S. Arai, A. Nishizawa, *Phys. Rev. D* 97 (2018) 104038, arXiv:1711.03776.
- [33] M. Lagos, M. Fishbach, P. Landry, D.E. Holz, *Phys. Rev. D* 99 (2019) 083504, arXiv:1901.03321.
- [34] S. Mastrogiovanni, L. Haegel, C. Karathanasis, I. Magana-Hernandez, D.A. Steer, arXiv:2010.04047, 2020.
- [35] G. Ashton, K. Ackley, I.M.n. Hernandez, B. Piotrkowski, arXiv:2009.12346, 2020.
- [36] W. Del Pozzo, *Phys. Rev. D* 86 (2012) 043011, arXiv:1108.1317.
- [37] A. Finke, S. Foffa, F. Iacovelli, M. Maggiore, M. Mancarella, arXiv:2101.12660, 2021.
- [38] M. Sasaki, T. Suyama, T. Tanaka, S. Yokoyama, *Class. Quantum Gravity* 35 (2018) 063001, arXiv:1801.05235.
- [39] T. Callister, M. Fishbach, D. Holz, W. Farr, *Astrophys. J. Lett.* 896 (2020) L32, arXiv:2003.12152.
- [40] K.K.Y. Ng, S. Vitale, W.M. Farr, C.L. Rodriguez, arXiv:2012.09876, 2020.
- [41] H.-Y. Chen, D.E. Holz, J. Miller, M. Evans, S. Vitale, J. Creighton, *Class. Quantum Gravity* 38 (2021) 055010, arXiv:1709.08079.
- [42] M. Mapelli, N. Giacobbo, F. Santoliquido, M.C. Artale, *Mon. Not. R. Astron. Soc.* 487 (2019) 2, arXiv:1902.01419.
- [43] C.L. Rodriguez, M. Zevin, P. Amaro-Seoane, S. Chatterjee, K. Kremer, F.A. Rasio, C.S. Ye, *Phys. Rev. D* 100 (2019) 043027, arXiv:1906.10260.
- [44] B.P. Abbott, et al., INTEGRAL LIGO Scientific, Virgo, Fermi-GBM, *Astrophys. J. Lett.* 848 (2017) L13, arXiv:1710.05834.
- [45] C.J. Moore, E. Finch, R. Buscicchio, D. Gerosa, arXiv:2103.16486, 2021.
- [46] N. Aghanim, et al., Planck, *Astron. Astrophys.* 641 (2020) A6, arXiv:1807.06209.
- [47] E. Bellini, A.J. Cuesta, R. Jimenez, L. Verde, *J. Cosmol. Astropart. Phys.* 02 (2016) 053, Erratum: *J. Cosmol. Astropart. Phys.* 06 (2016) E01, arXiv:1509.07816.
- [48] D. Alonso, E. Bellini, P.G. Ferreira, M. Zumalacárregui, *Phys. Rev. D* 95 (2017) 063502, arXiv:1610.09290.
- [49] J. Gleyzes, *Phys. Rev. D* 96 (2017) 063516, arXiv:1705.04714.
- [50] P. Madau, M. Dickinson, *Annu. Rev. Astron. Astrophys.* 52 (2014) 415, arXiv:1403.0007.
- [51] S. Vitale, W.M. Farr, K. Ng, C.L. Rodriguez, *Astrophys. J. Lett.* 886 (2019) L1, arXiv:1808.00901.
- [52] J.B. Jiménez, J.M. Ezquiaga, L. Heisenberg, *J. Cosmol. Astropart. Phys.* 04 (2020) 027, arXiv:1912.06104.
- [53] J.M. Ezquiaga, W. Hu, M. Lagos, M.-X. Lin, arXiv:2108.10872, 2021.
- [54] E.S. Phinney, arXiv:astro-ph/0108028, 2001.
- [55] A. Maselli, S. Marassi, V. Ferrari, K. Kokkotas, R. Schneider, *Phys. Rev. Lett.* 117 (2016) 091102, arXiv:1606.04996.
- [56] A. Saffer, K. Yagi, *Phys. Rev. D* 102 (2020) 024001, arXiv:2003.11128.
- [57] R.C. Nunes, *Phys. Rev. D* 102 (2020) 024071, arXiv:2007.07750.
- [58] R. Abbott, et al., LIGO Scientific, Virgo, *Phys. Rev. Lett.* 125 (2020) 101102, arXiv:2009.01075.
- [59] M.C. Straight, J. Sakstein, E.J. Baxter, arXiv:2009.10716, 2020.
- [60] C. Messenger, J. Read, *Phys. Rev. Lett.* 108 (2012) 091101, arXiv:1107.5725.
- [61] C. Ye, M. Fishbach, arXiv:2103.14038, 2021.
- [62] R. Farmer, M. Renzo, S.E. de Mink, P. Marchant, S. Justham, arXiv:1910.12874, 2019.
- [63] M. Dominik, E. Berti, R. O'Shaughnessy, I. Mandel, K. Belczynski, C. Fryer, D.E. Holz, T. Bulik, F. Pannarale, *Astrophys. J.* 806 (2015) 263, arXiv:1405.7016.
- [64] M. Fishbach, Z. Doctor, T. Callister, B. Edelman, J. Ye, R. Essick, W.M. Farr, B. Farr, D.E. Holz, arXiv:2101.07699, 2021.
- [65] D. Traykova, E. Bellini, P.G. Ferreira, *J. Cosmol. Astropart. Phys.* 08 (2019) 035, arXiv:1902.10687.
- [66] T. Baker, I. Harrison, *J. Cosmol. Astropart. Phys.* 01 (2021) 068, arXiv:2007.13791.
- [67] S. Mastrogiovanni, D. Steer, M. Barsuglia, *Phys. Rev. D* 102 (2020) 044009, arXiv:2004.01632.
- [68] H. Audley, et al., LISA, arXiv:1702.00786, 2017.
- [69] D. Foreman-Mackey, D.W. Hogg, D. Lang, J. Goodman, *Publ. Astron. Soc. Pac.* 125 (2013) 306, arXiv:1202.3665.
- [70] D. Foreman-Mackey, *J. Open Sour. Softw.* 1 (2016) 24, <https://doi.org/10.21105/joss.00024>.
- [71] I. Mandel, W.M. Farr, J.R. Gair, *Mon. Not. R. Astron. Soc.* 486 (2019) 1086, arXiv:1809.02063.