Probing gravity and growth of structure with gravitational waves and galaxies' peculiar velocity

A. Palmese[1,2] and A. G. Kim[3]

[1]*Fermi National Accelerator Laboratory, P. O. Box 500, Batavia, Illinois 60510, USA*
[2]*Kavli Institute for Cosmological Physics, University of Chicago, Chicago, Illinois 60637, USA*[*]
[3]*Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

(Received 13 May 2020; accepted 7 April 2021; published 7 May 2021)

The low-redshift velocity field is a unique probe of the growth of cosmic structure and gravity. We propose to use distances from gravitational wave (GW) detections, in conjunction with the redshifts of their host galaxies from wide field spectroscopic surveys (e.g., DESI, 4MOST, TAIPAN), to measure peculiar motions within the local Universe. Such measurement has the potential to constrain the growth rate $f\sigma_8$ and test gravity through determination of the gravitational growth index γ, complementing constraints from other peculiar velocity measurements. We find that binary neutron star mergers with associated counterpart at $z \lesssim 0.2$ that will be detected by the Einstein Telescope (ET) will be able to constrain $f\sigma_8$ to $\sim 3\%$ precision after 10 years of operations when combined with galaxy overdensities from DESI and TAIPAN. If a larger network of third generation GW detectors is available (e.g., including the Cosmic Explorer), the same constraints can be reached over a shorter timescale ($\sim$5 years for a 3 detectors network). The same events (plus information from their hosts' redshifts) can constrain γ to $\sigma_\gamma \lesssim 0.04$. This constraint is precise enough to discern general relativity from other popular gravity models at 3σ. This constraint is improved to $\sigma_\gamma \sim 0.02$–0.03 when combined with galaxy overdensities. The potential of combining galaxies' peculiar velocities with gravitational wave detections for cosmology highlights the need for extensive optical to near-infrared follow-up of nearby gravitational wave events, or exquisite GW localization, in the next decade.

DOI: 10.1103/PhysRevD.103.103507

I. INTRODUCTION

The motion of galaxies on top of the cosmological expansion, i.e., their peculiar velocity field, follows the inhomogeneous clustering of structure in the Universe and the laws of gravity. Peculiar velocities of galaxies thus encode important information about large scale structure and its growth, and can probe models of gravity. The velocity field can be studied in several ways. One possibility is through redshift space distorsions (RSD), since peculiar motions alter the correlations between galaxies along the line of sight. Another option is to derive peculiar velocity measurements from redshift surveys using galaxies that also have a distance estimate, so that the contribution due to the cosmological expansion can be subtracted out. Such distances can be estimated from the fundamental plane relation for elliptical galaxies, the Tully–Fisher relation for spiral galaxies, and, as it has been suggested more recently, from type Ia supernovae (SNe Ia) (e.g., [1–5]).

Gravitational wave (GW) detections also provide a distance measurement. If a single galaxy can be associated with a GW event, either through an observable electromagnetic (EM) counterpart, or thanks to an exquisite localization by the GW detectors, the host galaxy will also inherit the GW distance measurement. Such association was possible for GW170817 [6] for the first time, whose counterpart [7] was found in the old elliptical galaxy NGC 4993 [8]. The first cosmological parameter estimates from GW detections ([9]; [10]; [11]) rely on this association between the binary merger and the host galaxies, and are performed using the "standard siren" method [12], which makes use of the distance-redshift relation to probe the expansion of the Universe. In other words, the GW distance estimates are used with host galaxies' redshifts to populate the Hubble diagram. In these works, the peculiar motion of galaxies can give rise to systematic uncertainties and biases, and several works have studied how these can be properly taken into account in standard siren measurements [13,14]. More generally, multiple studies have discussed the importance of distorsions in the luminosity distance due to peculiar velocities and gravitational lensing of foreground large scale structure (e.g., [15–19]). It is important to note that peculiar motions also contain interesting cosmological information that we could exploit.

[*]palmese@fnal.gov

The authors of [20] recently suggested using gravitational waves with counterparts as distance indicators to measure peculiar velocities, and discussed implications for estimates of ΛCDM cosmological parameters using a comparison method between the density field of galaxies and peculiar velocities from GWs using measurements up to 190 Mpc. In this work, we extend this method to include autocorrelations of overdensities and of the velocity field, other than cross-correlations, including information from the CMB as explained in more detail in Appendix A. Moreover, we show how this method is interesting beyond 190 Mpc, and beyond ΛCDM to determine the gravitational growth index γ as a tool to test general relativity (GR).

We focus on the results that will be obtained using third generation (3G) gravitational wave detectors, such as the Einstein Telescope (ET; [21]) and the Cosmic Explorer (CE; [22]), because current-generation detectors would not produce luminosity distance measurements precise enough to be competitive for this analysis (currently $\sigma_d > 10\%$; [23]). Moreover the limited distance horizon of the LIGO/Virgo/KAGRA network ($\lesssim$200 Mpc for binary neutron star mergers; [23]) results in small number statistics. On the other hand, 3G GW experiments will detect up to millions of compact object binary mergers across the history of the Universe every year, including binary neutron star mergers (BNS) out to $z \sim 2$ and binary black hole mergers out to $z \sim 20$. Detections of BNSs' GW signal in the local Universe will be highly complete over the full sky, within the redshift range we are interested in for our purposes ($z < 0.3$). Several works have shown how these sources can be used as standard sirens for precision cosmology (e.g., [24–27]), although attention has to be paid at the way the analysis is carried out to avoid introducing biases [28,29].

The promise of future ground-based GW detectors is also particularly compelling for peculiar velocity studies since the expected precision on luminosity distance measurements can reach few per cent in the nearby Universe [27], and even sub–percent with a space-based experiment such as DECIGO [30,31]. A large number of facilities is expected to be involved in following up these events (e.g., [32,33]), and it is reasonable to assume that the next generation of telescopes will be able to detect optical counterparts to nearby $z < 0.3$ events (e.g., LSST, WFIRST; [34]). These data can be combined with precise host-galaxy redshifts from ongoing and planned wide–field spectroscopic galaxy surveys, such as the Dark Energy Spectroscopic Instrument (DESI; [35]), the 4-metre Multi-Object Spectroscopic Telescope (4MOST; [36]) and TAIPAN [37], to derive the constraints presented in this work.

This article is structured as follows. In Sec. II we define the specifics of the GW experiment and the galaxy survey that are relevant for the forecast. In Sec. III we present the method used to derive the constraints, including the Fisher formalism. Section IV contains our forecasted constraints on the gravitational growth index for 3G GW detectors, and a comparison to the expected precision using SNe Ia with peculiar velocities. In Sec. V we present the conclusions. Throughout this article, we assume a flat ΛCDM cosmology with $H_0 = 100\ h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $\Omega_{m0} = 0.3$. Quoted uncertainties are 1σ.

II. EXPERIMENTS SPECIFICS

In this work, we focus on gravitational wave detections of binary neutron star (BNS) mergers. This is justified by the fact that BNSs are expected to be accompanied by an EM counterpart [38,39], and this has been observed for the event GW170817 [7]. BNS rates are estimated to be 1 to 2 order of magnitude lower than SNe Ia: $\sim 0.25\text{–}2.81 \times 10^{-6}\ \mathrm{Mpc^{-3}\,yr^{-1}}$ [40] versus $2.69 \times 10^{-5}(h/0.70)^3\ \mathrm{Mpc^{-3}\,yr^{-1}}$ [41]. Note that we use this effective rate for the $z < 0.3$ considered in this article, and do not account for the redshift-dependent increase in rest-frame rates. While the rates imply that the number statistics of GWs will be less constraining than SNe, the promise of more precise distance measurements in the local Universe and full sky detection sensitivity can still make the method presented here a competitive probe. The expected distance precision for an ET-CE network is a few percent in the local Universe, depending on the network configuration [27], while SN standardization methods leave an ∼5% distance uncertainty floor that is not improved with precision photometry [e.g., [42]].

Moreover, a CE-ET network of multiple detectors can be considered sensitive to all sky directions to first order, in particular for the nearby BNS considered here, resulting effectively in a 4π experiment. However, we still need to identify EM counterparts, meaning that even if we assume that there is an instrument on the ground and/or in space that is able to identify them at any time, the discovery would still be challenging close to the Galactic plane. We will make the assumption that ∼10% of the sky area is lost because of this. We consider different time windows for the experiment, up to 18 years, to show the scaling of the constraints with the number of detected sources.

In addition, we consider perspectives for current generation GW detectors. We take into account a LIGO/Virgo configuration (HLV) at design sensitivity (expected for 2022+), which is able to detect BNS events out to 190 Mpc [23], with a ∼20% uncertainty on the distance on average (e.g., [43]). Several works have shown how independent EM observations can constrain the binary viewing angle [44–47], thus providing an external constraint that breaks the inclination angle–distance degeneracy, one of the main sources of uncertainty for the distance precision. When estimated independently of the distance, viewing angle constraints can be as good as ∼5 deg for GW170817 when using x-ray and radio data, and ∼10 deg from optical-NIR

data [47], leading to a factor 2 or better improvement on the distance precision.

In all of the scenarios considered, the EM counterparts need to be accompanied by host galaxy redshifts. While it is realistic to assume that they will be measured as part of the GW follow–up campaigns, it is worth noting that most of these galaxies in the local Universe are likely to already have a spectroscopic redshift measured as part of ongoing and upcoming galaxy surveys such as DESI and TAIPAN. The same surveys will also measure RSDs, and in the following we consider their complementarity with the proposed peculiar velocity surveys. In particular, we consider a galaxy survey with number density of $n_g \sim 10^{-3}\ h^3\,\mathrm{Mpc}^{-3}$ out to $z < 0.2$, and $10^{-4}\ h^3\,\mathrm{Mpc}^{-3}$ at $0.2 < z < 0.3$, which is realistic for TAIPAN [48], and it is a conservative lower bound for the DESI Bright Galaxy Survey (BGS; [35]).

The last ingredient needed in our analysis is the range of scales to consider in the power spectra. Following [49], the minimum and maximum modes k are set by the linear extent of the GW experiment volume ($k_{\min} = (\pi/r_{\max})h/\mathrm{Mpc}$, with $r_{\max}$ being the radius of the volume under consideration) and by the smallest scales that are confidently modeled ($k_{\max} = 0.1\ h/\mathrm{Mpc}$), respectively.

III. METHOD

The connection between peculiar velocities and the growth of structure is straightforward to understand in linear theory. The amplitude of overdensities scales with the growth factor D, which evolves as $f \equiv \frac{d\ln D}{d\ln a}$, the linear growth rate. From the conservation of mass, the velocity field scales with the overdensity field by a factor f. The peculiar velocity power spectrum is related to the overdensity power spectrum at the time of the CMB as $P_{vv} \propto (fD\mu)^2 P_{\delta\delta}(z = \mathrm{CMB})$. (See Appendix B for an overview).

The growth of structure depends on gravity. [50,51] show that the linear growth rate is well–approximated by $f \approx \Omega_{\mathrm{m}}^{\gamma}$ for several gravity models, where γ is the growth index. They also find that for general relativity, $f(R)$, and DGP gravity (see [52] for a review), $\gamma = 0.55, 0.42, 0.68$, respectively. It is thus interesting to constrain γ with a precision of $\sigma_\gamma/\gamma \lesssim 20\%$ at 3σ (in other words, $3\sigma_\gamma \sim 0.1$) to be able to discern between GR and other gravity models at $\sim$99% CL. Using this parameterization, the peculiar velocity power spectrum probes gravity through γ in:

$$fD = a_{\mathrm{CMB}}\Omega_{\mathrm{m}}^{\gamma} e^{\int_{a_{\mathrm{CMB}}}^{a} \Omega_{\mathrm{m}}^{\gamma} \mathrm{d}\ln a} \tag{1}$$

where $\Omega_{\mathrm{m}}(a)$ is the matter density at $a(t)$, and it also depends on the gravity model. We choose to anchor the linear growth factor at the CMB $D(a_{\mathrm{CMB}})a_{\mathrm{CMB}}^{-1} = 1$.

The same GW events used to measure peculiar velocities (or some other source) can also serve as tracers of mass overdensities. The overdensity power spectrum in redshift space for mode $\vec{k}$ also depends on gravity through $P_{\delta\delta} \propto (bD + fD\mu^2)^2$ where b is the GW host galaxies' bias and $\mu \equiv \cos(\hat{k}\cdot\hat{r})$ where $\hat{r}$ is the direction of the line of sight [53]. The bias is a nuisance parameter in our analysis, so we marginalize over it when inferring γ. Future analyses of GW events could establish a prior for this parameter, which is likely to depend on the formation channels for the class of mergers considered. The other power spectrum to consider is the galaxy-velocity cross-correlation, which goes as $P_{v\delta} \propto (bD + fD\mu^2)fD\mu$. For a derivation of the power spectra see [54].

In this work we have decided to focus on peculiar velocities, but the formalism could be also expressed in terms of "peculiar distances," and an example of this is provided in [55] for the case of luminosity distance space distorsions.

Forecasts on the growth index are then computed using the Fisher matrix formalism, following [48,56]. The Fisher information matrix can be written as:

$$F_{ij} = \frac{\Omega}{8\pi^2}\int_{r_{\min}}^{r_{\max}}\int_{k_{\min}}^{k_{\max}}\int_{-1}^{1} r^2k^2\mathrm{Tr}\left[C^{-1}\frac{\partial C}{\partial\lambda_i}C^{-1}\frac{\partial C}{\partial\lambda_j}\right]d\mu dk dr \tag{2}$$

where

$$C(k,\mu,a) = \begin{bmatrix} P_{\delta\delta}(k,\mu,a) + \frac{1}{n} & P_{v\delta}(k,\mu,a) \\ P_{v\delta}(k,\mu,a) & P_{vv}(k,\mu,a) + \frac{\sigma^2_{v,\mathrm{eff}}}{n} \end{bmatrix}, \tag{3}$$

Ω is the solid angle over which GW sources are detected, n is the GW events number density, r is the comoving distance and $r_{\max}$ (corresponding to the redshift $z_{\max}$) is the maximum comoving distance at which events are detected. The minimum distance, $r_{\min}$, is set to $z = 0.01$ for SNe to ensure small propagated velocity uncertainties, and to $z = 0.001$ for GWs because the distance uncertainty is smaller than SNe, especially at these low distances. However, we find that our results are not significantly affected by a more conservative $z = 0.01$ cut.

One parameter set we consider normalizes growth to the amplitude of clustering today, such that $D(z) = \sigma_8(z)$, noting that the shape of the power spectrum does not change in our model. In this case, $\lambda \in \{\{\langle f\sigma_8\rangle\}, b\sigma_8\}$ where $\langle f\sigma_8\rangle$'s refer to the effective $f\sigma_8$'s in a set of redshift bins, and the combination $b\sigma_8$ is constant in all redshift bins. We also consider the parameter set $\lambda \in \{\gamma, \Omega_{\mathrm{m}0}, b\}$ where $\Omega_{\mathrm{m}0}$ is the matter density at $z = 0$. Taking ΛCDM as our fiducial model, $\Omega_{\mathrm{m}} = \frac{\Omega_{\mathrm{m}0}}{\Omega_{m0}+(1-\Omega_{\mathrm{m}0})a^3}$. The uncertainty in the growth index is then given by $\sqrt{F^{-1}_{\gamma\gamma}}$.

The $\sigma^2_{v,\mathrm{eff}}$ in the shot-noise term is the velocity variance derived from the ensemble of sources within a differential cell volume [57,58]. For indicators whose distance

uncertainties are dominated by intrinsic magnitude dispersion, all objects in the same cell share the same relative distance uncertainty and $\sigma^2_{v,\mathrm{eff}} = \sigma^2_v$ [see Eq. (4)]. In the case of GW sources, whose different inclination angles result in a broad range of distance uncertainties, the appropriate number to use for the effective per-object velocity variance is $\sigma^2_{v,\mathrm{eff}} = \langle \sigma_v^{-2} \rangle^{-1}$.

For the Fisher calculation we take $\gamma = 0.55$, $\Omega_{m0} = 0.3$, $b = 1.2$ (as expected from [37]), and the equivalent $f\sigma_8$. We use an external prior for Ω_{m0} with 0.005 uncertainty. The matter power spectrum at $P_{\delta\delta}$ at $z = 0$ is calculated using the default configuration of CAMB [59]. This is then propagated to redshift z through $D(z;\gamma)$, as explained in detail in Appendix B.

We consider the contribution from the non-linear velocity which can affect scales of $k > 0.1\ h/\mathrm{Mpc}$ by adding in quadrature to the velocity uncertainty a velocity dispersion term of $300\ \mathrm{km\,s^{-1}}$ (which corresponds to the typical peculiar velocity dispersion found for galaxy clusters; [60–62]). Note that we are highly shot-noise dominated at $k > 0.1\ h\,\mathrm{Mpc}^{-1}$ and so our results are insensitive to the nonlinear regime. The per-object peculiar velocity uncertainty σ_v is then related to the distance uncertainty through:

$$\sigma_v^2 = \left(1 - \frac{1}{aH\chi}\right)^{-2} \left(\frac{\sigma_d}{d}\right)^2 + \left(\frac{300\ \mathrm{km\,s^{-1}}}{c}\right)^2, \quad (4)$$

where a is the scale factor, H is the Hubble parameter and χ is the comoving distance, all computed at the distance d.

In the following, we will assume that the fractional relative distance uncertainty from the GW experiment scales as the inverse of the SNR of the GW signal ρ: $\sigma_d/d \propto 1/\rho \propto d$ (e.g., [26,63]), and show results for a range of different normalizations to this relation. We ignore the weak-lensing contribution to the distance uncertainty since it does not have a significant contribution at the redshifts considered here.

We also show results for specific network configurations that have been studied in the literature. At the time of writing, realistic distance uncertainty distributions for 3G experiments such as those presented in [27] were not available. We therefore use approximations for $\sigma^2_{v,\mathrm{eff}}/n$ in Eq. (3), based on the available information.

The first approximation is to use only the mean $\langle (\sigma_d/d)^{-2} \rangle^{-1}$ in redshift bins to describe a sample of events with a distribution of different (σ_d/d):

$$\sigma^2_{v,\mathrm{eff}} = \langle \sigma_v^{-2} \rangle^{-1} \quad (5)$$

$$\approx \left(1 - \frac{1}{aH\chi}\right)^{-2} \left\langle \left(\frac{\sigma_d}{d}\right)^{-2} \right\rangle^{-1} \quad (6)$$

This approximation is useful when considering the ET configuration (three interferometers with 60 deg opening angles and 10 km arms, arranged in a triangle) studied in [25]. We use their fit for the redshift dependence of $\langle (\sigma_d/d)^{-2} \rangle^{-1} \equiv A(z)^{-1}$ [their Eq. (32)], which holds for a subsample $n = (1 - \cos 20°) n_T$ of face-on systems with inclination angles $\iota < 20$ deg, out of the total number of systems n_T. This is a pessimistic scenario as the low-inclination subset has the worst distance precision [27,43], and the full population must have a smaller $\sigma^2_{v,\mathrm{eff}}$. As a less pessimistic bound, we use the above approximation for $\sigma^2_{v,\mathrm{eff}}$, but we use the full population n_T. We note that the quoted distance uncertainty estimates are marginalized over the inclination angle.

The second estimate that we consider consists in limits. From information such as the median, we know that a subpopulation f e.g., 0.5 of the GW-population distances would have $\frac{\sigma_d}{d} < 0.01$ at $z = 0.1$. Then we can estimate

$$\frac{\sigma^2_{v,\mathrm{eff}}}{n} \lesssim \frac{(1 - \frac{1}{aH\chi})^{-2} (0.01^2 + (\frac{300\ \mathrm{km\,s^{-1}}}{c})^2)}{f n_T}.$$

Note that this is a pessimistic estimate when considering the median, given that the distribution does go to lower uncertainties for edge–on GW systems, and that we only use a fraction of the available data. For this reason we treat it as an upper limit.

IV. RESULTS

A. Growth of structure in GR

We first compute the expected constraint on $f\sigma_8$ in three redshift bins in the range $0.0 < z < 0.3$. The fiducial $f\sigma_8$ values correspond to the expectation from ΛCDM.

Our result for a 10 year ET configuration as in [25] is $\sigma(f\sigma_8)/f\sigma_8 = 0.0513,\ 0.0485,\ 0.0921$ at $z = 0.05,\ 0.15$ and 0.25, respectively, as shown in Fig. 1 by the blue triangles with light blue errorbars. The best constraint is reached in the second bin because at the lowest redshifts cosmic variance and intrinsic velocity dispersion give a larger contribution to the final uncertainty, while in the highest redshift bin the relative distance uncertainty is ∼3 times larger than in the second bin. This result is compared to our expected constraints from a 10 year SN survey covering 2π of the sky that is able to recover a SN magnitude dispersion of $\sigma_M = 0.08$ (red triangles). Although GW distances can be significantly better than those of SNe at $z < 0.1$, the SN survey shows slightly better constraints over the first redshift bins because of the higher rate of these transients. The difference between the two results becomes more significant in the last bin, because the SN precision on the distance is assumed to be constant with redshift.

In Fig. 1 these constraints (light blue triangles, outer error bars) are compared to existing $f(z)\sigma_8(z)$ measurements from 6dF [64], WiggleZ [65], SDSS-II LRG [66], SDSS-II Main Galaxy sample [67], BOSS [68], VIPERS [69] and eBOSS-CMASS [70].

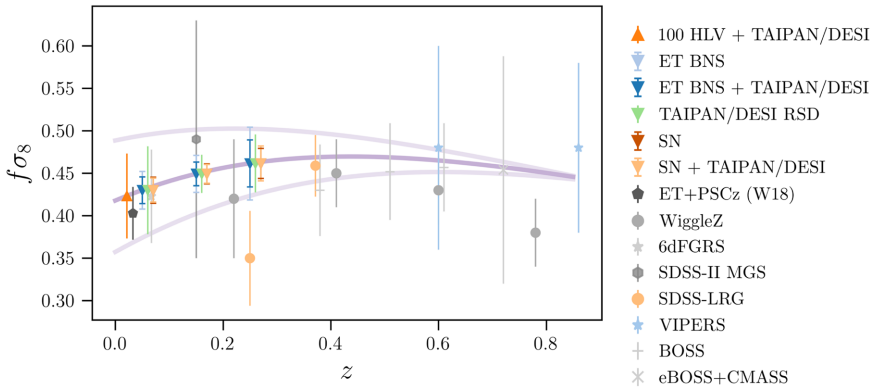

FIG. 1. Expected constraints on $f(z)\sigma_8(z)$ using peculiar velocities measured using binary neutron star mergers from a 10 year 3G GW experiment (light blue errorbars with triangles) and in combination with DESI/TAIPAN (dark blue errorbars). DESI/TAIPAN RSD-only conservative expected constraints are shown by the green triangles, shifted to higher z for visualization purposes. These are compared to constraints from a 10 year SN survey covering 2π of the sky that is able to recover a SN magnitude dispersion of $\sigma_M = 0.08$ in the same redshift bins (red triangles, shifted to higher z for visualization purposes). The results from a combination of such survey with DESI/TAIPAN RSDs is shown by the light orange triangles, and they are very similar to the SN-only case. Prospects for 100 BNSs from LIGO/Virgo (HLV) at design sensitivity are also reported (darker orange triangle), in combination with galaxy surveys. All the results represented by triangles are computed in this work. The remaining data points represent existing $f(z)\sigma_8(z)$ measurements from 6dF [64], WiggleZ [65], SDSS-II LRG [66], SDSS-II Main Galaxy sample [67], BOSS [68], VIPERS [69] and eBOSS-CMASS [70]. We also report the forecast from [20] (black pentagon), who use a combination of GW and PSCz galaxy survey data. The darker purple line is the theoretical prediction for $f(z)\sigma_8(z)$ in a Flat ΛCDM Universe with $\gamma = 0.55$ (GR), while the other curves show the theoretical prediction from $\gamma = 0.42$ and $\gamma = 0.68$ (which are the values predicted for $f(R)$ and DGP gravity, respectively).

The above results show that GW distances and hosts *alone* can already place interesting constraints on $f\sigma_8$. It is important to note that galaxies residing inside a galaxy cluster will require spectroscopic observations of the whole cluster, in order to enable a measurement of the velocity dispersion of cluster galaxies and of the redshift of the cluster. In fact, in these cases the redshift of the galaxy cluster, rather than the redshift of the member galaxy, should be used in the measurement of the power spectra discussed here. It is also worth noting that the velocity of the center of mass of the binary inside the host galaxy does not need to be taken into account in this analysis, as the redshift used is measured from the galaxy (or galaxy cluster), not from the binary.

More precise results can be reached by the addition of a dense galaxy survey, such as those mentioned in Sec. II. The number density of events n considered in the overdensity power–spectrum in Eq. (3) is replaced by the much larger galaxy number density n_g. Our Fisher matrix constraints are reported in Table I and shown by the dark blue error bars in the Figure. With a ∼3% precision at $z \lesssim 0.2$, these bounds are competitive with the constraints expected from the aforementioned SN experiment, and with the forecast by [48] for a combination of TAIPAN with HI surveys WALLABY (Widefield ASKAP L-band Legacy All-sky Blind Survey; [71]) and Westerbork Northern Sky HI Survey (WNSHS). In Table I, we also report the expected constraints from SNe and RSD from a TAIPAN/DESI-like experiment. The addition of galaxies brings marginal improvement in the first two z bins, while the third bin is better constrained by the SN hosts' rather than the survey galaxies' overdensity power spectrum. This is due to the fact that for a 10 year survey, the number density of SNe is ∼5 times larger than $n_g = 10^{-4}\ h^3\,\mathrm{Mpc}^{-3}$.

The green triangles show our expected constraints from RSD only for a galaxy survey with the number densities mentioned above. It is clear that the GW measurement can bring significant additional information to RSDs in the

TABLE I. Expected constraints on $f(z)\sigma_8(z)$ using peculiar velocities measured using binary neutron star mergers from a 10 year ET and in combination with DESI/TAIPAN out to $z < 0.3$. For comparison, we also report constraints from a 10 year SN survey.

Data	$\frac{\sigma(f\sigma_8)}{f\sigma_8}(z = 0.05)$	$\frac{\sigma(f\sigma_8)}{f\sigma_8}(z = 0.15)$	$\frac{\sigma(f\sigma_8)}{f\sigma_8}(z = 0.25)$
ET GW BNS	0.0513	0.0485	0.0921
ET GW BNS + TAIPAN/DESI	0.0365	0.0311	0.0598
SN	0.0357	0.0261	0.0381
SN + TAIPAN/DESI	0.0327	0.0248	0.0445

lowest redshift bin (reducing the RSD-only 1σ by $\sim 70\%$), while the two probes provide similar constraints in the remaining bins, where their combination provides a 38 and 20% improvement respectively. Note that these results hold for an ET-only configuration. If a larger network of 3G detectors is built, and can reach an average 1% uncertainty in distance at $z = 0.1$ (as it is realistic for three detectors), the same constraints can be reached after only ~ 5 years.

GW forecasts have been previously made by [20] (black pentagon), who use a similar method to what proposed in this work. We note that the number of events they use out to $z = 0.045$ corresponds to a ~ 3–4 year experiment with the configuration assumed here, but the competitive constraints are achieved by combining the GW data with the point source catalog redshifts (PSCz; [72,73]) galaxies. Unlike this article, they do not use velocity-velocity correlations to inform $f\sigma_8$. We consider a [20]-like survey in our formalism by considering the PSCz number density out to $z = 0.045$, and a 3 year GW experiment with distance errors $\sigma_d/d = 1\%$ for all events out to $z = 0.045$. We find that $\sigma(f\sigma_8)/f\sigma_8 = 0.0474$, versus $\sigma(f\sigma_8)/f\sigma_8 = 0.0769$ from [20] (38% improvement).

In addition, we show constraints from 100 events detected by HLV at design sensitivity, combined with a galaxy survey with $n_g = 10^{-3}\ h^3\,\mathrm{Mpc}^{-3}$. We find that this constraint is mostly dominated by the overdensity power spectrum, with the peculiar velocities only providing a $\sim 3\%$ improvement. When additional constraints on the viewing angle are available, the improvement can reach 8%–15% if the constraint is 5–10 deg at 1σ, or if the viewing angle has an upper bound of 30 deg or less, using the derived distance precision from [74].

The dark purple line in Fig. 1 is the $f(z)\sigma_8(z)$ evolution computed from theory assuming a flat ΛCDM with $\gamma = 0.55$ (GR), while the other curves show the theoretical prediction from $f(R)$ and DGP gravity. It is clear that the measurement proposed in this work will allow us to place interesting constraints on these gravity models.

B. Testing GR

In the second part of this work, in place of determining $f\sigma_8$ in different redshift bins, we let γ be a free parameter. Expected constraints on the growth index are shown in Fig. 2. The left hand plot shows the precision for different values of effective fractional distance uncertainty for different volumetric rates integrated over time (i.e., $n \times t$). The distance uncertainty σ_{d_*} is the uncertainty at a reference distance d_*, here corresponding to $z_* = 0.1$. Events at distances different from d_* have a distance precision that scales as described in Sec. II. The number of events considered corresponds to the total of an experiment length up to 18 years for a BNS volumetric restframe rate corresponding to the maximum *a posteriori* from the latest estimate, $1.09 \times 10^{-6}\ \mathrm{Mpc}^{-3}\,\mathrm{yr}^{-1}$ [40].

If the rate is lower, it will take more time to reach the same constraints. The effect of the shift in number density from the high to the low bound of the 90% CL range is shown by the white line on the left hand plot for a 5 years experiment, for GW sources out to $z = 0.3$, with a distance precision of 1% (or following a distribution of distance uncertainties with $\langle(\sigma_d/d)^{-2}\rangle^{-1/2} = 0.01$) at $z = 0.1$. We find that data from a 5 year GW experiment that is able to reach a $\lesssim$ few per cent precision on average for sources at $z = 0.1$ will be enough to constrain the growth index to $\sigma(\gamma) \lesssim 0.04$ if the BNS rate is on the high end, allowing us to discern between GR and other popular gravity models. This precision level is in fact reached by the ET configuration studied in [25]. When using their approximation for the redshift dependance of the distance uncertainty, we recover $\sigma(\gamma) = 0.039$ after 5 years if we detect EM counterparts out to $z_{\max} = 0.3$, and $\sigma(\gamma) = 0.041$ if $z_{\max} = 0.2$.

Expected limits for other different configurations of detectors (estimated with the approximation described in Sec. III from the median of a population) are given by the boxes in Fig. 2. Using the median values from [27], we can provide an extreme pessimistic bound using only the best 50% of events, that will have a distance precision equal or better than what reported on the y-axis. The extreme optimistic bound is chosen as the lowest precision considered in [27] for the redshift bins in consideration. The other two bounds of the boxes can be drawn by considering all sources having the median and lowest precision of the sample, respectively. The configurations considered include one ET detector, one ET with ideal low–frequency noise described in [27], 2 ET detectors (which is similar to the ET + CE case), and the 3 ET (or 3CE) network.

The right hand side plot of Fig. 2 shows the expected precision on γ for different values of distance uncertainty and different maximum redshift values, starting from the same assumptions of the left hand side plot. The rapid flattening of the σ_γ contours with the redshift show that, for a given distance precision, adding events at higher redshifts only very slowly helps constraining the growth index, due to the increasing distance uncertainty for more distant events. The assumption that we have made on the detectability of KNe out to $z \sim 0.3$ can be therefore relaxed to reach only $z \sim 0.15$ for an experiment with $\sigma_{d_*}/d_* = 5\%$ or $z \sim 0.2$ for an experiment with $\sigma_{d_*}/d_* = 1\%$, since the constraining power flattens out at higher redshifts for these distance precisions, and in both cases recovering $\sigma(\gamma) \lesssim 0.04$. On the low-redshift end of the Figure, for $z_{\max} \lesssim 0.05$ BNSs cannot even reach a 0.1 uncertainty on γ and thus cannot provide an interesting constraint using LIGO/Virgo at design sensitivity. The A+ upgrade of current detectors could provide more interesting constraints at the level of $\sigma(\gamma) \gtrsim 0.07$ only if improved distance measurements can be achieved by EM constraints on the geometry of the systems, as mentioned in Sec. II. This can be seen by following the arrow in the right panel of Fig. 2.

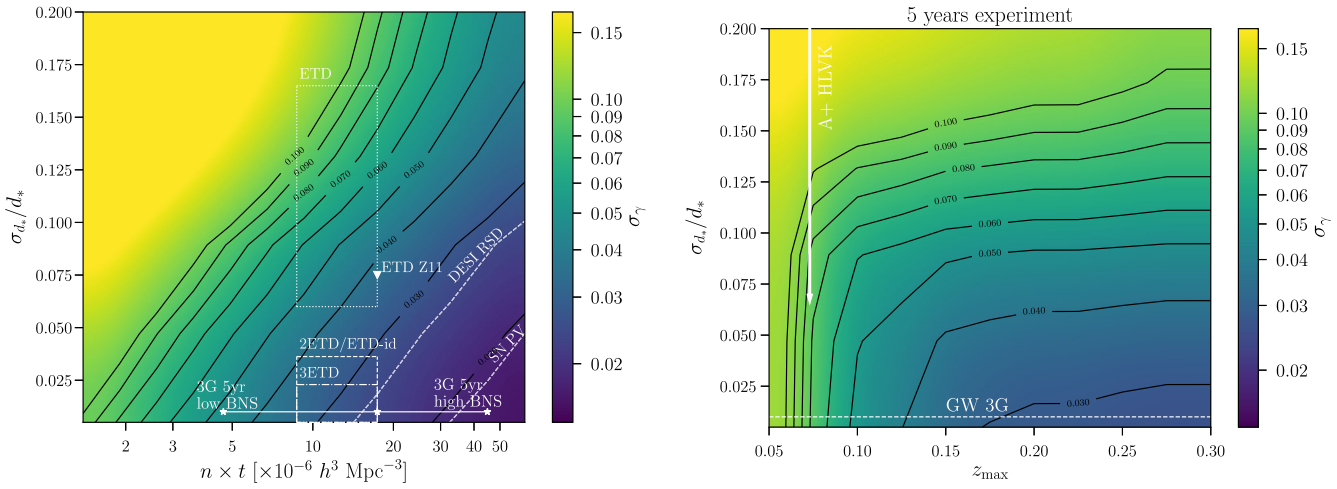

FIG. 2. Growth index uncertainty for different values of distance precision as a function of BNS volumetric rates integrated over time out to $z_{\max} < 0.3$ (left panel), and as a function of the maximum redshift $z_{\max}$ out of which we consider GW events. The distance uncertainty σ_{d_*} is the uncertainty at a reference distance d_*, here corresponding to $z_* = 0.1$. The BNS volumetric rate considered for the specific estimates on the left, and all points on the right panel, is the maximum *a posteriori* value from the latest GW estimate, $1.09 \times 10^{-6} \times (h/0.679)^3$ Mpc^{-3} yr^{-1} [40]. On the left panel, also the low and high 90% CL limits in the rate are shown for an ideal 3G experiment (where the distance precision is 1% at $z = 0.1$) after 5 years. The effect of the shift in number density from the high to the low bound of the 90% CL range is shown by the white line. The boxes represent limits within which we expect the constraint to fall for the various ET configurations studied in [27]. The white triangle shows our result using the [25] approximation for 1 ET. For comparison, we also show the constraints forecasted for DESI RSD and for an LSST–like SN peculiar velocity survey from [5] (lighter dashed lines). The distance precision for the ideal 3G experiment is also shown for reference on the right panel by the dashed line. The arrow shows possible constraints from the A+ LIGO/Virgo/KAGRA network, if improved viewing angle constraints can be derived from EM observations, reducing the distance uncertainty from order ∼20% down to 10% or better.

However, such a measurement would not be precise enough to discern between popular gravity models. Even if the measured distance precision improves, at low-redshift the 300 km s^{-1} random noise and sample variance dominate the dispersion in peculiar velocities.

Similarly to what presented in the previous subsection, we combine the peculiar velocity field measured from GWs with the overdensities from a RSD galaxy survey. In this case, the constraints are improved to $\sigma_\gamma \sim 0.02$–0.03 when combined with galaxy overdensities. On the other hand, a combination of RSD surveys covering ∼90% of the sky such as DESI + TAIPAN can only reach $\sigma_\gamma \sim 0.05$ at $z < 0.3$. However, note that DESI is expected to constrain $\sigma_\gamma = 0.026$ by measuring the growth of structure at larger redshifts out to $z < 1.6$ [5].

In Fig. 3 we show a comparison between the constraints from a GW and a SN experiment for different distance uncertainties. In this idealized case, BNSs and SNe experiments cover 90% of the sky for 10 years, and they have the same $\frac{\sigma_v^2}{n}$ entering in Eq. (3) at $z = 0.1$. In other words, the reported σ_d/d precision corresponds to a magnitude uncertainty at all redshifts $\sigma_M = \frac{5}{\ln(10)}\frac{\sigma_d}{d}$ for the SNe. The different behavior observed between the GW and the SN approach in measuring the growth index is driven by the different scaling of the distance measurement precision with redshift. The redshift dependence of the distance

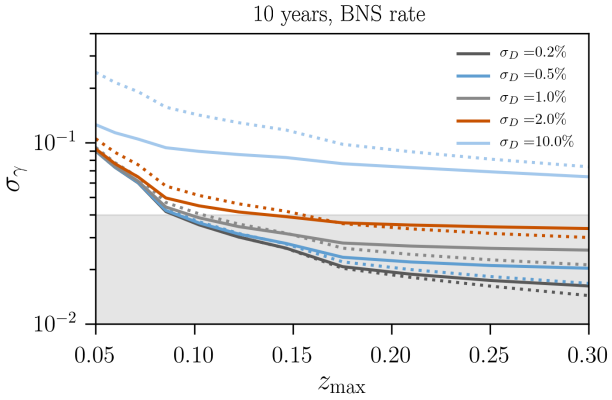

FIG. 3. Comparison between a SN (dashed lines) and a GW (solid lines) experiment to measure the growth index from peculiar velocities, for different distance uncertainties at $z = 0.1$, as a function of the maximum distance reach of the SNe/GW mergers. In this idealized case, BNSs and SNe have the same value of $\frac{\sigma_v^2}{n}$ at $z = 0.1$, implying that the reported σ_d/d precision corresponds to a magnitude uncertainty at all redshifts $\sigma_M = \frac{5}{\ln(10)}\frac{\sigma_d}{d}$ for the SNe. Both SN and GW experiments cover 90% of the sky. Differences in the γ constraints are driven by the different redshift dependence of the distance uncertainty measurement. The shaded region shows where constraints on γ can discern between popular gravity models ($\sigma_\gamma \lesssim 0.04$).

uncertainty makes the GW measurements more precise at $z < 0.1$, thus experiment is more constraining in γ than the SN case at the lowest redshift. At higher redshifts, adding more events does not improve significantly the constraints from GWs because the distance uncertainty is much less constraining than the events around $z \sim 0.1$, and the curve starts to flatten. This is not the case for SNe, since the intrinsic scatter in the magnitude does not depend on redshift, so including higher z events still provides a significant contribution. We also note that at $z < 0.1$ the expected constraint on γ is very similar for all distance uncertainties $\sigma_d < 1\%$, and for both the SN and GW cases. At low-redshifts $z < 0.05$, the advantage of a small σ_d/d quickly saturates out due to sample variance.

There are numerous works presenting a wide range of methods to test GR using GW sources (for a broad review, see [75,76]). One powerful constraint on gravity theories comes from the only event with counterpart, GW170817 (e.g., [77,78]). However, the method applied in those works are not able to probe gravity theories predicting a speed of gravitational waves equal to the speed of light, such as $f(R)$ theories, leaving many gravity theories still unexplored. This motivates studies of further methods. Another constraint that is already possible from GW170817 for $f(R)$ theories that change the effective Plank mass [79]. Moreover, [80] show that it is possible to constrain anisotropic stress as a modification of GW propagation, and that a comparison with measurements from large-scale-structure can lead to interesting constraints on gravity theories. Here we compare our findings with some of those works that are most similar to our method, and make use of the large scale structure of the Universe. First, most works on this subject probe a typical redshift which is higher than the one considered here, making our probe complementary to those. For example, [18] uses weak lensing magnification due to intervening large scale structure to discern between dark energy and modified gravity cosmological models, and [81] explores the effect of lensing within different gravity theories directly on the GW waveforms. The effect of lensing typically becomes relevant at redshifts that are larger than those considered here. [82] probe modified gravity theories using massive black hole binaries from the Laser Interferometer Space Antenna (LISA), which are more frequent at $z > 1$, where deviations from GR for Horndeski models are more evident. On the other hand, it is clear from Fig. 1 that the effect of most popular modified gravity theories on the growth of structure is more evident at the redshifts probed with this analysis. Moreover, to the best of our knowledge, no other method that constrains the growth index using gravitational waves has been presented to date, making the analysis presented here compelling and complementary to other methods probing other parameters.

V. CONCLUSIONS

In this article, we presented an alternative method to test GR and constrain the growth of structure by measuring the galaxies' peculiar velocity field using gravitational wave measurements of compact binary mergers. We find that data from a 5 year GW experiment, in conjunction with follow-up facilities, that are able to detect BNS mergers out to $z \sim 0.2$ and reach a $\lesssim$ few percent precision on average for sources at $z = 0.1$ will constrain $f\sigma_8$ to $\sigma(f\sigma_8) \sim 3$–4% in that redshift range. For a single ET, $\sigma(f\sigma_8) \sim 5\%$ is reached in 10 years.

Moreover, this method will constrain the growth index to $\sigma(\gamma) \lesssim 0.04$, allowing us to discern between GR and other popular gravity models at $\sim 3\sigma$. The distance precision and the detection horizon fit well with what is expected for a network of 3G GW detectors, such as ET and CE. On the other hand, the potential of the cosmological probe presented here highlights the need for extensive optical to near-infrared follow-up of nearby BNSs to identify the associated kilonovae. Once identified, it is likely that large number of host galaxies for these events will be already observed by planned spectroscopic surveys (DESI, 4MOST, TAIPAN) in the coming decade.

These results show that a self-contained 3G GW experiment combined with adequate optical-NIR follow-up efforts can provide interesting results to probe the growth of structure in the local Universe. Nevertheless, galaxy surveys are much more effective at measuring the overdensity field, and we find that they are complementary to the peculiar velocity field probed by the GWs. When we combine GWs and DESI+TAIPAN-like surveys, $f\sigma_8$ can be constrained to $\sim$2–3% at $z < 0.2$ depending on the network configuration, and $\sigma(\gamma) = 0.02$–0.03, which could be a decisive test for general relativity. We also find that the same RSD surveys alone can only reach $\sigma_\gamma \sim 0.05$ at these redshifts, demonstrating the value of adding peculiar velocities. Note that these results are competitive with other probes from upcoming experiments, such as from a combination of SNe from LSST and DESI [5].

There is a significant gain in constraining power ($\sim$38% at $z < 0.045$) when considering the peculiar velocity power spectrum and the overdensity power spectrum in addition to their cross–correlation. This is an advantage of the method presented here, compared to the one proposed in [20].

Our results show that events from LIGO/Virgo at design sensitivity cannot provide interesting constraints on the growth of structure, unless improved distance precision from EM observations can be pursued to provide a $\lesssim$10% distance precision for $\mathcal{O}(100)$ events, and these are combined with galaxy surveys overdensities. If this is possible, a $\gtrsim$10% improvement on the RSD-only 1σ measurements on $f\sigma_8$ from upcoming galaxy surveys can be achieved. Improved constraints on the viewing angle could improve prospects for the proposed method during the current decade. Optical-NIR data could be a particularly powerful tool for this purpose, since kilonovae can be observed from all directions and thus can possibly

be identified for most GW detections, unlike counterparts at other wavelengths [83].

In this work, we focused on BNS mergers. However, there are additional compact object mergers that could improve the statistics presented. In particular, it has been shown that neutron star–black hole mergers (NSBH) can be accompanied by an EM counterpart (e.g., [84]), and that they can provide improved distance constraints compared to BNSs [85], leading to improved cosmological constraints. Binary black hole mergers could also produce an electromagnetic counterpart within scenarios that can predict the presence of material around the binary, for example in AGN disks (e.g., [86,87]), or when the components themselves are the central black holes of dwarf galaxies [88,89]. Also NSBH and binary black hole mergers without counterpart can contribute to this analysis if their localization is so accurate to only fit one galaxy (a small fraction of events is expected to satisfy this condition already with current generation detectors; [90]).

In the analysis, we have assumed a constant bias, and this is a reasonable assumption since the peculiar velocity field measurement is only useful over a limited redshift range in the nearby Universe (our results hold for $z \lesssim 0.2$). The observed distribution of short gamma-ray bursts implies that BNS mergers simply follow the stellar mass content in the Universe, as only a fraction 5–13% resides in galaxy clusters [91], implying that it is reasonable to assume that the GW BNS bias could simply follow the bias of luminous galaxies. According to several works, the galaxy bias evolution of galaxies does not change dramatically (usually at the $\lesssim$10% level) since redshift 0.2, and it is close to 1 around $z = 0$ (e.g., [35,92,93]). We have tested different bias evolution relations and absolute values from [35,92,93] and conclude that the expected variations do not significantly affect the results reported here. We also note that assuming a constant bias over redshift bins of width $\sim$0.2 (note that our results are most interesting at $0 < z < 0.2$) is a common assumption in state-of-the-art dark energy experiments such as the Dark Energy Survey (see, e.g., [94,95]), and that a similar approach can be extended to GW sources, given the possible relation to luminous galaxies. When a peculiar velocity measurement with GW events will be possible, we would recommend using future galaxy bias constraints at the redshift range of interest, marginalizing over it with its uncertainty. We have verified that even a 20% change in the absolute value of the bias from 1.2 does not affect the main results of this work, and thus we expect the marginalization over an even more precise bias measurement that will be available in the future, to not significantly affect the constraints. From a theoretical standpoint, one would expect $b \gtrsim 1$, at least for binaries merging at the endpoint of stellar evolution ([96,97]), which justifies our choice of the bias value. In fact, since star formation efficiency increases with halo mass up to $\sim 10^{12}\ M_\odot$ (e.g., [98]), if the binaries are formed through the isolated binary scenario and are thus related to the star formation history, one would expect them to be more likely in galaxies that live in halos close to the stellar-to-halo mass relation (SHMR) peak, which is mostly unchanged up to $z \lesssim 1$ [99]. They would therefore be biased tracers of the dark matter distribution. Note that the SHMR is not expected to change significantly up to $z \sim 0.3$, and so would the bias, according to this scenario. However, it is unclear if the BNSs originate from the isolated binary scenario based on observations of GW170817 (e.g., [8,100]). A preliminary distinction between different formation channels for BNS, and thus between different types of host galaxies as tracers of the dark matter distribution, will be possible in the nearby Universe with $\mathcal{O}(10)$ events using current generation GW detectors [101]. Precision measurements of the bias of GW BNS mergers at the redshift of interest for our work will be enabled by 3G detectors (e.g., [97]). In the future, once more information is available for a broader population of GW BNS mergers, another possibility would be to fit the bias along with the cosmological parameters. In the meantime, a better treatment of the GW bias parameter could follow the findings of [102], who argue that the redshift evolution of the bias is well described by a linear relation. We leave an analysis of this kind for future work.

Another interesting aspect to be explored in the future concerns the small scales of larger scale structure. There is in fact significant information at scales smaller than those considered in the projections of this article. We anticipate that continuing advances in the modeling of nonlinear modes will allow for virial cluster velocities to be included in the analysis model.

Note that the luminosity distances from GWs discussed in this work are estimated assuming GR. The method presented here aims at testing GR, rather than constraining different gravity models. A significant deviation of γ from the value expected in GR using this method, would mean that a modification of gravity exists. On the other hand, a result consistent with $\gamma = 0.55$ could still hide modifications of gravity due to the GR assumption in the GW distance estimates. In this case, inconsistencies in γ with precision measurements from EM-based distances would hint at such modification, since the luminosity distance in the GW measurement would have a different meaning (e.g., [80,103,104]). However, [105] shows that differences between GW and EM distances are at the subpercent level in the local Universe for popular gravity models, so we expect this effect to be negligible at the redshifts probed here.

We conclude that cosmology can largely benefit from the combination of dark energy experiments with gravitational wave data through a number of different probes, now including peculiar velocities, and that in the future we should seek for common strategies to maximize the scientific outcome from these synergies. In the coming

decades, a simultaneous measurement of the peculiar velocity field and the Hubble diagram using gravitational waves as standard candles, will fully exploit the cosmological information enclosed in the connection between gravitational wave sources and the large scale structure of the Universe.

ACKNOWLEDGMENTS

We thank Marica Branchesi, Zoheyr Doctor, Jan Harms, Cullan Howlett and Eric Linder for very useful discussion. Work supported by the Fermi National Accelerator Laboratory, managed and operated by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for U.S. Government purposes. This work is supported in part by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award No. DE-SC-0007867 and Contract No. DE-AC02-05CH11231.

APPENDIX A: CONTRAST BETWEEN OUR ANALYSIS AND THAT OF WANG *ET AL.*

Conservation of mass establishes a relationship between density and momentum fields through the continuity equation, which to first order gives

$$\begin{aligned} Haf\delta(\mathbf{x}) + \nabla\cdot\mathbf{v}(\mathbf{x}) &= 0 \\ Ha\beta\delta^g(\mathbf{x}) + \nabla\cdot\mathbf{v}(\mathbf{x}) &= 0. \end{aligned} \tag{A1}$$

[20] use this relationship between galaxy overdensity and velocity fields to determine $\beta = f/b$. Combining this with $\sigma_{8,g}$ as measured from the galaxy survey yields $f\sigma_8 = \beta\sigma_{8,g}$. Conceptually (though not exactly in practice), the model for the data includes β and a model for the underlying galaxy overdensity field. Equation (A1) is used to calculate the velocity field (taking care with the constant of integration). While the authors of [20] do not assign any error to this process, the similar analysis of [106] quadratically adds a 150 $\mathrm{km\,s^{-1}}$ uncorrelated uncertainty to each velocity measurement.

The approach in this article uses the above comparison of density and velocity fields and adds independent information from the anisotropy power spectrum at the time of the CMB. A local peculiar velocity survey does not cover the same volume of the CMB so a direct comparison of the density and velocity fields is not possible. Nevertheless, the correlations in matter underdensities in the different volumes should be consistent.

The diagonal terms of the matrix in Eq. (3) contain the autocorrelations in galaxy counts and peculiar velocities for a single k-mode measured in Fourier space as predicted from the anisotropy power spectrum at the time of the CMB. We include in the velocity shot-noise a 300 $\mathrm{km\,s^{-1}}$ term that represents sources of peculiar velocity that are not represented in our model, e.g., contributions from k-modes excluded in our calculations. It does not include errors that are in common between galaxy counts and velocities, for example differences between CAMB predictions and the true overdensity field today. The off-diagonal term gives the cross-correlation that accounts for the fact that galaxy counts and velocities arise from the same overdensity field. We do not add an error that corresponds to [106], but by limiting our calculations to $k_{\rm min} \le k \le k_{\rm max}$ we use less of the available data relative to their analysis, which uses all the information in real space.

APPENDIX B: RELATIONSHIPS BETWEEN THE PECULIAR VELOCITY AND PECULIAR MAGNITUDE POWER SPECTRA AND THE CMB MATTER POWER SPECTRUM

Predictions for the overdensity power spectrum are made by solving the Boltzmann equations. For our Fisher calculations, we do not recalculate this power spectrum as our model parameters change. While imperfect, this approach is motivated as follows. In linear theory the linear growth factor D is introduced to represent the time evolution of the density field. Its normalization is arbitrary, but a common practice is to normalize it to the CMB, such that $D(z_{\rm CMB})/a_{\rm CMB} = 1$. This choice is useful because it anchors D, independent of the parameters upon which it may depend (e.g., Ω_{M0}, γ), to a redshift with a precise measurement of $P_{\delta\delta}$. With this convention

$$P_{\delta\delta}(k;z,\gamma) = a_{\rm CMB}^{-2} D^2(z;\gamma) P_{\delta\delta}(k;z={\rm CMB}) \tag{B1}$$

and

$$P_{vv}(k,\mu;z,\gamma) = (\mu Haf)^2 k^{-2} P_{\delta\delta}(k;z,\gamma) \tag{B2}$$

$$= (\mu HfD)^2 \left(\frac{a}{a_{\rm CMB}}\right)^2 k^{-2} P_{\delta\delta}(k;z={\rm CMB}). \tag{B3}$$

For distance indicators it is useful to work with peculiar magnitudes [107], i.e., magnitude deviations from the background cosmological expansion. At low redshift and for small peculiar velocities the peculiar magnitude power spectrum is

$$P_{\delta m\delta m}(k,\mu;z,\gamma) \approx \left(\frac{5}{\ln 10}\right)^2 (\mu f)^2 (d_L k)^{-2} P_{\delta\delta}(k;z,\gamma). \tag{B4}$$

Going beyond linear theory, CAMB provides a more precise calculation of $P_{\delta\delta}$ using the CMB as an initial condition and standard gravity. Using the CAMB power spectrum as reference

$$P_{\delta m \delta m}(k,\mu;z,\gamma) = \tag{B5}$$

$$= \left(\frac{5}{\ln 10}\right)^2 (\mu f)^2 (d_L k)^{-2} \frac{P_{\delta\delta}(k;z,\gamma)}{P_{\delta\delta}^{\rm CAMB}(k;z=0)} P_{\delta\delta}^{\rm CAMB}(k;z=0). \tag{B6}$$

For the normalization term, taking $P_{\delta\delta}^{\rm CAMB}(k;z) \approx P_{\delta\delta}(k;z,\gamma=0.55)$ and $f=\Omega(z)^\gamma$

$$\frac{P_{\delta\delta}(k;z,\gamma)}{P_{\delta\delta}^{\rm CAMB}(k;z=0)} = \left(\frac{D(z;\gamma)}{D^{\rm CAMB}(z=0)}\right)^2 \tag{B7}$$

$$\approx \left(\frac{D(z;\gamma)}{D(z=0;\gamma=0.55)}\right)^2 \tag{B8}$$

$$= \exp\left[2\left(\int_{a_{\rm CMB}}^{a} \Omega_m^\gamma d\ln a - \int_{a_{\rm CMB}}^{1} \Omega_m^{0.55} d\ln a\right)\right]. \tag{B9}$$

For the special case of calculating the Fisher Matrix at $\gamma = 0.55$ the normalization of D is irrelevant as

$$\frac{P_{\delta\delta}(k;z,\gamma)}{P_{\delta\delta}^{\rm CAMB}(k;z=0)} \approx \exp\left(2\int_1^a \Omega_m^{0.55} d\ln a\right). \tag{B10}$$

For some applications, normalizing $D=1$ at $z=0$ is appropriate. Then

$$P_{\delta\delta}(k;z,\gamma) = D^2(z;\gamma) P_{\delta\delta}(k;z=0). \tag{B11}$$

For this choice of normalization, $P_{\delta\delta}(k;z={\rm CMB},\gamma)$ does depend on γ through $D^2(z={\rm CMB};\gamma)$.

APPENDIX C: EXPLANATION OF OUR ANALYSIS

Optimal estimation of the cosmology/gravity parameters would require the analysis of model predictions of primary observables, e.g., the GW waveforms. The standard procedure for using distance indicators for cosmological applications such as peculiar velocity and expansion history, however, proceeds as a two step process. In the first step, an estimate of the distance to the system is derived from the primary data. In SN Ia cosmology for example, light curves are fit to a SN light curve model and the fit parameters are combined to estimate a distance. In this application, the waveforms are fit to a model of the properties of the binary system, including the viewing angle, and the distance. In the second step, the estimated distance is used as a secondary data observable, derived from the primaries, that are used in the cosmology analysis. This is not optimal, in that it does not use the full suite of available data in the cosmology analysis. However, most of the important information is captured in the distance estimator. The challenge is then in the determination of the likelihood for the distance estimators, whose values are calculated through a highly nonlinear fitting process. In supernova cosmology it is assumed and validated through simulation, that the posterior for the distance derived from the light-curve fitting process is a good approximation for the likelihood of most probable distance (this assumption does break down when the primary data uncertainties are large). SN cosmology then proceeds using the best-fit distances as the data, and the posterior credible interval as their uncertainties, in the cosmology fit. We use the same procedure in this GW prediction.

[1] C. Gordon, K. Land, and A. Slosar, Cosmological Constraints from Type Ia Supernovae Peculiar Velocity Measurements, Phys. Rev. Lett. **99**, 081301 (2007).

[2] A. Abate and O. Lahav, The three faces of Ω_m: Testing gravity with low- and high-redshift SNe Ia surveys, Mon. Not. R. Astron. Soc. **389**, L47 (2008).

[3] D. Huterer, D. L. Shafer, D. M. Scolnic, and F. Schmidt, Testing ΛCDM at the lowest redshifts with SN Ia and galaxy velocities, J. Cosmol. Astropart. Phys. 5 (2017) 015.

[4] A. Kim, G. Aldering, P. Antilogus, A. Bahmanyar, S. BenZvi, H. Courtois, T. Davis, S. Ferraro, S. G. A. Gontcho, O. Graur *et al.*, Testing gravity using type Ia supernovae discovered by next-generation wide-field imaging surveys, Bull. Am. Astron. Soc. **51**, 140 (2019), arXiv:1903.07652.

[5] A. G. Kim and E. V. Linder, Complementarity of peculiar velocity surveys and redshift space distortions for testing gravity, Phys. Rev. D **101**, 023516 (2020).

[6] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, C. Affeldt *et al.* (LIGO Scientific and Virgo Collaborations), Gw170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral, Phys. Rev. Lett. **119**, 161101 (2017).

[7] LIGO Scientific, Virgo Collaborations, F. GBM, INTEGRAL, IceCube Collaboration, AstroSat Cadmium Zinc Telluride Imager Team, IPN Collaboration, The Insight-Hxmt Collaboration, ANTARES Collaboration, The Swift Collaboration *et al.*, Multi-messenger observations of a binary neutron star merger, Astrophys. J. Lett. **848**, L12 (2017).

[8] A. Palmese, W. Hartley, F. Tarsitano, C. Conselice, O. Lahav, S. Allam, J. Annis, H. Lin, M. Soares-Santos, D. Tucker *et al.*, Evidence for dynamically driven formation of the GW170817 neutron star binary in NGC 4993, Astrophys. J. Lett. **849**, L34 (2017).

[9] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya *et al.*, A gravitational-wave standard siren measurement of the Hubble constant, Nature (London) **551**, 85 (2017).

[10] M. Soares-Santos, A. Palmese, W. Hartley, J. Annis, J. Garcia-Bellido, O. Lahav, Z. Doctor, M. Fishbach, D. E. Holz, H. Lin *et al.*, First Measurement of the Hubble Constant from a Dark Standard Siren using the Dark Energy Survey Galaxies and the LIGO/Virgo Binary-Black-hole Merger GW170814, Astrophys. J. Lett. **876**, L7 (2019).

[11] A. Palmese, J. deVicente, M. E. S. Pereira, J. Annis, W. Hartley, K. Herner, M. Soares-Santos, M. Crocce, D. Huterer, I. Magaña Hernandez *et al.*, A statistical standard Siren measurement of the Hubble constant from the LIGO/Virgo gravitational wave compact object merger GW190814 and dark energy survey galaxies, Astrophys. J. Lett. **900**, L33 (2020).

[12] B. F. Schutz, Determining the Hubble constant from gravitational wave observations, Nature (London) **323**, 310 (1986).

[13] C. Howlett and T. M. Davis, Standard siren speeds: Improving velocities in gravitational-wave measurements of H_0, Mon. Not. R. Astron. Soc. **492**, 3803 (2020).

[14] C. Nicolaou, O. Lahav, P. Lemos, W. Hartley, and J. Braden, The impact of peculiar velocities on the estimation of the hubble constant from gravitational wave standard sirens, Mon. Not. R. Astron. Soc. **495**, 90 (2020).

[15] C. Cutler and D. E. Holz, Ultrahigh precision cosmology from gravitational waves, Phys. Rev. D **80**, 104009 (2009).

[16] C. M. Hirata, D. E. Holz, and C. Cutler, Reducing the weak lensing noise for the gravitational wave hubble diagram using the non-gaussianity of the magnification distribution, Phys. Rev. D **81**, 124046 (2010).

[17] C. Shang and Z. Haiman, Cosmology with standard sirens: The importance of the shape of the lensing magnification distribution, Mon. Not. R. Astron. Soc. **411**, 9 (2011).

[18] S. Camera and A. Nishizawa, Beyond Concordance Cosmology with Magnification of Gravitational-Wave Standard Sirens, Phys. Rev. Lett. **110**, 151103 (2013).

[19] P. Zhang, The large scale structure in the 3d luminosity-distance space and its cosmological applications, arXiv:1810.11915.

[20] Y. Y. Wang, F. Y. Wang, and Y. C. Zou, Precise peculiar velocities from gravitational waves accompanied by electromagnetic signals and cosmological applications, Phys. Rev. D **98**, 063503 (2018).

[21] M. Punturo, M. Abernathy, F. Acernese, B. Allen, N. Andersson, K. Arun, F. Barone, B. Barr, M. Barsuglia, M. Beker, N. Beveridge *et al.*, The Einstein telescope: A third-generation gravitational wave observatory, Classical Quant. Grav. **27**, 194002 (2010).

[22] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, K. Ackley, C. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, C. Affeldt *et al.*, Exploring the sensitivity of next generation gravitational wave detectors, Classical Quant. Grav. **34**, 044001 (2017).

[23] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari *et al.*, Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA, Living Rev. Relativity **21**, 3 (2018).

[24] B. S. Sathyaprakash, B. F. Schutz, and C. Van Den Broeck, Cosmography with the Einstein telescope, Classical Quant. Grav. **27**, 215006 (2010).

[25] W. Zhao, C. Van Den Broeck, D. Baskaran, and T. G. F. Li, Determination of dark energy by the Einstein telescope: Comparing with Cmb, Bao, and Snia observations, Phys. Rev. D **83**, 023005 (2011).

[26] R.-G. Cai and T. Yang, Estimating cosmological parameters by the simulated data of gravitational waves from the einstein telescope, Phys. Rev. D **95**, 044024 (2017).

[27] W. Zhao and L. Wen, Localization accuracy of compact binary coalescences detected by the third-generation gravitational-wave detectors and implication for cosmology, Phys. Rev. D **97**, 064031 (2018).

[28] R. E. Keeley, A. Shafieloo, B. L'Huillier, and E. V. Linder, Debiasing cosmic gravitational wave sirens, Mon. Not. R. Astron. Soc. **491**, 3983 (2020).

[29] A. Shafieloo, R. E. Keeley, and E. V. Linder, Will cosmic gravitational wave sirens determine the hubble constant?, J. Cosmol. Astropart. Phys. 03 (2020) 019.

[30] S. Kawamura, T. Nakamura, M. Ando, N. Seto, K. Tsubono, K. Numata, R. Takahashi, S. Nagano, T. Ishikawa, M. Musha *et al.*, The Japanese space gravitational wave antenna—DECIGO, Classical Quant. Grav. **23**, S125 (2006).

[31] R. Takahashi and T. Nakamura, The decihertz laser interferometer can determine the position of the coalescing binary neutron stars within an arcminute a week before the final merging event to the black hole, Astrophys. J. **596**, L231 (2003).

[32] R. Chornock, P. S. Cowperthwaite, R. Margutti, D. Milisavljevic, K. D. Alexander, I. Andreoni, I. Arcavi, A. Baldeschi, J. Barnes, E. Bellm *et al.*, Multi-messenger astronomy with extremely large telescopes, arXiv:1903.04629.

[33] A. Palmese, O. Graur, J. T. Annis, S. BenZvi, E. Di Valentino, J. Garcia-Bellido, S. G. A. Gontcho, R. Keeley, A. Kim, O. Lahav, S. Nissanke, K. Paterson, M. Sako, A. Shafieloo, and Y.-D. Tsai, Gravitational wave cosmology and astrophysics with large spectroscopic galaxy surveys, Bull. Am. Astron. Soc. **51**, 310 (2019), https://baas.aas.org/pub/2020n3i310/release/1.

[34] D. Scolnic, R. Kessler, D. Brout, P. S. Cowperthwaite, M. Soares-Santos, J. Annis, K. Herner, H.-Y. Chen, M. Sako, Z. Doctor *et al.*, How many kilonovae can be found in past, present, and future survey data sets?, Astrophys. J. **852**, L3 (2017).

[35] DESI Collaboration, The DESI Experiment Part I: Science, targeting, and survey design, arXiv:1611.00036.

[36] E. Swann, M. Sullivan, J. Carrick, S. Hoenig, I. Hook, R. Kotak, K. Maguire, R. McMahon, R. Nichol, and S. Smartt, 4MOST Consortium Survey 10: The time-domain extragalactic survey (TiDES), Messenger **175**, 58 (2019).

[37] E. da Cunha, A. M. Hopkins, M. Colless, E. N. Taylor, C. Blake, C. Howlett, C. Magoulas, J. R. Lucey, C. Lagos, K. Kuehn *et al.*, The taipan galaxy survey: Scientific goals and observing strategy, Pub. Astron. Soc. Aust. **34**, e047 (2017).

[38] B. D. Metzger, G. Martínez-Pinedo, S. Darbha, E. Quataert, A. Arcones, D. Kasen, R. Thomas, P. Nugent, I. V. Panov, and N. T. Zinner, Electromagnetic counterparts of compact object mergers powered by the radioactive decay of r-process nuclei, Mon. Not. R. Astron. Soc. **406**, 2650 (2010).

[39] J. Barnes and D. Kasen, Effect of a high opacity on the light curves of radioactively powered transients from compact object mergers, Astrophys. J. **775**, 18 (2013).

[40] The LIGO Scientific and the Virgo Collaborations, GW190425: Observation of a compact binary coalescence with total mass $\sim 3.4\ M_{\odot}$, Astrophys. J. Lett. **892**, L3 (2020).

[41] B. Dilday, M. Smith, B. Bassett, A. Becker, R. Bender *et al.*, Measurements of the rate of type Ia supernovae at redshift $\lesssim 0.3$ from the Sloan digital sky survey II supernova survey, Astrophys. J. **713**, 1026 (2010).

[42] M. Betoule, R. Kessler, J. Guy, J. Mosher, D. Hardin, R. Biswas, P. Astier, P. El-Hage, M. Konig, S. Kuhlmann *et al.*, Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples, Astron. Astrophys. **568**, A22 (2014).

[43] H.-Y. Chen, S. Vitale, and R. Narayan, Viewing Angle of Binary Neutron Star Mergers, Phys. Rev. X **9**, 031028 (2019).

[44] C. Guidorzi, R. Margutti, D. Brout, D. Scolnic, W. Fong, K. D. Alexander, P. S. Cowperthwaite, J. Annis, E. Berger, P. K. Blanchard *et al.*, Improved constraints on h 0 from a combined analysis of gravitational-wave and electromagnetic emission from gw170817, Astrophys. J. **851**, L36 (2017).

[45] K. Hotokezaka, E. Nakar, O. Gottlieb, S. Nissanke, K. Masuda, G. Hallinan, K. P. Mooley, and A. T. Deller, A Hubble constant measurement from superluminal motion of the jet in GW170817, Nat. Astron. **3**, 940 (2019).

[46] D. Dobie, D. L. Kaplan, K. Hotokezaka, T. Murphy, A. Deller, G. Hallinan, and S. Nissanke, Constraining properties of neutron star merger outflows with radio observations, Mon. Not. R. Astron. Soc. **494**, 2449 (2020).

[47] S. Dhawan, M. Bulla, A. Goobar, A. Sagués Carracedo, and C. N. Setzer, Constraining the observer angle of the kilonova AT2017gfo associated with GW170817: Implications for the Hubble constant, Astrophys. J. **888**, 67 (2020).

[48] C. Howlett, L. Staveley-Smith, P. J. Elahi, T. Hong, T. H. Jarrett, D. H. Jones, B. S. Koribalski, L. M. Macri, K. L. Masters, and C. M. Springob, 2MTF—VI. Measuring the velocity power spectrum, Mon. Not. R. Astron. Soc. **471**, 3135 (2017).

[49] C. Howlett, L. Staveley-Smith, and C. Blake, Cosmological forecasts for combined and next-generation peculiar velocity surveys, Mon. Not. R. Astron. Soc. **464**, 2517 (2017).

[50] E. V. Linder, Cosmic growth history and expansion history, Phys. Rev. D **72**, 043529 (2005).

[51] E. V. Linder and R. N. Cahn, Parameterized beyond-Einstein growth, Astropart. Phys. **28**, 481 (2007).

[52] D. Huterer, D. Kirkby, R. Bean, A. Connolly, K. Dawson, S. Dodelson, A. Evrard, B. Jain, M. Jarvis, E. Linder, R. Mandelbaum, M. May, A. Raccanelli, B. Reid, E. Rozo, F. Schmidt, N. Sehgal, A. Slosar, A. van Engelen, H.-Y. Wu, and G. Zhao, Growth of cosmic structure: Probing dark energy beyond expansion, Astropart. Phys. **63**, 23 (2015), dark Energy and CMB.

[53] N. Kaiser, Clustering in real space and in redshift space, Mon. Not. R. Astron. Soc. **227**, 1 (1987).

[54] Y.-S. Song and W. J. Percival, Reconstructing the history of structure formation using redshift distortions, J. Cosmol. Astropart. Phys. 10 (2009) 004.

[55] T. Namikawa, Analyzing clustering of astrophysical gravitational-wave sources: Luminosity-distance space distortions, J. Cosmol. Astropart. Phys. 01 (2021) 036.

[56] C. Howlett, A. S. G. Robotham, C. D. P. Lagos, and A. G. Kim, Measuring the growth rate of structure with type ia supernovae from LSST, Astrophys. J. **847**, 128 (2017).

[57] A. Abate, S. Bridle, L. F. A. Teodoro, M. S. Warren, and M. Hendry, Peculiar velocities into the next generation: Cosmological parameters from large surveys without bias from non-linear structure, Mon. Not. R. Astron. Soc. **389**, 1739 (2008).

[58] C. Adams and C. Blake, Improving constraints on the growth rate of structure by modelling the density-velocity cross-correlation in the 6dF Galaxy Survey, Mon. Not. R. Astron. Soc. **471**, 839 (2017).

[59] A. Lewis and S. Bridle, Cosmological parameters from CMB and other data: A Monte Carlo approach, Phys. Rev. D **66**, 103511 (2002).

[60] N. A. Bahcall and S. P. Oh, The peculiar velocity function of galaxy clusters, Astrophys. J. Lett. **462**, L49 (1996).

[61] D. A. Dale, R. Giovanelli, M. P. Haynes, L. E. Campusano, and E. Hardy, Seeking the local convergence depth. V. Tully-Fisher peculiar velocities for 52 Abell clusters, Astron. J. **118**, 1489 (1999).

[62] K. L. Masters, C. M. Springob, M. P. Haynes, and R. Giovanelli, SFI++ I: A new I-band Tully-Fisher template, the cluster peculiar velocity dispersion, and H_0, Astrophys. J. **653**, 861 (2006).

[63] D. J. Mortlock, S. M. Feeney, H. V. Peiris, A. R. Williamson, and S. M. Nissanke, Unbiased Hubble constant estimation from binary neutron star mergers, Phys. Rev. D **100**, 103523 (2019).

[64] F. Beutler, C. Blake, M. Colless, D. H. Jones, L. Staveley-Smith, G. B. Poole, L. Campbell, Q. Parker, W. Saunders, and F. Watson, The 6dF Galaxy Survey: $z \approx 0$ measurements of the growth rate and σ_8, Mon. Not. R. Astron. Soc. **423**, 3430 (2012).

[65] C. Blake, S. Brough, M. Colless, C. Contreras, W. Couch, S. Croom, T. Davis, M. J. Drinkwater, K. Forster, D. Gilbank, M. Gladders, K. Glazebrook, B. Jelliffe, R. J. Jurek, I. H. Li, B. Madore, D. C. Martin, K. Pimbblet, G. B. Poole, M. Pracy, R. Sharp, E. Wisnioski, D. Woods, T. K. Wyder, and H. K. C. Yee, The WiggleZ dark energy survey: The growth rate of cosmic structure since redshift $z = 0.9$, Mon. Not. R. Astron. Soc. **415**, 2876 (2011).

[66] A. Oka, S. Saito, T. Nishimichi, A. Taruya, and K. Yamamoto, Simultaneous constraints on the growth of structure and cosmic expansion from the multipole power spectra of the SDSS DR7 LRG sample, Mon. Not. R. Astron. Soc. **439**, 2515 (2014).

[67] C. Howlett, A. J. Ross, L. Samushia, W. J. Percival, and M. Manera, The clustering of the SDSS main galaxy sample—II. Mock galaxy catalogues and a measurement of the growth of structure from redshift space distortions at $z = 0.15$, Mon. Not. R. Astron. Soc. **449**, 848 (2015).

[68] S. Satpathy, S. Alam, S. Ho, M. White, N. A. Bahcall, F. Beutler, J. R. Brownstein, C.-H. Chuang, D. J. Eisenstein, J. N. Grieb, F. Kitaura, M. D. Olmstead, W. J. Percival, S. Salazar-Albornoz, A. G. Sánchez, H.-J. Seo, D. Thomas, J. L. Tinker, and R. Tojeiro, The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: On the measurement of growth rate using galaxy correlation functions, Mon. Not. R. Astron. Soc. **469**, 1369 (2017).

[69] S. de la Torre *et al.*, The VIMOS Public Extragalactic Redshift Survey (VIPERS). Gravity test from the combination of redshift-space distortions and galaxy-galaxy lensing at $0.5 < z < 1.2$, Astron. Astrophys. **608**, A44 (2017).

[70] M. Icaza-Lizaola, M. Vargas-Magaña, S. Fromenteau, S. Alam, B. Camacho, H. Gil-Marin, R. Paviot, A. Ross, D. P. Schneider, J. Tinker, Y. Wang, C. Zhao, A. Prakash, G. Rossi, G.-B. Zao, I. Cruz-Gonzalez, and A. de la Macorra, The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 LRG sample: Structure growth rate measurement from the anisotropic LRG correlation function in the redshift range $0.6 < z < 1.0$, Mon. Not. R. Astron. Soc. **492**, 4189 (2020).

[71] S. Johnston *et al.*, Science with ASKAP. The Australian square-kilometre-array pathfinder, Exp. Astron. **22**, 151 (2008).

[72] W. Saunders, S. Oliver, O. Keeble, M. Rowan-Robinson, A. Canavezes, S. Maddox, W. Sutherland, G. Efstathiou, R. McMahon, V. Springel, S. White, H. Tadros, C. Frenk, E. Branchini, A. Taylor, B. Ballinger, and A. Heavens, First results from the PSC-z survey, in *Wide Field Surveys in Cosmology*, Vol. 14, edited by S. Colombi, Y. Mellier, and B. Raban (Editions Frontieres, 1998), p. 71, https://inspirehep.net/literature/1637505.

[73] E. Branchini, L. Teodoro, C. S. Frenk, I. Schmoldt, G. Efstathiou, S. D. M. White, W. Saunders, W. Sutherland, M. Rowan-Robinson, O. Keeble, H. Tadros, S. Maddox, and S. Oliver, A non-parametric model for the cosmic velocity field, Mon. Not. R. Astron. Soc. **308**, 1 (1999).

[74] H.-Y. Chen, M. Fishbach, and D. E. Holz, A two per cent Hubble constant measurement from standard sirens within five years, Nature (London) **562**, 545 (2018).

[75] N. Yunes and X. Siemens, Gravitational-wave tests of general relativity with ground-based detectors and pulsar-timing arrays, Living Rev. Relativity **16**, 9 (2013).

[76] Z. Carson and K. Yagi, Testing general relativity with gravitational waves, arXiv:2011.02938.

[77] T. Baker, E. Bellini, P. Ferreira, M. Lagos, J. Noller, and I. Sawicki, Strong Constraints on Cosmological Gravity from GW170817 and GRB170817a, Phys. Rev. Lett. **119**, 251301 (2017).

[78] J. M. Ezquiaga and M. ZumalacÃ¡rregui, Dark Energy After GW170817: Dead Ends and the Road Ahead, Phys. Rev. Lett. **119**, 251304 (2017).

[79] M. Lagos, M. Fishbach, P. Landry, and D. E. Holz, Standard sirens with a running Planck mass, Phys. Rev. D **99**, 083504 (2019).

[80] I. D. Saltas, I. Sawicki, L. Amendola, and M. Kunz, Anisotropic Stress as a Signature of Nonstandard Propagation of Gravitational Waves, Phys. Rev. Lett. **113**, 191101 (2014).

[81] J. M. Ezquiaga and M. ZumalacÃ¡rregui, Gravitational wave lensing beyond general relativity: Birefringence, echoes and shadows, Phys. Rev. D **102**, 124048 (2020).

[82] T. Baker and I. Harrison, Constraining scalar-tensor modified gravity with gravitational waves and large scale structure surveys, J. Cosmol. Astropart. Phys. 01 (2021) 068.

[83] B. D. Metzger and E. Berger, What is the most promising electromagnetic counterpart of a neutron star binary merger?, Astrophys. J. **746**, 48 (2012).

[84] D. Kasen, B. Metzger, J. Barnes, E. Quataert, and E. Ramirez-Ruiz, Origin of the heavy elements in binary neutron-star mergers from a gravitational-wave event, Nature (London) **551**, 80 (2017).

[85] S. Vitale and H.-Y. Chen, Measuring the Hubble Constant with Neutron Star Black Hole Mergers, Phys. Rev. Lett. **121**, 021303 (2018).

[86] 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. R. Astron. Soc. **425**, 460 (2012).

[87] B. McKernan, K. E. S. Ford, I. Bartos, M. J. Graham, W. Lyra, S. Marka, Z. Marka, N. P. Ross, D. Stern, and Y. Yang, Ram-pressure stripping of a kicked Hill sphere: Prompt electromagnetic emission from the merger of stellar mass black holes in an AGN accretion disk, Astrophys. J. Lett. **884**, L50 (2019).

[88] C. J. Conselice, R. Bhatawdekar, A. Palmese, and W. G. Hartley, Gravitational waves from black holes in merging ultra-dwarf galaxies, Astrophys. J. **890**, 8 (2020).

[89] A. Palmese and C. J. Conselice, GW190521 from the merger of ultra-dwarf galaxies, arXiv:2009.10688.

[90] H.-Y. Chen and D. E. Holz, Finding the one: Identifying the host galaxies of gravitational-wave sources, arXiv:1612.01471.

[91] A. E. Nugent, W.-f. Fong, Y. Dong, A. Palmese, J. Leja, A. Rouco Escorial, P. K. Blanchard, K. Paterson, R. Chornock, A. Monson, M. Nicholl, and E. Berger, The distant, galaxy cluster environment of the short GRB 161104A at $z \sim 0.8$ and a comparison to the short GRB host population, Astrophys. J. **904**, 52 (2020), https://iopscience.iop.org/article/10.3847/1538-4357/abc24a/pdf.

[92] J. N. Fry, The evolution of bias, Astrophys. J. Lett. **461**, L65 (1996), https://iopscience.iop.org/article/10.1086/310006/pdf.

[93] M. Rahman, A. J. Mendez, B. Ménard, R. Scranton, S. J. Schmidt, C. B. Morrison, and T. Budavári, Exploring the SDSS photometric galaxies with clustering redshifts, Mon. Not. R. Astron. Soc. **460**, 163 (2016).

[94] M. Crocce, J. Carretero, A. H. Bauer, A. J. Ross, I. Sevilla-Noarbe, T. Giannantonio, F. Sobreira, J. Sanchez, E. Gaztanaga, M. C. Kind *et al.*, Galaxy clustering, photometric redshifts and diagnosis of systematics in the des

science verification data, Mon. Not. R. Astron. Soc. **455**, 4301 (2016).
[95] J. Elvin-Poole, M. Crocce, A. Ross, T. Giannantonio, E. Rozo, E. Rykoff, S. Avila, N. Banik, J. Blazek, S. Bridle *et al.*, Dark energy survey year 1 results: Galaxy clustering for combined probes, Phys. Rev. D **98**, 042006 (2018).
[96] G. Scelfo, N. Bellomo, A. Raccanelli, S. Matarrese, and L. Verde, GWxLSS: Chasing the progenitors of merging binary black holes, J. Cosmol. Astropart. Phys. 09 (2018) 039.
[97] F. Calore, A. Cuoco, T. Regimbau, S. Sachdev, and P. D. Serpico, Cross-correlating galaxy catalogs and gravitational waves: A tomographic approach, Phys. Rev. Research **2**, 023314 (2020).
[98] R. H. Wechsler and J. L. Tinker, The connection between galaxies and their dark matter halos, Annu. Rev. Astron. Astrophys. **56**, 435 (2018).
[99] P. Behroozi, R. H. Wechsler, A. P. Hearin, and C. Conroy, UniverseMachine: The correlation between galaxy growth and dark matter halo assembly from $z = 0$–10, Mon. Not. R. Astron. Soc. **488**, 3143 (2019).
[100] Y.-D. Tsai, A. Palmese, S. Profumo, and T. Jeltema, Is gw170817 a multimessenger neutron star-primordial black hole merger?, arXiv:2007.03686.
[101] S. Adhikari, M. Fishbach, D. E. Holz, R. H. Wechsler, and Z. Fang, The binary-host connection: Astrophysics of gravitational wave binaries from their host galaxy properties, Astrophys. J. **905**, 21 (2020).
[102] S. Libanore, M. C. Artale, D. Karagiannis, M. Liguori, N. Bartolo, Y. Bouffanais, N. Giacobbo, M. Mapelli, and S. Matarrese, Gravitational wave mergers as tracers of large scale structures, J. Cosmol. Astropart. Phys. 02 (2021) 035.
[103] A. Nishizawa, Generalized framework for testing gravity with gravitational-wave propagation. i. Formulation, Phys. Rev. D **97**, 104037 (2018).
[104] E. Belgacem, Y. Dirian, S. Foffa, and M. Maggiore, Gravitational-wave luminosity distance in modified gravity theories, Phys. Rev. D **97**, 104066 (2018).
[105] E. V. Linder, Limited modified gravity, J. Cosmol. Astropart. Phys. 10 (2020) 042.
[106] S. S. Boruah, M. J. Hudson, and G. Lavaux, Cosmic flows in the nearby Universe: New peculiar velocities from SNe and cosmological constraints, Mon. Not. R. Astron. Soc. **498**, 2703 (2020).
[107] L. Hui and P. B. Greene, Correlated fluctuations in luminosity distance and the importance of peculiar motion in supernova surveys, Phys. Rev. D **73**, 123526 (2006).