



Universiteit
Leiden
The Netherlands

Prospects of testing late-time cosmology with weak lensing of gravitational waves and galaxy surveys

Balardo, A.; Garoffolo, A.; Martinelli, M.; Mukherjee, S.; Silvestri, A.

Citation

Balardo, A., Garoffolo, A., Martinelli, M., Mukherjee, S., & Silvestri, A. (2023). Prospects of testing late-time cosmology with weak lensing of gravitational waves and galaxy surveys. *Journal Of Cosmology And Astroparticle Physics*, 2023(6).
doi:10.1088/1475-7516/2023/06/050

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/3718826>

Note: To cite this publication please use the final published version (if applicable).

PAPER

Prospects of testing late-time cosmology with weak lensing of gravitational waves and galaxy surveys

To cite this article: Anna Balaudo *et al* JCAP06(2023)050

View the [article online](#) for updates and enhancements.

You may also like

- [IDENTIFYING ELUSIVE ELECTROMAGNETIC COUNTERPARTS TO GRAVITATIONAL WAVE MERGERS: AN END-TO-END SIMULATION](#)
Samaya Nissanke, Mansi Kasliwal and Alexandra Georgieva
- [Crack propagation monitoring in a full-scale aircraft fatigue test based on guided wave-Gaussian mixture model](#)
Lei Qiu, Shenfang Yuan, Qiao Bao et al.
- [HIGH-ENERGY ELECTROMAGNETIC OFFLINE FOLLOW-UP OF LIGO-VIRGO GRAVITATIONAL-WAVE BINARY COALESCENCE CANDIDATE EVENTS](#)
L. Blackburn, M. S. Briggs, J. Camp et al.

Prospects of testing late-time cosmology with weak lensing of gravitational waves and galaxy surveys

Anna Balaudo ^{a,b}, Alice Garoffolo ^b, Matteo Martinelli ^{c,d},
Suvodip Mukherjee ^e and Alessandra Silvestri ^b

^aLeiden Observatory, Leiden University, PO Box 9513, Leiden 2300 RA, The Netherlands

^bInstitute Lorentz, Leiden University, PO Box 9506, Leiden 2300 RA, The Netherlands

^cINAF-Osservatorio Astronomico di Roma,

Via Frascati 33, I-00078 Monteporzio Catone, Italy

^dINFN — Sezione di Roma — c/o Dipartimento di Fisica,

Piazzale Aldo Moro 2, Edificio G. Marconi, I-00185 Roma, Italy

^eDepartment of Astronomy & Astrophysics, Tata Institute of Fundamental Research,
1 Homi Bhabha Road, Colaba, Mumbai 400005, India

E-mail: balaudo@strw.leidenuniv.nl, garoffolo@lorentz.leidenuniv.nl,

matteo.martinelli@inaf.it, suvodip@tifr.res.in, silvestri@lorentz.leidenuniv.nl

Received October 14, 2022

Revised April 5, 2023

Accepted May 29, 2023

Published June 21, 2023

Abstract. We investigate the synergy of upcoming galaxy surveys and gravitational wave (GW) experiments in constraining late-time cosmology, examining the cross-correlations between the weak lensing of gravitational waves (GW-WL) and the galaxy fields. Without focusing on any specific GW detector configuration, we benchmark the requirements for the high precision measurement of cosmological parameters by considering several scenarios, varying the number of detected GW events and the uncertainty on the inference of the source luminosity distance and redshift. We focus on Λ CDM and scalar-tensor cosmologies, using the Effective Field Theory formalism as a unifying language. We find that, in some of the explored setups, GW-WL contributes to the galaxy signal by doubling the accuracy on non- Λ CDM parameters, allowing in the most favourable scenarios to reach even percent and sub-percent level bounds. Though the most extreme cases presented here are likely beyond the observational capabilities of currently planned individual GW detectors, we show nonetheless that — provided that enough statistics of events can be accumulated — GW-WL offers the potential to become a cosmological probe complementary to LSS surveys, particularly for those parameters that cannot be constrained by other GW probes such as standard sirens.

Keywords: cosmological parameters from LSS, gravitational lensing, Gravitational waves in GR and beyond: theory, gravitational waves/experiments

ArXiv ePrint: [2210.06398](https://arxiv.org/abs/2210.06398)

Contents

1	Introduction	1
2	Gravitational waves observation	3
3	Theoretical framework	5
3.1	Luminosity distance in scalar-tensor theories	6
3.2	Lensing convergence estimators for gravitational waves	7
4	Tomographic observables	8
5	Models, analysis method and surveys specifications	10
5.1	Models	10
5.2	Analysis method	13
5.3	Galaxy and GW surveys	14
6	Results	16
6.1	Λ CDM	17
6.2	Scalar-tensor models	20
7	Discussion and conclusions	26
A	Impact of GW binning	28
B	Impact of GW sources redshift distribution	29

1 Introduction

The historical direct detection of a gravitational wave (GW) by the LIGO-Virgo collaboration in 2015 [1], marked the beginning of a series of observing campaigns that led to the detection of around one hundred sources [2–4], one of which (GW170817) with a confirmed electromagnetic counterpart [5–7]. In the near future, the KAGRA interferometer [8] will join LIGO-Virgo in their observing runs, while the space interferometer LISA [9] is expected to launch in the late ’30s. Further down the lane, the third generation of ground-based GW detectors will see light with the network of Einstein Telescope (ET) [10] and Cosmic Explorer (CE) [11], which can drastically improve the sensitivity to GW signals and measure hundreds of thousands of events over 10 years of observations. While GW170817 led the way, allowing the first GW measurement of the Hubble constant [12], multi-band GW observations will effectively open a new window in observational cosmology, targeting several classes of sources among which stellar black holes binaries (BBH), neutron star-black hole systems (NS-BH), binary neutron stars (BNS), massive black hole binaries (MBHB), extreme mass-ratio inspirals and a stochastic gravitational wave background.

Meanwhile, the intense worldwide effort towards mapping the Universe via wide galaxy and weak lensing surveys has started to deliver large scale structure (LSS) data of unprecedented precision and extension, such as those provided by KiDS [13, 14] and DES [15, 16]

collaborations. With Stage IV missions (such as *Euclid* [17], the Vera C. Rubin Observatory [18–20] and Nancy Grace Roman Space Telescope [21]) we will see a paradigm shift in the volume of data available. This will lead to a new level of scrutiny of the standard model of cosmology, Lambda Cold Dark Matter (Λ CDM).

This model successfully describes the Universe in terms of few parameters, yet the rise of precision cosmology has seen the emergence of some tensions between datasets when interpreted within it [22, 23], which could simply signal the first cracks in Λ CDM as we are achieving a new level of precision in the measurement of its parameters [24].

With the ongoing/upcoming cosmological surveys, we have the opportunity to probe gravity on cosmological scales, and shed light on the nature of dark energy, sourcing cosmic acceleration, and dark matter, supporting gravitational clustering. In this context, many alternatives to the Λ CDM model have been proposed, trying to fill the gaps in the interpretation of the available datasets. We focus on extensions of Λ CDM that address the phenomenon of cosmic acceleration by means of a dynamical dark energy (DE) component or modifying the laws of gravity on large scales [25, 26]. Effectively adding one scalar degree of freedom to the gravitational theory, we adopt the formalism of effective field theory of dark energy (EFT of DE) [27–32] to explore the class of Horndeski gravity [33], which will undergo a high degree of scrutiny with the upcoming cosmological missions. We refer to these models with the common designation of *scalar-tensor theories* (ST), or simply *modified gravity* (MG); popular examples are $f(R)$ gravity, generalized Brans-Dicke models and Generalized Galileons [26].

Wide surveys that combine galaxy clustering (GC) and weak lensing (WL), such as the ESA-led *Euclid* mission [34] and the Vera Rubin Observatory [35], are expected to test gravity on cosmological scales at unprecedented precision, thus providing important constraints on ST theories, see e.g. [36]. In order to further tighten such constraints, and better understand the nature of DE and the theory of gravity, it will be crucial to bring in new probes able to provide independent measurements and also break degeneracies with the cosmological parameters. Multi-band GW observations are very interesting to this extent, as they offer complementary probes that are characterized by a different set of systematic errors with respect to galaxy surveys.

The dynamics of GWs depends on the extended parameters on two level: explicitly, as they enter directly their propagation equation, and implicitly via the different expansion rate of the Universe and growth of gravitational potentials. In this work, we consider gravitational theories in which photons are not directly coupled to the additional scalar field, which is a typical assumption also in the EFT of DE description. As a result, the dynamics of photons is affected only implicitly, while gravitational waves can display an explicitly modified behavior due to the direct couplings between the metric and the scalar field, which results in a modification of the equations for the propagation of tensors. This opens the possibility of gaining constraining power on DE/MG through probes that combine electromagnetic and GW sources. One example of such breaking of degeneracies lies in the cosmological friction term which enters the equation for propagation of photons and GWs on cosmological background; in ST theories, this term is modified by a factor depending on the conformal coupling only for GWs, while it remains unaffected for photons. This results in a difference between the luminosity distance inferred from electromagnetic sources and the one inferred from the amplitude of GWs (see section 3) [37–39] and can be used to test theories of gravity as shown in [37, 40–47].

In this paper, we focus on the potentialities of weak lensing of GWs (GW-WL). For high enough redshift sources ($z \gtrsim 0.5$), GW-WL can cause up to $\sim 5 - 10\%$ distortion in the GW strain, posing a serious limitation to the precision with which the luminosity distance

of the sources is measured [48, 49]. However, it can also be exploited as an interesting signal [50], especially when cross-correlated with galaxies [51–53], CMB lensing [54], HI intensity mapping [55, 56] and other GW probes [57, 58]. Cross-correlation of the GW sources with galaxies is shown to be a powerful technique for several other science topics ranging from cosmology and astrophysics [55, 59–63] to its recent application to GWTC-3 [64]. Building on the relativistic corrections to the luminosity distance of GWs in MG, we introduce a lensing convergence estimator for GWs which contains explicitly a contribution from the conformal coupling characteristic of ST theories, thus maximizing the constraining power on cosmological parameters when cross correlating GW-WL with the galaxy fields (GC,WL).

As we will discuss in our analysis, an accurate measurement of the GW-WL signal will require the detection of numerous GW events, as well as high precision in the determination of the GW source luminosity distance and in the inference of its redshift. We choose to focus this work on bright GW sources, i.e. sources for which the redshift is known with high accuracy, and we consider separately the cases in which the source redshift is observed photometrically or spectroscopically. Additionally, we decide not to specialize to a given GW detector with its noise and sensitivity. Rather, we investigate the requirements on both the number of GW sources and the tolerance in the luminosity distance measurement uncertainty for GW-WL to contribute significantly to the GC+WL signal, improving on the constraints on gravity that will become available with upcoming galaxy surveys. We then compare our benchmarks with the current observational expectations for planned GW detectors.

We show that the information contained in GW-WL and its cross-correlations with the galaxies can push cosmological bounds significantly beyond the capabilities of galaxy surveys, but only if high precision measurement of distances are possible for bright sirens, in addition to the detection of a consistent statistics of events. Though we have limited our work to bright GW events and its synergies with galaxies, combination with other EM probes can also be explored, and can further improve the measurement of beyond-GR effects (see e.g. [54]).

The paper is organized as follows: in section 2 we shortly review the principal features of GW observations; in section 3 we setup our theoretical framework and define the estimator that can be used to extract the lensing convergence from GW data; in section 4 we present the observables used to probe the cosmological models, while in section 5 we report the details of our analysis, and collect our results in section 6; finally in section 7 we draw our conclusions.

2 Gravitational waves observation

General relativity predicts two polarization modes, h_+ and $h_\times$, for the strains of GW emitted from coalescing compact objects. The strain for a binary system of masses m_1 and m_2 can be modelled in terms of the redshifted chirp mass $\mathcal{M}_z = (1+z)\mathcal{M} \equiv (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$. Assuming GR, and under the approximation of circular orbits, at leading order in frequency this reads

$$h_+(f) \propto \frac{1 + \cos^2 \iota}{2d_L^{\text{gw}}} (\mathcal{M}_z^{5/6}) (f^{-7/6}) e^{i\psi_+(\mathcal{M}_z f)}, \quad (2.1)$$

$$h_\times(f) \propto \frac{\cos \iota}{d_L^{\text{gw}}} (\mathcal{M}_z^{5/6}) (f^{-7/6}) e^{i\psi_\times(\mathcal{M}_z f)}, \quad (2.2)$$

where ι is the inclination angle and ψ denotes the phase of the GW strain, with $\psi_\times = \psi_+ + \pi/2$. The observed GW signal $h_{\text{obs}}(t) = F_+(t)h_+(t) + F_\times(t)h_\times(t)$ is a linear composition of the two polarization strain of the GW modulated by the antenna pattern $F_{+,\times}$. Thus, the

luminosity distance d_L to the GW sources can be measured directly from the GW signal, as it is inversely proportional to the amplitude of the strain [65]. Its accuracy can be increased by combining the two polarizations to mitigate the degeneracy between d_L and the inclination angle. This same degeneracy can also be reduced through the measurement of higher modes of the GW signal [66, 67] or exploiting external information such as, in case the binary merger is accompanied by the emission of a relativistic jet, the angle of the jet with respect to the line of sight [68–71]

As the GW strain can be parameterized in terms of the redshifted masses of the binary objects, GW sources cannot probe independently the redshift of the source unless there is a known mass-scale or physical scale which can be used to break the degeneracy between mass and redshift [72–76]. For some GW sources, a detectable electromagnetic counterpart can potentially be observed and used to measure its redshift directly with a spectroscopic (or photometric) follow-up of the host galaxy. Among these sources are BNS like GW170817, whose coincident jet emission has been observed as a gamma-ray burst [5–7] and later in X-rays [77] and radio [69, 78, 79], while the kilonova emission resulting from the synthesis of heavy elements in the merger ejecta has been followed up in optical and near-infrared [7, 80–82]. NS-BH are expected to produce EM emission through similar mechanisms [83], while stellar origin and intermediate mass BBH that merge in a gas-rich environment, such as an AGN disk, could emit X-rays in the late inspiral phases as a consequence of super-Eddington accretion [84, 85]. Finally, MBHB embedded in accretion disks can potentially emit X-rays and gamma rays immediately prior to or after the merger through two main channels: the perturbations induced on the disk by the recoil kick imparted on the merger product by the GW emission [86, 87] and the possible formation of relativistic jets [88–91]. These sources are typically dubbed ‘bright’ sirens. Beside the hunt for EM signatures accompanying the GW emission, if the direction of the incoming GW is sufficiently well localized it is in principle possible to identify the galaxy hosting the merging binary using the GW signal alone. In practice, this is not going to be possible with third-generation detectors, as they are expected to localize GW event in a sky area of several squared degrees, encompassing thousands of galaxies [10, 92, 93]. However, fourth-generation detectors like DECIGO [94], BBO [95] and ALIA [96] are expected to provide localization areas of the order of squared arcminutes, even arcseconds in the best scenarios [95, 96], which will largely facilitate the identification of individual host galaxies within the catalogues of surveys like Euclid or LSST.

For the purposes of this study, we will focus on bright standard sirens, broadly defined as GW detections for which a photometric or spectroscopic redshift can be measured, including the case of host identification when possible.

However, if the GW sources are such that a direct EM counterpart can be expected, then the total number and redshift distribution of bright detections will depend on various astrophysical assumptions. These assumptions concern the EM emission mechanism, the intensity and frequency of the emitted EM radiation and the availability and characteristics of the EM detectors tasked with the follow-up observation of the GW signal. Similarly, if the redshifts are determined through the identification of the host galaxy, the redshift distribution of bright (in the broad sense that we intend here) sources will depend on the completeness and depth of the galaxy catalogues used to identify the hosts.

Realistically, we can envision a future scenario in which the final catalogue of GW events usable for analyses similar to the one presented in this paper will be made up of sources whose redshifts have been measured in different ways. These include both host identification and direct EM counterparts that are emitted through different astrophysical channels. Thus, we

do not make any restrictive assumptions regarding a specific type of counterpart observation. Rather, we directly make an assumption on the redshift distribution of the final catalogue of GW sources. We detail such assumption in section 5.3, and investigate its impact on the cosmological constraints in appendix B.

Since the weak lensing phenomenon will affect GW events generated by binary systems of all masses, a higher statistics of events can be stacked by combining observations from space-based detectors such as LISA [9], Taiji [97], TianQin [98] or DECIGO, and ground-based detectors like LVK [99], CE [11, 100] and the ET [101]. These observatories can detect numerous bright GW sources, among which BNS, NS-BH and massive BBH. Moreover, a limited number of these sources might be observed in different frequency bands by both space and ground-based detectors, whose observations can then be combined to reduce uncertainties in the inference of the GW source parameters [102]. This combined analysis is expected to provide better determination of the luminosity distance measurement, though it might be applicable only to a small number of events [103–105].

3 Theoretical framework

With the new generation of LSS and GWs data, we will have the opportunity to test gravity on cosmological scales with unprecedented precision. Alternatives to Λ CDM introducing a dynamical dark energy or modifications of GR, will be explored and constrained. In this work, we focus on dynamical models of dark energy and modifications of GR that could be relevant at late times, on large cosmological scales. While the theory landscape is vast, most of the models in this category introduce a propagating scalar degree of freedom, either directly or modifying GR equations from constraints to dynamical ones [26]. In order to satisfy some minimal stability requirements, they must lead to second order equations of motion for the propagating degrees of freedom: one massless tensor field (with the two standard polarization modes) and one scalar. Models that meet these characteristics are broadly referred to as *scalar-tensor* theories, or Horndeski gravity [33].

We restrict our study to Horndeski models in which tensor modes propagate luminally, i.e. $c_T^2 = 1$, at all redshifts, satisfying the bound from GW170817 [6] as studied in [106–108]. In fact, with the advent of space based GW detectors, constraints on c_T could be further narrowed through multi-wavelength observations, see e.g. [109]. Rather than choosing a specific Horndeski model, we opt for a more agnostic exploration based on the effective field theory of dark energy (EFT of DE) [27–32] which spans a vast range of beyond- Λ CDM models, while containing Λ CDM as a subcase. For a thorough review of the EFT of DE formalism and its applications to cosmological tests of gravity, we refer the reader to [110]. EFT of DE offers a powerful unifying framework in terms of a handful of free functions of time, commonly referred to as *EFT functions*. A peculiar feature of EFT of DE is that it allows to formulate a set of conditions that the EFT functions need to satisfy in order for the resulting theory to be viable, i.e. for instance to not develop instabilities [31, 110–113]. These can be quite powerful in restricting a priori the parameter space for which it makes sense to explore the phenomenology [114, 115].

The EFT of DE action for Horndeski models can be written in unitary gauge, upon identifying the slices of constant time with the hypersurfaces of a uniform scalar field. Imposing second-order equations of motion, and additionally a luminal speed of propagation

for tensors, the resulting quadratic action reads

$$S = \int d^4x \sqrt{-g} \left\{ \frac{m_0^2}{2} [1 + \Omega(\tau)] R + \Lambda(\tau) - c(\tau) a^2 \delta g^{00} + \right. \\ \left. + \gamma_1(\tau) \frac{m_0^2 H_0^2}{2} (a^2 \delta g^{00})^2 - \gamma_2(\tau) \frac{m_0^2 H_0}{2} (a^2 \delta g^{00}) \delta K_\mu^\mu \right\} + S_m[g_{\mu\nu}, \chi_i]. \quad (3.1)$$

where $m_0^2 = (8\pi G)^{-2}$ and $\delta g^{00}, \delta K_\mu^\mu$ are, respectively, the reduced Planck mass, the perturbations of the time-time component of the metric, and the trace of the perturbations to the extrinsic curvature of constant-time hypersurfaces; S_m is the action for matter fields χ_i , minimally coupled to the metric (for more details on the construction of this action see [30, 31]). The free functions of time $\Omega(a), \Lambda(a), c(a)$ and $\gamma_1(a), \gamma_2(a)$ are the EFT functions; the first three affect the dynamics both of the background and linear cosmological perturbations, while the latter two affect only perturbations. Λ CDM is included in this framework, and it corresponds to the choice $\Lambda(a) = \text{const}$, with the rest of EFT functions being zero.

Different EFT functions correspond to different characteristics of the theory: the non-minimal coupling $\Omega(a)$, leads to a running Planck mass; the kineticity $\gamma_1(a)$, quantifies the independent dynamics of the scalar field; the braiding $\gamma_2(a)$, broadly signals a coupling between the metric and the scalar degree of freedom. Note that we adopt the convention of [116, 117] for the EFT functions. Alternative conventions are found in the literature, most commonly the so-called α_i parametrizations [118] in terms of $\{\alpha_M, \alpha_K, \alpha_B\}$ (while $\alpha_T = c_T^2 - 1$ is zero for our case) for which there is a simple direct correspondence with $\{\Omega, \gamma_1, \gamma_2\}$ (see e.g. [110] and references therein). For reference, non-minimally coupled quintessence, f(R) gravity and Brans-Dicke theories would be characterized by the equation of state parameter of DE, w_{DE} , and Ω , while having $\gamma_1 = \gamma_2 = 0$; k-essence would further have $\gamma_1 \neq 0$ and k-mouflage correspond to all functions being non-zero.

3.1 Luminosity distance in scalar-tensor theories

In scalar-tensor scenarios, the non-minimal coupling, parameterized by $\Omega(z)$, induces an additional dissipation term in the equation for propagation of GWs on the cosmological background [37, 41, 118]. This is not the case for photons; the dependence of their luminosity distance on the expansion rate remains unaltered because photons do not couple directly to the scalar field [37, 38, 41]. It follows that the luminosity distance to a given redshift differs if determined via an electromagnetic or GW wave, according to:

$$\bar{d}_L^{\text{gw}}(z) = \sqrt{1 + \Omega(z)} \bar{d}_L^{\text{em}}(z). \quad (3.2)$$

where

$$\bar{d}_L^{\text{em}}(z) = (1+z)\chi(z) = \frac{(1+z)}{H_0} \int_0^z \frac{dz'}{E(z')}, \quad (3.3)$$

with $E(z) \equiv H(z)/H_0$. The non-minimal coupling $\Omega(z)$ in eq. (3.2), is more commonly written in terms of the running Planck mass as $M_P(z) = M_P(z=0)\sqrt{1 + \Omega(z)}$.

On their journey through the expanding Universe, photons and GWs encounter structure which induces direction-dependent corrections to their amplitude, and correspondingly inferred luminosity distance. These include weak lensing, Doppler, Sachs-Wolfe, Integrated Sachs-Wolfe, Shapiro time delays and volume effects on the propagating waves, which translate into corrections to the luminosity distance to the corresponding source [119, 120]. The

luminosity distance becomes then a function not only of redshift but also of angular direction, i.e.

$$d_L^{\text{gw}}(z, \hat{n}) = \bar{d}_L^{\text{gw}}(z) + \Delta d_L^{\text{gw}}(z, \hat{n}), \quad (3.4)$$

where we have assumed that the effects due to the inhomogeneities are small compared to the background luminosity distance so that $\Delta d_L^{\text{gw}}(z, \hat{n})$ is the correction due to large scale structures.

For a detailed derivation of relativistic corrections to d_L^{gw} in scalar-tensor theories we refer the reader to [121]. These have been computed only in the subclass of Horndeski theories with $c_T^2 = 1$, and this is the reason we consider the same set of theories. They were derived, using the covariant formulation of Horndeski gravity, and have been subsequently mapped into the EFT of DE language, as well as implemented in EFTCAMB in [122].

Depending on the angular scale and redshift, some of the relativistic effects can be dominant. For instance, in [48, 119, 123] it was shown that for GW sources at redshifts higher than $z \sim 0.5$, WL is the dominant LSS correction and can quickly reach $\sim 5 - 10\%$ of the GW strain amplitude. Given its substantial magnitude, WL can also be a valuable signal to be exploited, rather than only a source of error. Indeed, in this work we use it to probe the growth history and pattern of the LSS in the Universe which depends on the gravitational model taken into consideration. Since WL is an integrated effect it builds up along the propagation. For this reason, we focus on GW sources at high redshift, reaching up to $z \simeq 2.5$, for which we can approximate

$$\frac{\Delta d_L^{\text{gw}}}{\bar{d}_L^{\text{gw}}} \simeq -\kappa_{\text{gw}}, \quad (3.5)$$

where κ_{gw} is the lensing convergence field of GWs, that for a populations of sources reads

$$\kappa_{\text{gw}}(\hat{n}) = -\frac{1}{2} \nabla_\theta^2 \phi_L(\hat{n}) = - \int_0^\chi \frac{d\chi'}{\chi'} \int_{\chi'}^\infty d\chi_* \left(\frac{\chi_* - \chi'}{\chi_*} \right) \frac{dn_{\text{gw}}}{d\chi_*} \nabla_\theta^2 \Psi_W(\chi' \hat{n}, z') \quad (3.6)$$

where $dn_{\text{gw}}/d\chi_*$ is the distribution of the GW sources, Ψ_W is the Weyl potential and ∇_θ^2 indicates the 2D Laplacian with respect to the angle between the image and optical axis. Note that we are neglecting the shear deformations of the signal since these are subdominant in linear perturbation theory where WL mainly affects the magnification of the GWs. The convergence κ_{gw} can probe the growth of matter perturbations in the Universe because it is given in terms of the Weyl potential, which in addition is sensitive to modifications of gravity affecting the Poisson equation.

3.2 Lensing convergence estimators for gravitational waves

In the case of bright events, building on eq. (3.5) we can construct an estimator to extract the WL convergence from GW data as follows

$$\hat{\kappa}_{\text{gw}}(z, \hat{n}) \equiv 1 - \frac{d_L^{\text{gw}}(z, \hat{n})}{\bar{d}_L^{\text{em}}(z)}, \quad (3.7)$$

namely as the fractional difference between the GW luminosity distance $d_L^{\text{gw}}(z, \hat{n})$, inferred from the measurement of the GW strain, and the electromagnetic background luminosity distance $\bar{d}_L^{\text{em}}(z)$, obtained from the known redshift via eq. (3.3) by choosing a fiducial cosmological model. This estimator can be biased by three factors: the experimental error on

the $d_L^{\text{gw}}(z, \hat{n})$ measurement, the error on the source redshift, or the choice of the wrong fiducial model. Indeed, the calculation of $\bar{d}_L^{\text{em}}$ depends not only on the measured z , but on a set of fiducial cosmological parameters that enter eq. (3.3) and whose chosen values affect the final result. In other words, the uncertainties on the parameters that characterize the fiducial model propagate, through eq. (3.3), into an uncertainty on $\bar{d}_L^{\text{em}}$. To account for the sources of bias mentioned above we introduce three parameters ϵ_{gw} , ϵ_z and ϵ_c , respectively representing the uncertainties due to the d_L^{gw} measurement error, the redshift error and the fiducial model, and modify the convergence estimator as

$$\hat{\kappa}_{\text{gw}} = 1 - \frac{\sqrt{1 + \Omega(z)} (1 - \kappa_{\text{gw}} + \epsilon_{\text{gw}}) \bar{d}_L^{\text{em}}}{(1 + \epsilon_z + \epsilon_c) \bar{d}_L^{\text{em}}} \sim 1 - \sqrt{1 + \Omega(z)} (1 - \kappa_{\text{gw}} + \epsilon_{\text{gw}} - \epsilon_z - \epsilon_c), \quad (3.8)$$

where we linearized to first order in κ_{gw} , ϵ_{gw} , ϵ_z and ϵ_c , assuming κ_{gw} to be small and of the same order as the ϵ_i .

The estimator in eq. (3.7) crucially relies on the availability of the information on the redshift of the source, to disentangle the lensing convergence from the measured luminosity distance. This is one of the reasons that prompted us to restrict the analysis to bright events. Not only, accurate redshift measurements also allow for a very fine binning of sources in redshift space. Exploiting the weak lensing measurement from dark events - which are expected to be significantly higher in number than bright sources - is not impossible, and has been tackled in the context of 3x2 point statistics of GW sources e.g. in [124], but it requires to work in luminosity distance space. This carries limitations due to the experimental uncertainty on d_L , which imposes a much larger bin size and consequent loss of information, as the lensing correlation gets partially smoothed out in wider tomographic bins (see appendix A). Arguably, however, the greatest advantage of employing bright events for our probes lays in the accessibility of smaller angular scales, where the lensing correlation is stronger, as we discuss in section 5. We show the effectiveness of $\hat{\kappa}_{\text{gw}}$ for cross-correlation purposes in the next section.

4 Tomographic observables

We consider cross-correlations between the density field of galaxies, δ_g , the weak lensing convergence fields as measured by galaxies, κ_g , and the weak lensing convergence field as measured by GW, κ_{gw} . The angular power spectrum for the cross-correlations is [125]

$$C_{ij}^{XY}(\ell) = \int_0^{z_{\text{max}}} \frac{dz}{\chi^2(z)H(z)} W_X^i(k(\ell, z), z) W_Y^j(k(\ell, z), z) P_P(k(\ell, z), z), \quad (4.1)$$

where $X, Y = [\delta_g, \kappa_g, \kappa_{\text{gw}}]$ at the i -th and j -th tomographic bin, $P_P(k) \propto A_s(k/k_*)^{n_s-1}$ is the primordial power spectrum, with A_s and n_s its amplitude and spectral index and k_* a pivot scale, and we applied the Limber and flat-sky approximations [126–131], which sets $k(\ell, z) = (\ell + 1)/\chi(z)$. For brevity, we indicate $k(\ell, z)$ as k throughout the rest of the paper.¹ $W_{X_i}(k, z)$ is the window function for the observable X in the i -th tomographic bin. For galaxies, the window function can be written as

$$W_{\delta_g}^i(k, z) = T_\delta(k, z) b_g^i(z) n_g^i(z) H(z), \quad (4.2)$$

¹Note that, with respect to the approach of [125], in eq. (4.1) we absorb the transfer functions in the window functions W_X^i , which makes them depend explicitly on the scale k and makes the integral depend on the primordial power spectrum P_P , rather than on the matter power spectrum $P(k, z)$.

where $T_\delta(k, z)$ is the matter transfer function, evolving the primordial power spectrum such that $P_\delta(k, z) = T_\delta^2(k, z)P_p(k)$. In eq. (4.2), $n_g^i(z)$ and $b_g^i(z)$ are, respectively, the galaxy redshift distribution and the linear galaxy bias in the i -th redshift bin which we model following [125], i.e. assuming the bias is constant in each redshift bin, with the values used as a free nuisance parameter in our analysis.

The window function of the GW convergence we adopt is

$$W_{\kappa_{\text{gw}}}^i(k, z) = T_{\Psi_W}(k, z) \int_z^\infty dz' \frac{\chi(z') - \chi(z)}{\chi(z')} n_{\text{gw}}^i(z'), \quad (4.3)$$

where $n_{\text{gw}}^i(z)$ is the redshift distribution of GW sources in the i -th bin, and $T_{\Psi_W}(k, z)$ is the Weyl potential transfer function, which allows to get its power spectrum $P_{\Psi_W}(k, z) = T_{\Psi_W}^2(k, z)P_p(k)$. The galaxy WL convergence window function is

$$W_{\kappa_g}^i(k, z) = T_{\Psi_W}(k, z) \int_z^\infty dz' \frac{\chi(z') - \chi(z)}{\chi(z')} n_g^i(z') + W_{\text{IA}}^i(k, z), \quad (4.4)$$

where, with respect to eq. (4.3) we introduce an extra term $W_{\text{IA}}^i(k, z)$ to include the Intrinsic Alignment (IA) systematic effect that we model following [125], i.e.

$$W_{\text{IA}}^i(k, z) = -T_\delta(k, z) \frac{A_{\text{IA}} C_{\text{IA}} \Omega_{\text{m},0} F_{\text{IA}}(z)}{D(z)} n_g^i(z) H(z), \quad (4.5)$$

where $D(z)$ is the growth factor, $\Omega_{\text{m},0}$ the current matter density, and the subscript IA highlights the terms including the nuisance parameters A_{IA} , β_{IA} and η_{IA} , the last two being contained in the F_{IA} function, while $C_{\text{IA}} = 0.0134$ is a fixed constant.

Scalar-tensor theories modify the growth pattern of perturbations, affecting differently the different observables. The window functions just described will be correspondingly modified, mainly through the transfer functions $T_\delta(k, z)$ and $T_{\Psi_W}(k, z)$, which encode the evolution of density perturbation and lensing potential, respectively. Modifications of the growth also affect the growth factor $D(z)$ entering the IA contribution to the power spectra. All these effects are included in our angular power spectra, and the cosmological functions needed for the calculations, i.e. the functions describing the background expansion ($H(z)$ and $\chi(z)$) and those related to the perturbation evolution (the growth factor $D(z)$ and the transfer functions $T_\delta(k, z)$ and $T_{\Psi_W}(k, z)$) are obtained numerically via **EFTCAMB** [116, 117], which modifies the public code **CAMB** [132, 133] to include modified gravity models.

Finally, we model the distribution of sources (src) in each redshift bin as

$$n_{\text{src}}^i(z) = \frac{dn_{\text{src}}}{dz} \left[\text{Erf} \left(\frac{z - z_-^i}{\sqrt{2}\sigma_z^{\text{src}}(z)} \right) - \text{Erf} \left(\frac{z - z_+^i}{\sqrt{2}\sigma_z^{\text{src}}(z)} \right) \right], \quad (4.6)$$

with $\text{src} = [\text{g}, \text{gw}]$, z_-^i and z_+^i the lower and upper limits of the i -th bin, $\sigma_z^{\text{src}}(z)$ the error on the redshift measurement for the considered source, and dn_{src}/dz is parameterised as

$$\frac{dn_{\text{src}}}{dz} \propto \left(\frac{z}{z_0} \right)^2 \exp \left[- \left(\frac{z}{z_0} \right)^{3/2} \right]. \quad (4.7)$$

The parameter z_0 entering eq. (4.7), as well as the redshift error σ_z^{src} of eq. (4.6), are survey dependent and we will specify them in the following section (section 5) where we introduce the surveys considered in this study.

We now focus on the GW lensing auto-correlation, $C^{\kappa_{\text{gw}}\kappa_{\text{gw}}}(\ell)$, and its cross-correlations with galaxies, $C^{\kappa_{\text{gw}}\delta}(\ell)$ and $C^{\kappa_{\text{gw}}\kappa_g}(\ell)$. In the case of GW events with electromagnetic counterpart, to build these correlations we can use the weak lensing convergence estimator of eq. (3.8). In this case, since the three sources of error $\{\epsilon_{\text{gw}}, \epsilon_z, \epsilon_c\}$ are not correlated between themselves or with either the convergence or the contrast density field, we have

$$\langle \hat{\kappa}_{\text{gw}} \delta_g \rangle \simeq \sqrt{1 + \Omega} \langle \kappa_{\text{gw}} \delta_g \rangle, \quad (4.8a)$$

$$\langle \hat{\kappa}_{\text{gw}} \kappa_g \rangle \simeq \sqrt{1 + \Omega} \langle \kappa_{\text{gw}} \kappa_g \rangle, \quad (4.8b)$$

$$\begin{aligned} \langle \hat{\kappa}_{\text{gw}} \hat{\kappa}_{\text{gw}} \rangle &\simeq (1 + \Omega) [\langle \kappa_{\text{gw}} \kappa_{\text{gw}} \rangle + \langle \epsilon_{\text{gw}} \epsilon_{\text{gw}} \rangle + \langle \epsilon_z \epsilon_z \rangle] = \\ &= (1 + \Omega) \langle \kappa_{\text{gw}} \kappa_{\text{gw}} \rangle + N_{\text{gw}} + N_z, \end{aligned} \quad (4.8c)$$

where we have kept only terms up to second order in perturbations and used that $\langle \epsilon_{\text{gw}} \rangle = \langle \epsilon_z \rangle = 0$ when the average is performed over large volumes. On the contrary, in general $\langle \epsilon_c \rangle \neq 0$, hence in a complete analysis the term proportional to $\langle \epsilon_c \rangle$ should be accounted for by introducing a proper additional term to the correlations noise. However, this correction is typically small and for the purposes of this work we will assume the noise to be dominated by other terms, that we model in section 5. Hence, we neglected the contributions $\langle \epsilon_c \rangle$, together with those $\propto \langle \epsilon_c \epsilon_c \rangle$ which are second order.

Exploiting the results of eq. (4.8), the angular power spectrum in eq. (4.1) can be re-written as

$$\hat{C}_{ij}^{XY}(\ell) = \int_0^{z_{\text{max}}} \frac{dz}{\chi^2(z)H(z)} [\Theta_X(z)\Theta_Y(z)] W_X^i(k, z) W_Y^j(k, z) P_p(k), \quad (4.9)$$

where again $X, Y = [\delta_g, \kappa_g, \hat{\kappa}_{\text{gw}}]$, $\Theta_{\hat{\kappa}_{\text{gw}}}(z) = \sqrt{1 + \Omega(z)}$ and $\Theta_{[\delta, \kappa_g]}(z) = 1$. We point out that now the angular correlations involving GW include factors $\sqrt{1 + \Omega(z)}$, which can make these probes promising for capturing deviations from Λ CDM. This multiplicative factor is even squared in the case of the auto-correlation of the GW weak lensing field.

5 Models, analysis method and surveys specifications

This section provides details of the gravitational models on which we forecast constraints, as well as the details on the analysis we performed to obtain them. Furthermore, we also provide the galaxy and GW survey specifications that are considered.

5.1 Models

The models we investigate in this work belong to the Horndeski class of theories introduced in section 3 and described by the action of eq. (3.1). We have already discussed at length the role of $\Omega(a)$ in the previous sections. As for $\gamma_1(a)$, past works have shown that it affects only negligibly linear perturbations in the range of our observables [134–136] and does not enter into the modified equations for the propagation of tensors [110]. The function $\gamma_2(a)$ is non-zero in models that have a kinetic mixing between the scalar field and the metric [110]; similarly to γ_1 , it does not enter explicitly in the linear equation for tensors, but contrary to γ_1 it has a non-negligible effect on the LSS [134–136] hence we expect our observables to be sensitive to it.

We work in the *designer* approach [117] and fix the background expansion history to a choice for the equation of state w_{DE} . In this framework, Λ and c become superfluous, as they can be determined in terms of the Hubble parameter $\mathcal{H}(a)$ and the non-minimal coupling

$\Omega(a)$, and the model is then fully specified by the EFT functions $\{w_{\text{DE}}, \Omega, \gamma_1, \gamma_2\}$. We model the dark energy equation of state using the CPL parametrization [137, 138]

$$w_{\text{DE}}(a) = w_0 + w_a(1 - a) \quad (5.1)$$

With this prescription, the continuity equation dictates that the background DE density evolves as

$$\rho_{\text{DE}}(a) = \rho_{\text{DE}}^0 a^{-3(1+w_0+w_a)} e^{-3w_a(1-a)}, \quad (5.2)$$

with $\rho_{\text{DE}}^0 = \rho_{\text{DE}}(a=1)$ its value today. We choose to parameterize the time dependencies of $\Omega(a)$, $\gamma_1(a)$ and $\gamma_2(a)$ as

$$\Omega(a) = \Omega_0 \frac{\rho_{\text{DE}}}{\rho_{\text{DE}}^0}, \quad \gamma_1(a) = \gamma_1^0 \frac{\rho_{\text{DE}}}{\rho_{\text{DE}}^0}, \quad \gamma_2(a) = \gamma_2^0 \frac{\rho_{\text{DE}}}{\rho_{\text{DE}}^0} \quad (5.3)$$

considering that any modification to Λ CDM is expected to be small when ρ_{DE} is not the dominant contribution to the energy density budget of the Universe. We consider three scenarios:

- **Λ CDM:** in this case, the cosmological model is determined by the standard set of cosmological parameters. This means that the set of free parameters in the analysis will consist of the matter and baryonic relative energy densities at present time, $\Omega_{\text{m},0}$ and $\Omega_{\text{b},0}$, the reduced Hubble constant h , the spectral index of the primordial power spectrum n_s and the amplitude σ_8 of the linear matter power spectrum measured at $8/h$ Mpc. We consider all MG parameters fixed to their GR values, i.e. $w_0 = -1$ and $w_a = \Omega_0 = \gamma_1^0 = \gamma_2^0 = 0$;
- **Model I (M1):** this model allows us to study theories falling within the Generalized Brans-Dicke class. In this class, the background expansion history deviates from that of Λ CDM and depends on the specific parameters of the selected model. As we aim here for a generic approach we leave the background parameters w_0 and w_a free to vary, we allow for a non-minimal coupling ($\Omega_0 \neq 0$ and free to vary), and keep the standard Λ CDM free to vary. In these kind of models there is no derivative interaction, thus we set $\gamma_1^0 = \gamma_2^0 = 0$;
- **Model II (M2):** where the set of available parameters includes $\{w_0, w_a, \Omega_0, \gamma_2^0\}$ and those of Λ CDM. The background expansion history follows the same parametrization of M1, but with a different fiducial for the DE equations of state, and w_0 and w_a are free to vary. M2 represents the class of kinetic gravity braiding [139] models, where we allow for both a conformal coupling and derivative coupling (i.e. $\Omega \neq 0$, $\gamma_2 \neq 0$). For this model, we also take a non-zero $\gamma_1(a)$, which is necessary to ensure the mathematical stability of this theory around the assumed fiducial. As past works have shown that LSS are not sensitive to γ_1 , the corresponding parameter is treated as a constant in the analysis.

We refer the reader to table 1 for a complete list of the fiducial values chosen for the different parameters. Parameters that are kept fixed in the analysis are indicated in parentheses. Our set of fiducial Λ CDM parameters are chosen to be compatible with current constraints from LSS and CMB [135, 140]. Existing bounds on the DE equation of state and the MG parameters come from a number of observations, including GW [44, 141, 142], Big Bang Nucleosynthesis and CMB measurements [143], local tests of gravity [144–146] and theoretical constraints obtained imposing on MG theories a set of stability requirements [146].

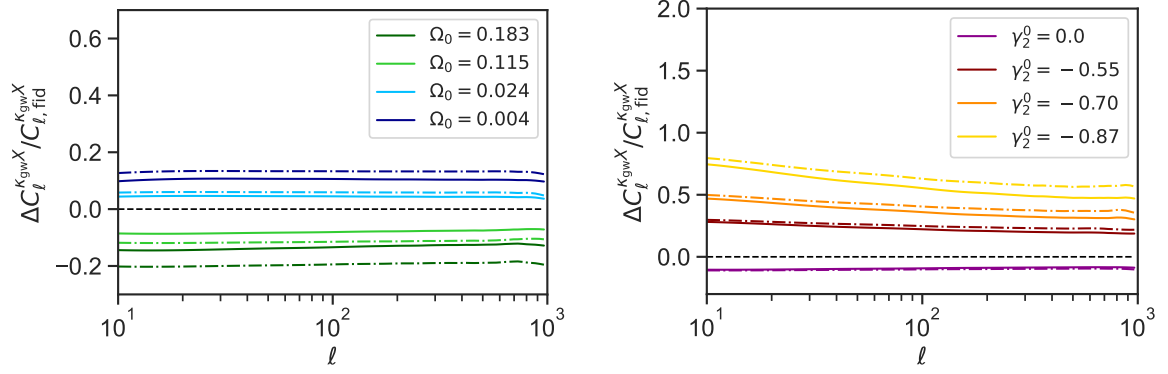


Figure 1. Relative deviations $(C_{\ell}^{\kappa_{\text{gw}}^X} - C_{\ell,\text{fid}}^{\kappa_{\text{gw}}^X})/C_{\ell,\text{fid}}^{\kappa_{\text{gw}}^X}$ as a function of multipole ℓ for model M2, where $C_{\ell,\text{fid}}^{\kappa_{\text{gw}}^X}$ are the correlations computed at the fiducial cosmology. We plot in the left (right) panel how the C_{ℓ} vary for different values of Ω_0 (γ_2^0), while we keep all other parameters fixed at the fiducial value. In each figure, the solid lines correspond to auto-correlations of GW-WL ($X = \kappa_{\text{gw}}$), while the dashed lines mark cross-correlations of GW-WL with galaxy WL ($X = \kappa_g$).

Model	h	$\Omega_{m,0}$	$\Omega_{b,0}$	n_s	σ_8	w_0	w_a	Ω_0	γ_1^0	γ_2^0
Λ CDM	0.6774	0.31	0.05	0.9667	0.8159	(-1.0)	(0.0)	(0.0)	(0.0)	(0.0)
Model I	0.6774	0.31	0.05	0.9667	0.8159	-0.946	-0.098	0.018	(0.0)	(0.0)
Model II	0.6774	0.31	0.05	0.9667	0.8159	(-0.94)	(-0.31)	0.047	(4.4)	-0.23

Table 1. Fiducial parameter values for the cosmological models considered in this work. The numbers in parenthesis correspond to parameters that were kept fixed in that specific model, while all other parameters were constrained simultaneously. Along with these parameters, we also let free to vary the galaxy bias in each bin and the IA parameters, for which we use the same fiducial values as [125].

However, constraints on the non- Λ CDM parameters, and especially on the MG parameters, are significantly model dependent, as the values of the EFT functions today are partially degenerate with their assumed evolution in time and with the assumed model for the DE equation of state. As an example, in [135] Ω_0 is constrained to be positive if its redshift evolution is the same as eq. (5.3), while negative values for Ω_0 are found to be favoured when the non-minimal coupling is assumed to evolve as a power law of the scale factor. Oftentimes, bounds are also placed on a re-parameterization of the EFT action in eq. (3.1) in terms of a different set of EFT functions, commonly dubbed $\{\alpha_i\}$ [118, 136, 140, 147–150] or on a set of phenomenological functions introduced specifically to capture deviations from Λ CDM in observations [140, 151–153]. These functions - and the resulting bounds - can be mapped into our set $\{\Omega, \gamma_i\}$ (see e.g. [117]), though the mapping is not trivial. For a complete overview of the existing cosmological bounds on EFT parameters, encompassing different parametrizations of the EFT action and several models, we refer the interested reader to the extensive review work of [110]. Our chosen fiducial values correspond to the best fit values obtained in [135], where the non- Λ CDM parameters were constrained against current cosmological data from CMB, BAO, SNIa and LSS. This work assumes the parametrization (5.3) for the EFT functions, and finds (at 2σ significance) $w_0 = -0.946^{+0.090}_{-0.060}$, $w_a = -0.098^{+0.25}_{-0.28}$ and $\Omega_0 = 0.018^{+0.032}_{-0.019}$ for our Model I, and $w_0 = -0.94^{+0.15}_{-0.13}$, $w_a = -0.31^{+0.48}_{-0.63}$, $\Omega_0 = 0.047^{+0.068}_{-0.051}$,

$\gamma_1^0 > 0.295$ and $\gamma_2^0 = -0.23_{-0.32}^{+0.26}$ for our Model II. These are, to our knowledge, the most stringent constraints to date corresponding to these models.

We sketch the impact of the EFT parameters on the $C^{XY}(\ell)$ in figure 1, where we plot the relative difference $(C_{\ell}^{\kappa_{\text{gw}}X} - C_{\ell, \text{fid}}^{\kappa_{\text{gw}}X})/C_{\ell, \text{fid}}^{\kappa_{\text{gw}}X}$ adopting model M2 and choosing as baseline the cross-correlation computed in the fiducial cosmology. We plot curves for $X = \kappa_{\text{gw}}$ (solid lines) and $X = \kappa_g$ (dashed lines) choosing different values of Ω_0 in the left panel and of γ_2^0 in the right panel, while all other parameters are kept fixed at the fiducial value. We observe that taking Ω_0 lower than the fiducial value (and closer to 0) acts to strengthen the signal, while values higher than the fiducial diminish the signal. On the contrary, higher values of γ_2^0 and closer to 0 tend to dampen the signal, while lower values enhance it. The values chosen for the left panel of figure 1 span the full range available to Ω_0 in this model, up to twice the 2σ constraints of [135], while causing variations in the cross-correlations of, at most, 20%. Similarly, values for γ_2^0 reaching down to twice the 2σ bounds of [135] can cause variations of up to 70% in the correlations. In figure 1, we excluded positive values of γ_2^0 , as we found that, if all other parameters remain set at the fiducial values, a positive γ_2^0 is theoretically forbidden for this model, giving rise to theories that are mathematically unstable.

5.2 Analysis method

To compute forecasts, we adopt the Fisher matrix formalism [125, 154], which allows us to obtain bounds on the free parameters of the analysis from the information matrix $\mathcal{F}_{\alpha\beta}$. Following [125], we define the Fisher matrix as

$$\mathcal{F}_{\alpha\beta} = \sum_{\ell=\ell_{\min}}^{\ell_{\max}} \frac{2\ell+1}{2} \sum_{i,j,m,n} \sum_{A,B,C,D} \frac{\partial C_{ij}^{AB}(\ell)}{\partial \theta_{\alpha}} [K^{-1}(\ell)]_{jm}^{BC} \frac{\partial C_{mn}^{CD}}{\partial \theta_{\beta}} [K^{-1}(\ell)]_{ni}^{DA}, \quad (5.4)$$

where α and β run over the set of free cosmological parameters θ (standard and MG), while A, B, C and D run over the density and convergence fields $[\delta_g, \kappa_g, \hat{\kappa}_{\text{gw}}]$ and finally i, j, m and n run over all unique pairs of tomographic bins. By the Cramer-Rao inequality, the lower bound on the standard deviation for the parameter θ_{α} is

$$\sigma_{\alpha} = \sqrt{\Sigma_{\alpha\alpha}} \quad \text{with} \quad \Sigma_{\alpha\alpha} \equiv [\mathcal{F}^{-1}]_{\alpha\alpha}. \quad (5.5)$$

To compute our matrices we use `CosmicFish` [155, 156],² that we extended to include GW weak lensing as an additional cosmological probe, incorporating also the MG contributions. The covariance matrix K is defined as

$$K_{ij}^{AB}(\ell) = \frac{C_{ij}^{AB}(\ell) + N_{ij}^{AB}(\ell)}{\sqrt[4]{f_{\text{sky}}^A f_{\text{sky}}^B}}, \quad (5.6)$$

where f_{sky}^A is the sky fraction covered by the detector measuring the observable A . Note that all currently planned GW interferometers present a full-sky area coverage ($f_{\text{sky}}^{\kappa_{\text{gw}}} = 1$), so there will always be total overlap with the region covered by EM surveys. N_{ij}^{AB} is the noise

²`CosmicFish` is publicly available at <https://cosmicfish.github.io/>. However, the version we are using is an update of the public one, which will be released in the near future.

of the correlation considered, which we model as

$$N_{ij}^{\delta_g \delta_g}(\ell) = \frac{1}{\bar{n}_g^i} \delta_{ij}, \quad (5.7)$$

$$N_{ij}^{\kappa_g \kappa_g}(\ell) = \frac{\sigma_\epsilon^2}{\bar{n}_{\kappa_g}^i} \delta_{ij}, \quad (5.8)$$

$$N_{ij}^{\kappa_{\text{gw}} \kappa_{\text{gw}}}(\ell) = \frac{1}{\bar{n}_{\text{gw}}^i} \left(\frac{\sigma_{d_L}^2}{d_L^2} + \frac{\sigma_s^2}{d_L^2} \right) e^{\frac{\ell^2 \theta_{\min}^2}{8 \ln 2}} \delta_{ij}, \quad (5.9)$$

where δ_{ij} is the Kronecker delta and $\bar{n}_A^i$ is the number of sources in the i -th redshift bin for the probe A . Here we assume that the instrument noise terms of the different probes are not correlated, i.e. $N_{ij}^{AB}(\ell) = 0$ for $A \neq B$, which is a reasonable assumption as GW detectors operate in a totally different manner than galaxy surveys. However, the terms in eq. (5.6) related to cosmic variance are still correlated between different probes, i.e. $C_{ij}^{AB}(\ell) \neq 0$ even when $A \neq B$. In these noise terms, σ_ϵ represents the intrinsic ellipticity affecting shear measurements, σ_{d_L} represents the average experimental error on the luminosity distance of the GW sources, while $\sigma_s = (\partial d_L / \partial z) \sigma_z^{\text{gw}}$ is the contribution to the luminosity distance error brought by the uncertainty on the merger redshift σ_z^{gw} , where the propagation is obtained assuming a fiducial cosmology. Lastly, $\theta_{\min}$ is the sky-localization area of the GW event, which also dictates the maximum available multipole for the analysis.

Finally, the combination of GW and galaxy tracers can in general be exploited in cosmic variance cancellation techniques for partially canceling the cosmic variance noise in the common Fourier modes [157]. This could further improve the predicted constraining power, though it is currently not included in our Fisher analysis and we leave it for future investigations.

5.3 Galaxy and GW surveys

The above discussion of the observables and methodology is quite generic and can be applied to nearly all of the upcoming galaxy surveys and GW observatories. We describe below the surveys specifications adopted to perform our forecasts.

For what concerns the galaxy distribution (eq. (4.7)) and the fraction of sky observed, in this work we use the specifications of [125], as representative of a Stage IV galaxy survey. We instead take a simpler approach for the photometric redshift error, using eq. (4.6) with $\sigma_z^g = 0.03(1+z)$, from which we obtain the distribution of sources in ten equipopulated redshift bins with $z \in [0, 2.5]$. For the limiting multipoles, we follow the “pessimistic” settings of [125], and we use a minimum multipole $\ell_{\min} = 10$ for both galaxy clustering and weak lensing, while we fix the maximum multipole to $\ell_{\max}^{\text{GC}} = 750$ for the former and $\ell_{\max}^{\text{WL}} = 1000$ for the latter. The $\ell_{\min}$ truncation eliminates the large scales, at which the Limber approximation breaks down and further relativistic effects become relevant alongside lensing and should be considered (see e.g. [158]). The $\ell_{\max}$ limit allows to account for both the uncertainties brought by the modelling of nonlinear scales, and the effects of non-Gaussian terms in the data covariance, which have a different impact on the two observables. However, with respect to the settings in [125], we further limit the weak lensing multipoles because of the lack, in the MG models considered, of a consolidated recipe for treating small scales, where the perturbations evolution enters the non-linear regime.

Concerning the GW survey, we take the parameter entering the source distribution in eq. (4.7) as $z_0 = 1.5$. This corresponds to a phenomenological parametrization for the

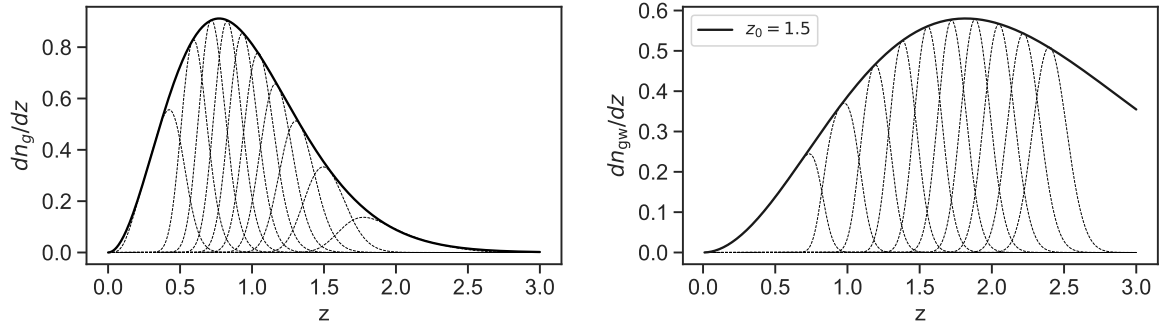


Figure 2. Normalized redshift distribution of galaxies and bright GW sources (left and right panel respectively). The dashed lines illustrate the redshift binning applied in our analysis.

normalized distribution of observed sources that we make compatible with the forecasts of [92] for the BNS observed by a network of ET in combination with two CE detectors. This parametrized redshift distribution is also similar to the one obtained in [159] for LISA luminous MBHB observations (i.e. MBHB with a detectable counterpart), though in that case the events number count peaks at higher redshifts, $z_0 \sim 2.5$. Even if our parametrization fails to accurately describe LISA sources, this is not extremely concerning for the purposes of the present study, as we expect - and will demonstrate below - that GW-WL will be effective in constraining cosmologies only if a high statistics of events is available. As current LISA forecasts at best predict only a few hundreds of total MBHB merger events over the span of 10 years [160], stellar mass binaries are likely to remain the primary target for GW-WL studies for the foreseeable future. Nonetheless, the class of experiments sensitive to these sources still encompasses most of the currently planned GW detectors, thus we feel safe in choosing $z_0 = 1.5$ without significant loss in generality. We characterize the impact on our analysis of choosing different values for z_0 in appendix B.

As in the case of galaxies, we consider $z \in [0, 2.5]$ and bin the GW sources in ten equipopulated redshift bins. We also explore different binning choices: we consider the cases with only eight and six equipopulated bins. We observe a very small deterioration in the cosmological constraints when decreasing the number of tomographic bins, as larger bins tend to smooth out the lensing signal, as we show in appendix A. However, our conclusions on the impact of GW-WL on the constraining power remains the same regardless of binning, thus in what follows we will report results only for the most favourable choice of 10 bins.

The main quantities that can impact the weak lensing estimation from bright GW sources are the uncertainty on the luminosity distance, the error on the source redshift and the total number of bright GW events. The peculiar velocity correction for sources at high redshift are going to be a less significant contamination [161]. Without specializing on the specifics of any particular detector, we consider several possibilities for the total number of detected GW events, N_{gw} , ranging from 10^3 to 10^6 , while we vary the average precision on the luminosity distance, σ_{d_L}/d_L , from 10% down to 0.5%. For what concerns the redshift error, we make the requirement that our sources will either have a detectable electromagnetic counterpart or an identifiable host galaxy, from which we can measure the redshift photometrically, with error $\sigma_z^{\text{gw}} = 0.03(1+z)$, or spectroscopically, with error $\sigma_z^{\text{gw}} = 0.0005(1+z)$ [17, 19, 162, 163]. In figure 2 we show the normalized source distribution of galaxies (left panel) and bright GW events (right panel) adopted in this analysis,

together with a representation of the redshift binning that is applied, and we summarise the specifications considered for galaxy and GW surveys in table 2.

As N_{gw} is the total number of events, $\bar{n}_{\text{gw}}^i$ in the noise eq. (5.9) can be written as the product between $N_{\text{gw}}/4\pi$ and the integral of the normalized source distribution over each bin. Since we are focusing on bright sources, we set $\theta_{\text{min}} = 0$ in eq. (5.9), thus implying that the event is perfectly localized by identifying the EM counterpart. This implies, in principle, that the GW analysis could include very small scales (very high values of ℓ), since the exponential term in eq. (5.9) which dampens the signal-to-noise ratio at high multipoles has no effect. Nevertheless, in our analysis we truncate the summation in eq. (5.4) at $\ell_{\text{max}} = 1000$ to match the limiting multipoles of the galaxy survey, for analogous reasons. This choice is equivalent to limit angular scales to $\theta_{\text{min}} \sim 11$ arcmin. Going up to $\ell_{\text{max}} > 1000$ would result in a further enhancement of the signal-to-noise ratio. This on one hand would imply tighter constraints on cosmological parameters when all sources are combined. At the same time though, the GC and WL signal themselves would benefit from the accessibility of higher multipoles, resulting in tighter galaxy-only constraints to begin with. Thus, we can expect that, qualitatively, the impact of GW on the galaxy bounds remains similar even if higher multipoles are accessible.

Pushing ℓ to such high values is a significant advantage of considering bright sirens. Including dark events limits significantly the highest reachable multipole, depending on the sky-localization Ω of the GW events as $\ell_{\text{max}} \sim 180/\sqrt{\pi\Omega}$. For example, for a 3G network composed by one ET and one or two CE detectors the localization of most sources is forecasted to be within $1 \text{ deg}^2 < \Omega < 100 \text{ deg}^2$ [10, 92, 93], which implies $10 \leq \ell_{\text{max}} \leq 100$. These scales are barely below the threshold at which the lensing correlations signal becomes relevant, thus reducing hugely the signal-to-noise. On the other hand, dark detections are expected to be much more numerous and to reach higher redshift depth, two factors which in contrast increase the SNR. It is not trivial to predict to which extent these advantages can counterbalance the lost multipoles. Preliminary to this work, we have conducted approximate explorations suggesting that the gain in constraining power coming from the higher statistics and redshifts might not be enough to compensate for the loss due to the inaccessibility of the smaller scales, at least until 4th generation detectors with very small localization volumes are considered. However, more work needs to be dedicated to properly characterize the potentialities of dark sirens, and we leave the extension of our analysis to include dark events for future investigation.

6 Results

Using the specifications, models and methodology described in section 5, we include in the Fisher matrix the contributions from all correlators C_{ℓ}^{XY} with X and Y in $[\delta_g, \kappa_g, \hat{\kappa}_{\text{gw}}]$, to explore the joint constraining power of GW-WL and galaxy surveys over cosmological parameters. We collect the fiducial values chosen for the parameters of the different models in table 1. We choose the fiducial cosmology in such a way that the parameters' values fall within the stable region of the models considered (identified via the stability sampler of EFTCAMB) and that the numerical derivatives, needed to obtain the Fisher matrices using eq. (5.4), can be performed without exiting this region.

In our analysis, we account for degeneracies between parameters by always attempting to constrain simultaneously the highest number of parameters possible for each model. For example, we always keep all the standard cosmological parameters free even when the GW-WL is used as a single probe, even though, since WL is mostly sensitive to the combination of Ω_m and σ_8 , we know a priori that it will not be much constraining of other parameters

Galaxy Clustering						
f_{sky}^g	σ_z^g	—	z_0	$\bar{n}_g \text{ [arcmin}^{-2}\text{]}$	ℓ_{min}	ℓ_{max}
0.35	$0.03(1+z)$		$0.9/\sqrt{2}$	30	10	750
Galaxy Weak Lensing						
$f_{\text{sky}}^{\kappa_g}$	σ_z^g	σ_ϵ	z_0	$\bar{n}_{\kappa_g} \text{ [arcmin}^{-2}\text{]}$	ℓ_{min}	ℓ_{max}
0.35	$0.03(1+z)$	0.3	$0.9/\sqrt{2}$	30	10	1000
Bright GW Weak Lensing						
$f_{\text{sky}}^{\kappa_{\text{gw}}}$	σ_z^{gw}	$\sigma_{d_L}/d_L (\%)$	z_0	N_{gw}	ℓ_{min}	ℓ_{max}
1	$0.03(1+z)$ $0.0005(1+z)$	$\in [0.5, 10]$	1.5	$\in [10^2, 10^6]$	10	1000

Table 2. Summary of the parameters entering the noise modeling for the probes considered and the binned source distributions of eq. (4.6). Parameters whose values change through the different analyses are reported with their range of variation.

such as, e.g., Ω_b . We choose to marginalize over these parameters rather than fixing them, not to neglect any possible degeneracy, as we will show and discuss below. However, in a few cases we instead decide to fix a parameter to its fiducial value, and put bounds only on the remaining set. Additionally, we always let free to vary the galaxy bias in each bin b_g^i and the IA parameters A_{IA} , β_{IA} and η_{IA} , for which we use the same fiducial values as [125].

6.1 Λ CDM

Let us start with the Λ CDM scenario. In figure 3 we show the marginalized 1σ relative bounds obtained on Ω_m and σ_8 varying the total number of observations N_{gw} and the luminosity distance precision σ_{d_L}/d_L . The first row shows the results for photometric observations of the electromagnetic counterparts or of the GW host galaxy, while the results with spectroscopic observations are shown in the second row.

For photometric observations we notice that, when GWs and galaxy surveys are combined, the bounds on both Ω_m and σ_8 do not change for about half of the configurations considered (yellow regions in the plots), regardless of the choices made for the GW sector. In these regions, the bounds on cosmological parameters are strongly dominated by galaxies, which contribute almost entirely to the constraining power. The black line marks the point at which GW events start to weigh in significantly. A high number of events is required for that: about 5×10^5 bright sources with luminosity distance determined at least at 1% precision, or 10^6 sources with d_L determined at 5% precision or better are necessary to be seeing a significant impact of GW-WL on the galaxy constraints.

The spectroscopic case is analogous, though there is one case where 5×10^4 GW events could already be sufficient to tighten the galaxy-only bounds on both Ω_m and σ_8 of $\sim 0.5\%$ and 0.25% respectively. For this to happen, however, the average σ_{d_L} is required to be no higher than 0.5% . Other than that, even in the spectroscopic scenario GW becomes complementary to galaxies only for $N_{\text{gw}} \geq 10^5$ for $\sigma_{d_L} \leq 1\%$ (or $N_{\text{gw}} \geq 10^6$ for $\sigma_{d_L} \leq 5\%$). We can though recognize the reduced noise affecting GW lensing in the overall better performance

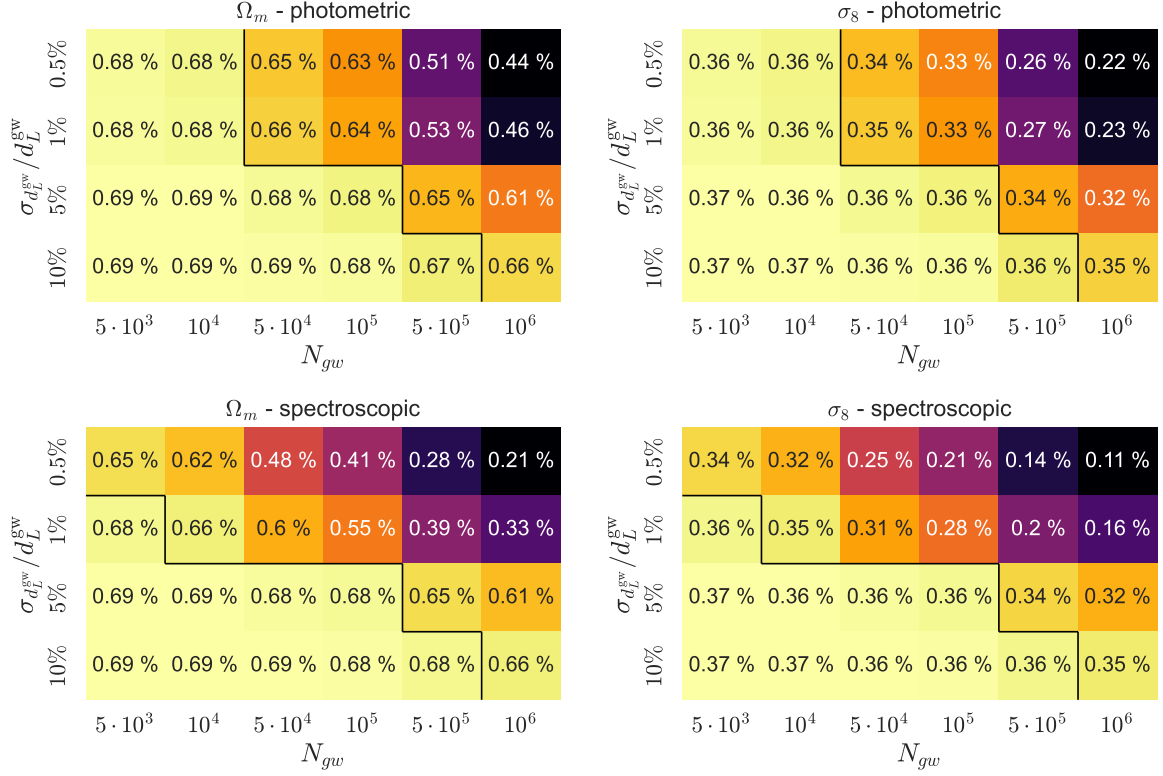


Figure 3. Marginalized 1σ relative confidence bounds for Ω_m (left panel) and σ_8 (right panel) in our Λ CDM model. In the top row, we consider bright sirens with EM counterparts observed photometrically, while in the bottom row we assume spectroscopic observations of the counterparts. In all tables are presented bounds from the combination of GW weak lensing, galaxies weak lensing and galaxy clustering. We vary the number of GW detections N_{gw} and the precision on the luminosity distance determination σ_{d_L} .

of the spectroscopic sirens: for all configurations in which GW contributes significantly to the constraining power, spectroscopic sources always provide bounds up to 0.1% tighter than the photometric ones. For example, the 0.7% ($\sim 0.4\%$) bound placed on Ω_m (σ_8) by galaxies can be reduced to 0.45% (0.2%) by factoring in the contribution of 10^6 photometric sirens measured with $\sigma_{d_L} = 1\%$, and further shrunk to $\sim 0.3\%$ (0.15%) considering instead the same number of spectroscopic events.

While for brevity we do not reproduce them here, we have performed the same analysis for the other free Λ CDM parameters. However, for the cases of h and Ω_b the impact of GW-WL remains mild for all configurations explored in terms of N_{gw} and σ_{d_L}/d_L . This is direct consequence, as anticipated, of GW-WL not being particularly sensitive to those parameters, which can instead be constrained much more efficiently through other GW probes, such as standard sirens. A different story holds for n_s , for which we find that GW-WL impacts the constraints in a way that is qualitatively absolutely analogous to Ω_m and σ_8 . To better understand this behaviour, we point to figure 4

In figure 4 we show the triangular plots for three free parameters of Λ CDM marginalising over the remaining two, in an idealistic scenario for which we choose $N_{gw} = 5 \cdot 10^5$ and $\sigma_{d_L}/d_L = 1\%$, showing the results from spectroscopic observations of the counterparts. We

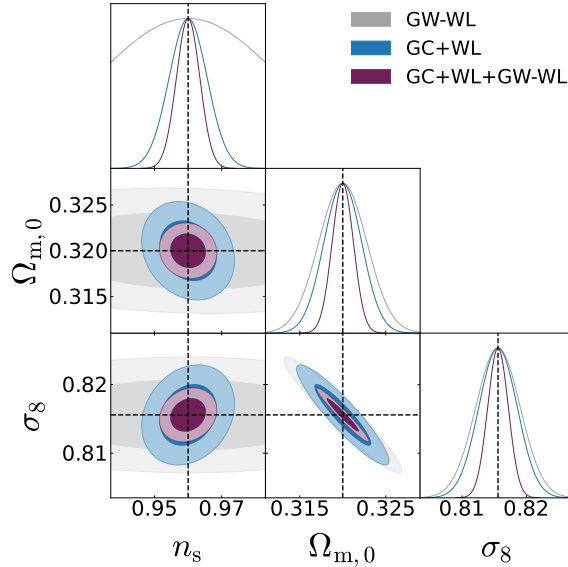


Figure 4. Marginalized forecasts for Λ CDM parameters in an idealistic scenario where $5 \cdot 10^5$ GW events are detected, with luminosity distance measured with 1% precision. All counterparts are assumed to be observed spectroscopically. Constraints come from GW weak lensing only (grey), galaxies WL and clustering (blue), and GW and galaxy probes combined, including contributions from the cross correlations (dark red).

plot bounds obtained considering GW-WL only, GC and galaxies WL, and the joint contribution of galaxy surveys and GW-WL. In the spectroscopic case the GW-WL constraints on $\Omega_{m,0}$ and σ_8 are comparable with those coming from galaxies, and they half when all probes are combined, also thanks to the reduced impact of nuisance parameters. Interestingly though, the GW-WL only constraints on n_s are quite large, showing that this probe is not particularly sensitive to variations of the spectral index. However, when all probes are combined, the impact of GW-WL cross-correlations with galaxies manifest by shrinking the bounds on σ_8 and Ω_m , in turn breaking the mild degeneracy that exists in the galaxy-only constraints between the couple of parameters (n_s, Ω_m) and (n_s, σ_8) . The net result is a strong reduction also of the bound on n_s , that still feels the impact of the new observables even while not being directly probed by GW-WL.

The analysis performed above shows that GW sources have the advantage of providing a completely independent measurement affected by different systematics and hence can shed light on the standard model of cosmology in a completely independent way. However, the precision of the measurement of the cosmological parameters can effectively improve over the constraints coming from galaxies only when the statistics of measured GW sources are large and the error on luminosity distance is small. Indeed, in Λ CDM where the non-minimal coupling $\Omega(a) = 0$, the estimator defined in eq. (3.7) corresponds to the same convergence measured by galaxy surveys. Hence, GW-WL and its cross-correlations with galaxies contributes to the constraining power only by effectively strengthening the WL statistics. Furthermore, we should emphasize that these constraints are compared with the measurement forecasted for galaxy survey only observations (GC+WL). However, as we show in figure 4, GW-WL can provide an independent measurement of the cosmological parameters that is, in some cases, comparable to those of galaxy surveys. Such independent measurement can be done

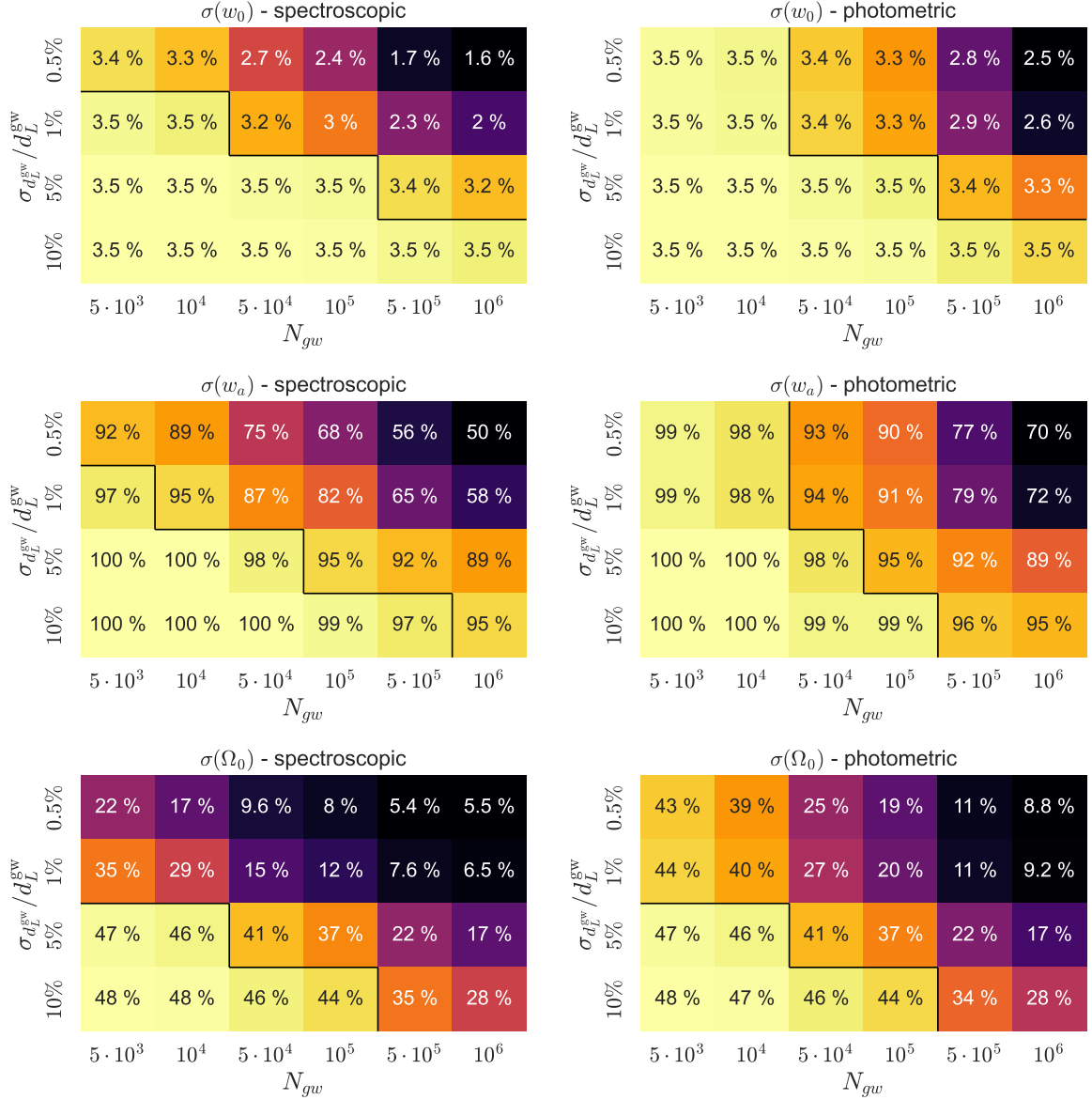


Figure 5. Marginalized 1σ confidence bounds on the MG parameters of Model I obtained from combining GW weak lensing with galaxy weak lensing and clustering. We consider bright sirens and vary the number of GW detections N_{gw} and the precision on the luminosity distance determination σ_{d_L} . Bounds are reported with the assumption that all GW sources will have a spectroscopic (left panels) or photometric (right panels) counterpart.

even with fewer GW sources, though the constraining power on cosmological parameters will be weak in comparison to the accuracy reached by next generation galaxy surveys.

6.2 Scalar-tensor models

We not turn our attention to MG models, for which the estimator (3.7) receives explicit contributions from the conformal coupling $\Omega(a)$. We wish to explore to which extent this helps in breaking degeneracies between Ω_0 and other cosmological parameters. Again, we vary

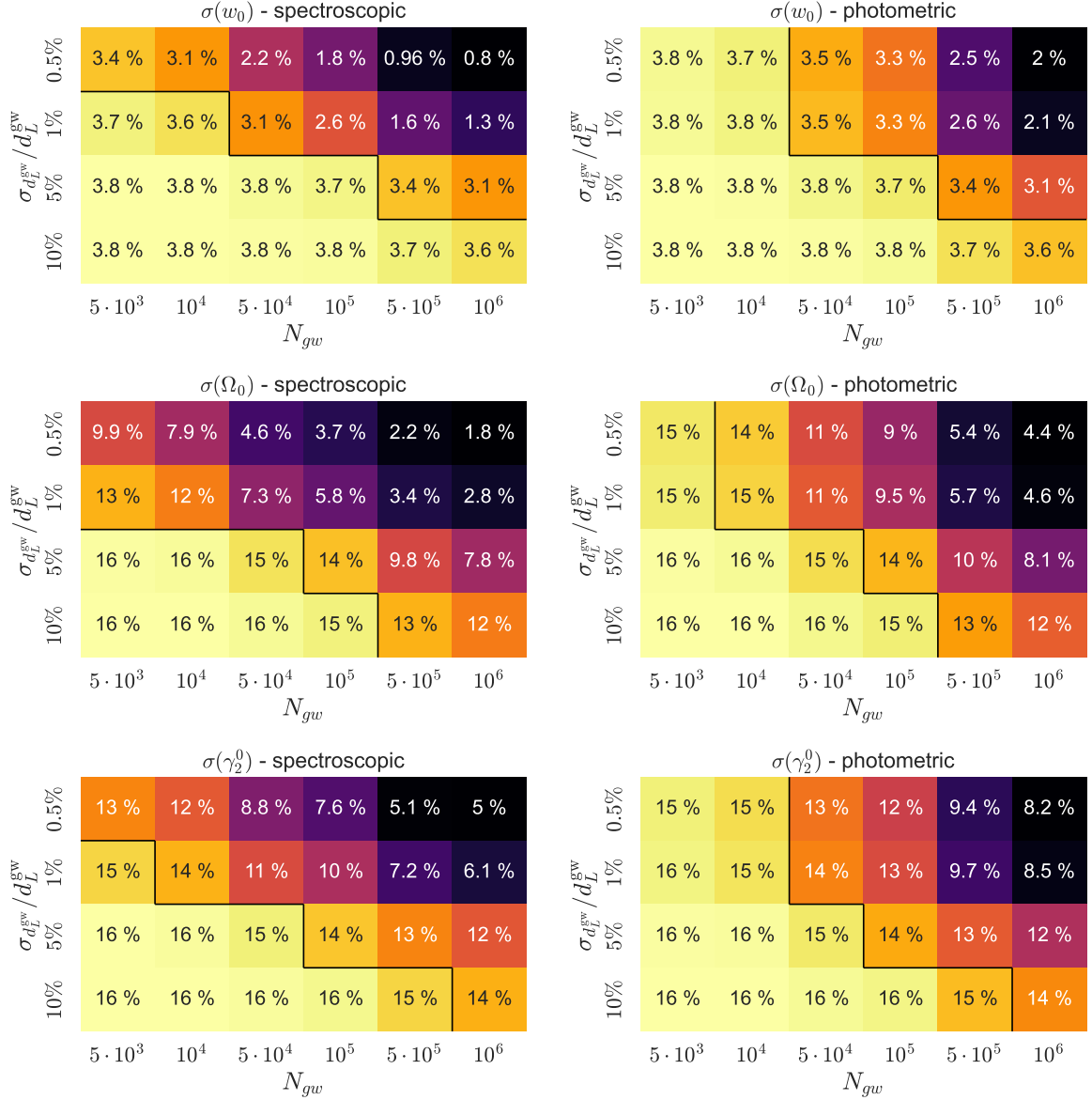


Figure 6. Marginalized 1σ confidence bounds on the MG parameters of Model II obtained from combining GW weak lensing with galaxy weak lensing and clustering. We consider bright sirens and vary the number of GW detections N_{gw} and the precision on the luminosity distance determination σ_{d_L} . Bounds are reported with the assumption that all GW sources will have a spectroscopic (left panels) or photometric (right panels) counterpart.

N_{gw} , σ_{d_L}/d_L and the error on the sources redshift determination σ_z . For each combination, we compute the Fisher forecasts on the cosmological parameters, including once again all correlators of the kind C_ℓ^{XY} , $X, Y \in [\delta_g, \kappa_g, \hat{\kappa}_{\text{gw}}]$.

In figure 5 we show the marginalized 1σ bounds on w_0 , w_a and Ω_0 in M1, for the case of photometric (top row) and spectroscopic (bottom row) redshifts. Similarly to the Λ CDM case, there are regions in which the constraining power comes almost exclusively from the galactic probes. In these regions, the number of available GW events is too low and

the accuracy of their luminosity distance determination is too poor for GW to improve on galaxies constraints. However, GWs start having an impact sooner this time, for $N_{\text{gw}} \sim 1$ order of magnitude lower than what we were observing in the Λ CDM scenario. In the case of photometric redshift, 5×10^4 GW sources determined with $\sigma_{\text{dL}} \leq 1\%$ are already sufficient to detect a tightening of the constraints on all MG parameters, that becomes significant at all σ_{dL} for $N_{\text{gw}} = 5 \times 10^5$ sources or more. The results are similar for w_0 and w_a , though for equal setups the improvement in constraining power is much lower than in the Ω_0 case.

Even more impressive are the spectroscopic results, for which at all N_{gw} there is at least one setup in which the GWs contribute to improving importantly the constraining power. In particular, with 10^4 GW events measured with $\sigma_{\text{dL}} \leq 1\%$ the constraints on Ω_0 are already almost halved. However, $\sim 5 \times 10^5$ sources are needed to reach $< 10\%$ level precision on Ω_0 .

Figure 6 shows analogous tables for w_0 , Ω_0 and γ_2^0 in the M2 case. Here, while in both photometric and spectroscopic scenarios our findings for the DE equation of state parameters are analogous to those of M1, the galaxy-dominated constraints are tighter on the EFT parameters than in the M1 case. This is in line with the findings of [135], which also finds tighter relative constraints on the EFT parameters for models in which γ_1^0 and γ_2^0 are not zero. However, the absolute bounds on Ω_0 do not change significantly between the two models: indeed as the percentage error on Ω_0 is smaller for M2 of a factor of ~ 3 , the fiducial value of Ω_0 in that model increases of approximately the same amount.

In the photometric scenario, about 10^5 sources or higher are required to reach $< 10\%$ level precision on Ω_0 , while the number drops to $N_{\text{gw}} \sim 10^4$ in the spectroscopic case. Again, spectroscopic sources perform better than photometric ones: with optimal luminosity distance determination ($\sigma_{\text{dL}} \leq 1\%$), 5×10^4 sources are enough for GW-WL to meaningfully contribute to the galaxy surveys constraints on all MG parameters. In both the photometric and spectroscopic cases, a large enough statistic of events can allow the determination of both Ω_0 and γ_2^0 at least at a few % level.

In figure 7 we report the triangular plots with marginalized constraints for the three MG parameters of M1 (left panel) and the four MG parameters of M2 (right panel). These results are obtained for an idealistic scenario in which 10^5 GW events are detected with a spectroscopic bright counterpart and their luminosity distance is determined at 1% accuracy. Besides confirming the findings reported above, the left panel of figure 7 highlights that the GW-only contours are still remarkably wide with respect to the galaxy bound, in particular concerning Ω_0 . We show this more evidently in the left panel of figure 8, where we plot the GW-WL-only constraints in gray compared to the GC+WL constraints in blue, zooming out to include the full extent of the GW bounds.

We impute these large constraints primarily to the degeneracy between w_0 and w_a that enters weak lensing self-correlations. As a result, GW-WL alone is not a good probe to constrain the DE equation of state, which in this model also determines the time evolution of the conformal coupling. The poor determination of w_{DE} leaves the coupling evolution unconstrained, resulting in very loose bounds on Ω_0 . Quite impressively though, the combined GW+galaxy constraints (dark red contours in figure 7) on the MG parameters of M1 are narrower than the galaxy-only bounds, which must mean that the increased constraining power relies heavily on the cross-correlations $C_\ell^{\kappa_{\text{gw}}\delta}$ and $C_\ell^{\kappa_{\text{gw}}\kappa_{\text{g}}}$. We verify this hypothesis in the right panel of figure 8, where we plot again in blue the marginalized galaxy bounds, this time compared with brown contours obtained including only GW-WL auto-correlation and its two cross-correlations with the galaxy fields, which we dub ‘XC’, while leaving out contributions from GC and WL auto and cross-correlations. We see that now the GW constraints become

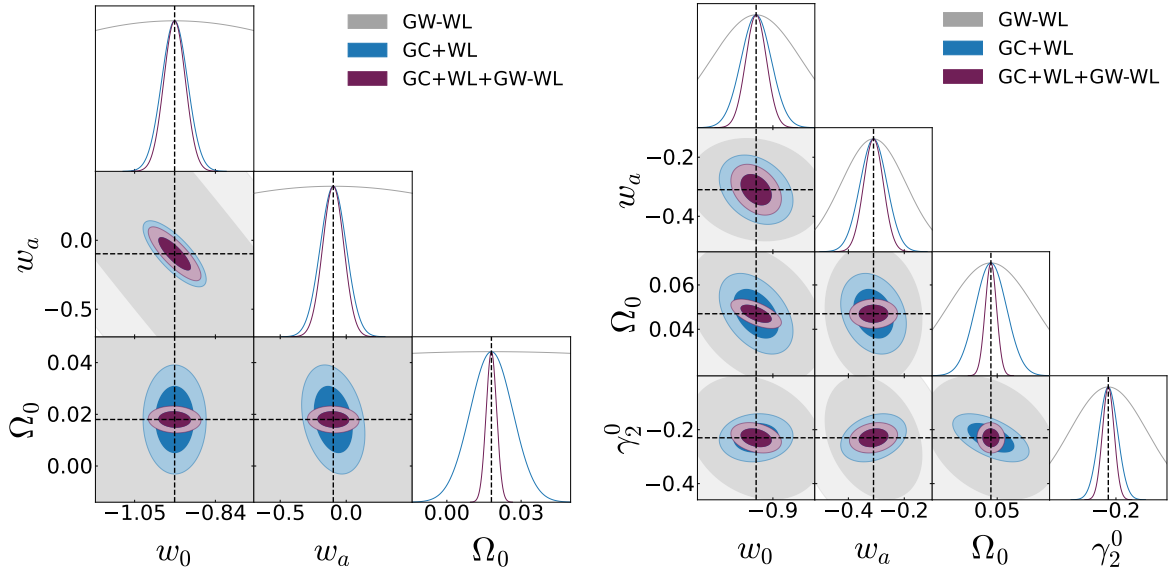


Figure 7. Marginalized forecasts for the MG parameters of model M1 (left panel) and M2 (right panel) in an idealistic scenario where $1 \cdot 10^5$ GWs events are detected with a bright counterpart measured spectroscopically and luminosity distance measured with 1% precision. Constraints come from GW weak lensing only (grey), galaxies WL and clustering (blue), and GW and galaxy probes combined, including cross-correlations between the probes (purple).

comparable with the galaxy bounds as the cross-correlation of GW-WL with, in particular, GC is breaking (or mitigating) the degeneracies in the MG sector. This confirms the essential role that GW-WL cross-correlations can have in impacting cosmological bounds.

As for model M2 (right panel in figure 7), we see that the GW-WL only contours are still pretty broad, though way tighter than in the M1 case. This is possibly because non-null γ_1 and γ_2 introduce additional EFT functions whose time evolution we also model as dependent on w_0 and w_a . If lensing is a probe overall efficient in constraining γ_2 , this might break some existing degeneracies between $\Omega(a)$ and the DE equation of state. Again, the cross-correlations play a striking role in tightening the bounds on the MG parameters, especially on Ω_0 , as the combination of GW and galaxies contributes to reducing its degeneracies with w_0 and γ_2^0 . Constraints on w_0 , w_a and γ_2^0 also improve with respect to the galaxy only case, though in the case of these parameters the effect is milder. Instead, we observe that GW-WL is not very efficient in breaking the remaining degeneracy in the (w_0, w_a) plane, as the constraints on these two parameters are still largely dominated by the galaxies. We can however deduce from figure 6 that this will eventually change if a higher number of bright GW events is considered.

Lastly, repeating the same analysis of figures 5 and 6 for the standard cosmological parameters $\{h, \Omega_{m,0}, \Omega_{b,0}, n_s, \sigma_8\}$ in both models, we find very similar results to those of subsection 6.1, i.e. $O(10^5)$ or more GW sources are needed to even to impact significantly the bounds placed on those by galaxies alone, $O(10^6)$ in the case of h and Ω_b . This is because these parameters are already determined by the galaxy probes with much more accuracy than the MG sector, and they benefit only very mildly from the increased constraining power introduced by the GWs.

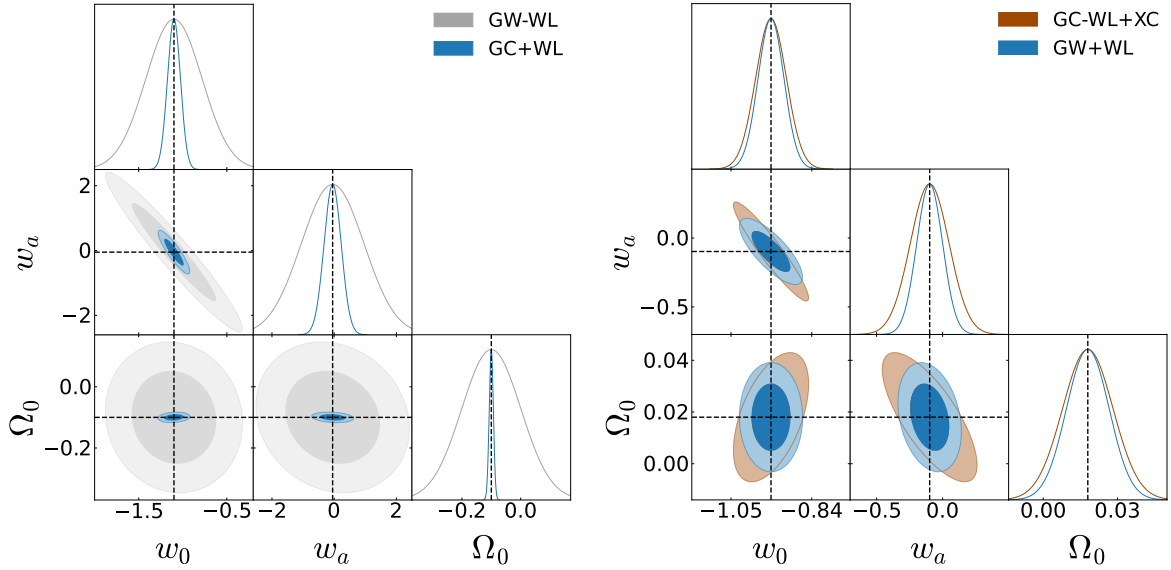


Figure 8. Comparison of the marginalized 1σ bounds on the MG parameters of M1. We consider the same setup of figure 7, i.e. $N_{\text{gw}} = 5 \cdot 10^5$, $\sigma_{d_L} = 1\%$ and spectroscopic counterparts. In both panels, the blue contours represent bounds placed through GC+WL only, matching the same blue contours of the left panel of figure 7. In the left panel, these are compared with GW-WL-only bounds, zooming out with respect to figure 7 to include the full extent of the grey contours. In the right panel instead, the blue contours are confronted with bounds obtained considering only the auto-correlation of GW-WL and its cross-correlations with galaxies (brown regions), while not including GC+WL auto and cross-correlations.

Lastly, for a better comparison of the results obtained with different setups and on the two MG models considered, we define a figure of Merit (FoM) on the line of what done in [164]:

$$\text{FoM} = \det(\mathcal{F}_{\alpha,\beta})^{\frac{1}{2N}}, \quad (6.1)$$

where $\mathcal{F}_{\alpha,\beta}$ is the Fisher matrix of eq. (5.4) marginalized over the bias and IA noise parameters, and N is the total number of parameters of the model. The FoM defined in this way is the $2N$ -th root of the product of the Fisher matrix eigenvalues. The squared root of this product is in turn inversely proportional to the volume of the N -dimensional ellipsoid delimiting the 1σ confidence region in the parameter space. The FoM thus allows us to investigate and quantify not only how the bounds on any individual parameter are affected by GW-WL, but more globally how the volume of the bounds in the parameter space is reduced. We compute the FoM for the Fisher matrices (5.4) of models M1 and M2, including all auto and cross-correlations of GW-WL, GC and galaxy WL and for all scenarios explored above in terms of N_{gw} and σ_{d_L} . We also considered separately the cases in which GW events redshifts are determined photometrically or spectroscopically. The results are plotted in figure 9.

We notice that the FoM confirms all of our findings above and more. In the photometric cases, for $N_{\text{gw}} \leq 10^4$ the overall improvement of bounds in the parameter space is only very mild. However, the FoM start increasing significantly from $N_{\text{gw}} = 5 \times 10^4$ or higher, for all choices of σ_{d_L} though significantly better performances occur for $\sigma_{d_L} \leq 1\%$. We also notice that there is only a small difference in the FoM curves for the configurations with 1% and 0.5% average uncertainties on the luminosity distance. This suggests that the photometric redshift error is starting to dominate the correlation noise in eq. (5.9), thus further improving

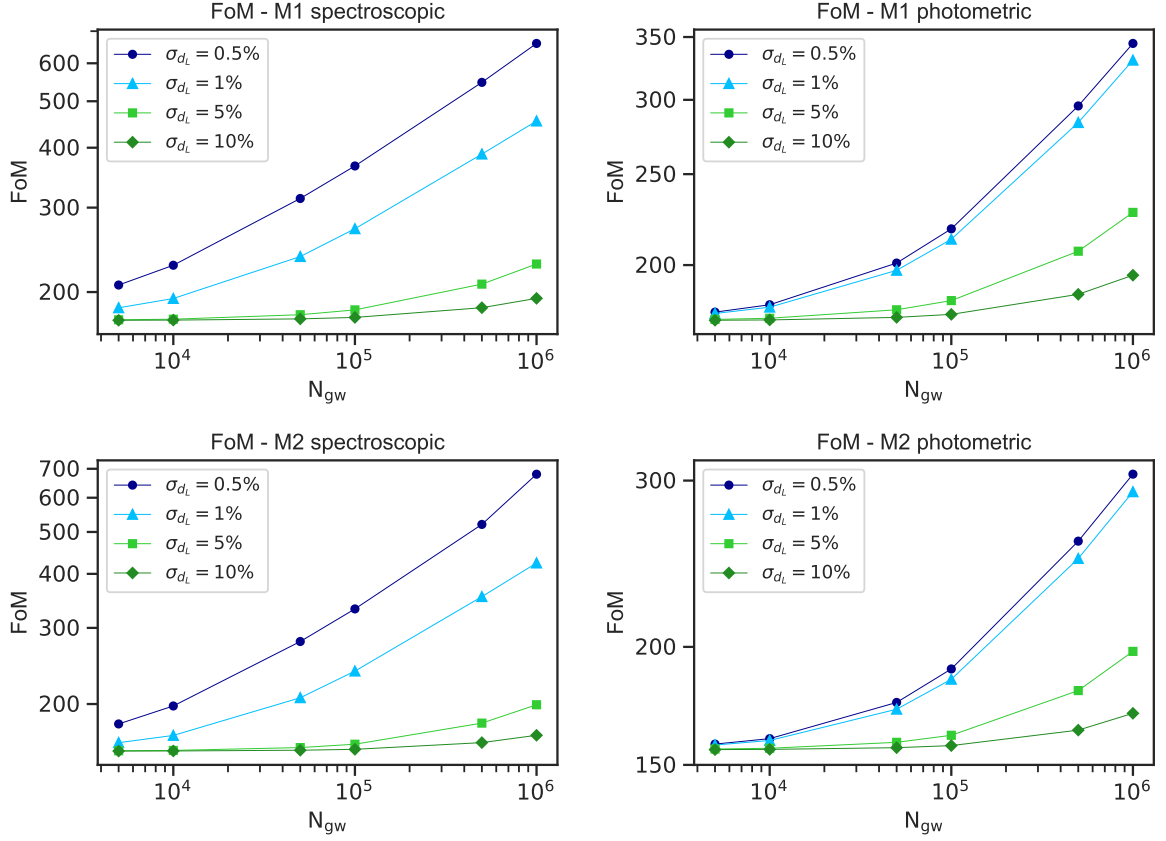


Figure 9. Figures of merit for different numbers of total detected events N_{gw} and several choices of the percentage luminosity distance uncertainty σ_{d_L} . We plot the FoM for model M1 in the top row and M2 in the bottom row, while in the left and right panels are grouped the scenarios with spectroscopic and photometric counterparts respectively.

the d_L measurement of photometric events beyond the 1% accuracy level only implies a minor increase in the GW-WL constraining power: the real game-changer becomes in this case the number of available sources.

On the contrary, because the spectroscopic redshift error is much lower, for spectroscopic events the correlation noise remains dominated by σ_{d_L} , and improvements to the FoM happen for every choice of σ_{d_L} . We observe, however, that for $\sigma_{d_L} \geq 5\%$ the values of the FoM of both photometric and spectroscopic scenarios are comparable. Only for lower percentage errors on d_L the spectroscopic FoM are higher than the photometric ones, almost doubling their value if the luminosity distance can be measured with sub-percent accuracy.

In summary, we conclude that GW-WL is a powerful probe to measure MG parameters and can in some cases outperform the measurements from the upcoming galaxy surveys. Moreover, GWs provide an independent measurement of the MG parameters, affected by completely different systematics and uncertainties, which can shed new lights on the physics of late-time cosmology.

7 Discussion and conclusions

We have investigated the synergy of the next generation of galaxy surveys and future GW observatories in constraining gravity on cosmological scales, with a focus on the role of weak lensing of gravitational waves. The latter represents a probe complementary to galaxy surveys but characterized by different systematics and thus able to contribute to cosmological constraints in an independent way.

At redshifts higher than $z \sim 0.5$, WL is the dominant contribution to the (linear) corrections to the luminosity distance of GWs induced by large-scale structure; building on this, we have constructed an estimator for the lensing convergence that in the case of bright GW events receives contributions from the conformal coupling of scalar-tensor theories $\Omega(z)$. In the Λ CDM case, where $\Omega(z) = 0$, GW lensing and galaxies eccentricities probe the same quantity and the contribution of GW is primarily that of increasing the WL statistics.

We have explored extensively cross-correlations between $\hat{\kappa}_{\text{gw}}$ and standard galaxy density fields δ_g and galaxy weak lensing κ_g , with the aim of determining the observational requirements for future GW detectors to make stringent constraints on late time cosmological models, reaching beyond the limits that will be placed by the next generation of galaxy surveys. To this extent, we did not restrict to any particular GW experiment, rather we considered several scenarios varying the total number of detected GW sources N_{gw} , the accuracy of such detection in terms of luminosity distance σ_{d_L} and the uncertainty on the estimation of the redshift of the GW event. We have considered separately, and compared, the case in which the GW counterparts (broadly including identified galaxy hosts) can be observed with a photometric survey and an optimistic case in which all redshifts can be determined spectroscopically.

As expected, we find that in the Λ CDM case GW-WL starts to improve on galaxy constraints only once a very high number of events is observed. We find that $O(10^5)$ detections are needed to start noticing an impact even on parameters such as σ_8 and Ω_m to which GW weak lensing is most sensitive, and $O(5 \times 10^5)$ for this impact to become meaningful. For ST theories with a non-zero conformal coupling, we show that introducing the cross-correlators of $\hat{\kappa}_{\text{gw}}$ with other probes can strongly break degeneracies among cosmological parameters, already impacting the constraints on non- Λ CDM parameters for a relatively small number of detected events. The order of magnitude of the necessary N_{gw} depends again on the accuracy of the luminosity distance measurement, but also non negligibly on the redshift determination and the chosen ST model.

The GW contribution can have a great impact on models in which the time evolution of EFT functions such as the non-minimal coupling is not fixed. These models, such as M1 explored in this work, tend to be more difficult to constrain with galaxies alone, and we find that if d_L^{gw} can be measured with an accuracy $\leq 1\%$, then $O(1 - 5 \times 10^4)$ GW photometric events are already enough for GW-WL to start improving on galaxy constraints when cross-correlated with the galaxy observables. This contribution becomes significant when a total of 10^5 sources, or higher, are detected. When spectroscopic counterparts are considered, the required N_{gw} reduces by about one order of magnitude, with $O(10^{4-5})$ sources already doubling the accuracy on the EFT parameters, according to the accuracy on the measurement of d_L .

For MG models with all three EFT functions open, such as our M2, we find that the addition of GW-WL to galaxy probes gives, in both spectroscopic and photometric case, results qualitatively similar to M1. We find that in the spectroscopic case GW-WL strongly contributes to reducing degeneracies in the MG sector already when $O(10^{4-5})$ are measured, again depending on the accuracy over the luminosity distance of such sources.

To put the presented numbers in perspective, LISA’s most optimistic estimates [165] foresee only a few hundreds of GW events over five years of observation, of which only a few tens with a bright counterpart, while recent estimates for a network combining ET and one or two CE detectors place the number of observed BNS at a few tens (or even hundreds) of thousands of events per year [10, 92, 93]. Similarly NS-BH systems are also expected to present electromagnetic counterparts [166, 167]. Based on the event rate currently measured by LVK [4], it is possible to forecast an NS-BH total rate of a few thousand events per year with the next-generation detectors. Over the span of 10 years of observations, up to $\sim 10^{5-6}$ GW events could be detected in total. Therefore, if these events can be associated with the detection of an electromagnetic counterpart, they will potentially be the best candidates to constrain late time cosmology through GW-WL. We must however remark that, according to the cited estimates, third generation ground based detectors are not expected to reach the average $\sim 1\%$ accuracy on the luminosity distance determination required by our more promising setups (see e.g. figures 5 and 6), likely not even operating as a network. The forecasted error on luminosity distance for a fraction of BNS is expected to be around 10% (slightly lower for nearby sources, but higher for most events) [92, 93]. The error bar on the luminosity distance for NS-BHs is estimated to be similar or slightly better (due to higher chirp mass and the relevance of higher-order modes due to higher mass ratios) [168, 169]. An average accuracy over d_L of $\sim 10\%$ could still lead to improvements on the galaxy-only cosmological constraints, but in that case we found that $O(10^6)$ GW bright sources are required. This number is likely beyond the capabilities of 3G ground-based detectors, as it is forecasted that in practice it will be possible to associate a counterpart only to a small fraction ($\sim 10^{3-4}$) of all ET and CE events [92], greatly reducing the available statistics. Nonetheless, independent measurements of cosmological parameters through GW-WL can still be carried out with the events from ET and CE, though the reached precision will not be competitive with that of galaxy surveys.

On the bright side, our forecasts could in principle be further improved by considering the cross-correlation signal up to higher $\ell_{\max}$, provided that a solid method to treat non-linearities is available, and by exploiting cosmic variance cancellation techniques. Hence, the chance that ET+2CE sources can contribute meaningfully in constraining cosmologies through the cross-correlations of GW-WL with galaxies is not yet completely ruled out, and deserves further investigation.

Moreover, stacking a large statistics of highly accurate GW events will be possible with proposed far future observatories like the space-based Big Bang Observer (BBO), DECi-hertz Interferometer Gravitational-wave Observatory (DECIGO) and Advanced Laser Interferometer Antenna (ALIA). These experiments are expected to reach sub-percent precision in the determination of the luminosity distance and an angular resolution of orders of fractions of degrees for ALIA [96] and arcseconds for BBO [95], with a total number of detected events of several hundreds of thousands. Such a high angular resolution will have the additional advantage of facilitating significantly the task of pinpointing host galaxies, and hence increase the fraction of available bright events.³

In conclusion, we find that the cross-correlations of galaxies and GW-WL have the potential to become, with time, crucial probes of cosmology and modified gravity, complemen-

³Note that the number of detected direct EM counterparts is limited, among other things, by the finite observational time that EM surveys can dedicate to the follow-up of GW detections: at present it seems unrealistic that orders of millions of detections per year can be efficiently followed up individually. Hence, host identification will become particularly important in the era of precision cosmology with GW.

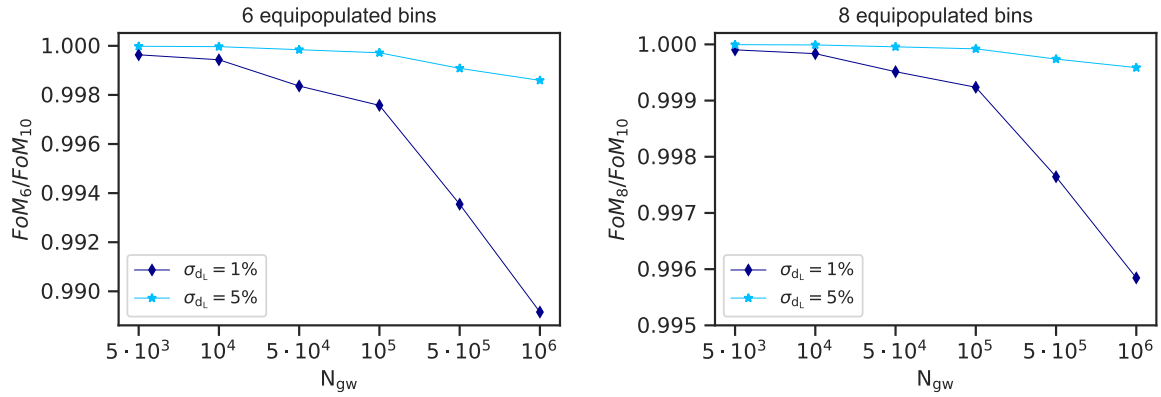


Figure 10. Ratio $\text{FoM}_{[6,8]}/\text{FoM}_{10}$ of the FoM computed binning the GW sources in 10 equipopulated tomographic bins and the FoM obtained for the same configurations in terms of number of GW sources N_{gw} and percentage uncertainty on the luminosity distance σ_{d_L} , but with a different number of bins (6 and 8 in the left and right panel respectively). The results regard Model I.

tary to galaxy surveys and other cosmological observables. Where sufficient statistics is available, these new probes can both help tighten constraints and make an independent measurement on some Λ CDM parameters, and strongly reduce existing degeneracies in the MG sector.

While in this paper we have focused on GW weak lensing, GW measurements of the luminosity distance-redshift relation, exploiting both bright ([40, 41, 43, 165, 170, 171]) and dark ([59, 61, 63, 64, 172]) events, will provide tight bounds on parameters affecting the cosmic expansion history. This potentially helps in breaking some degeneracies between background and growth parameters as measured by GW-WL, as explored in [57, 58]. These works have shown that the consequence of adding the standard sirens contributions is to effectively tighten the constraints on some of the parameters to which only GW-WL is sensitive. In this perspective, even with lower statistics GW-WL can play an essential role when one wishes to exploit exclusively GW observations to provide cosmological constraints that are independent of EM probes. We will explore the interplay of standard sirens with galaxy surveys and GW-WL applied to constrain ST theories in a companion work, currently in preparation [173].

Acknowledgments

We thank Marco Raveri and Fabrizio Renzi for useful discussions. AB is supported by a de Sitter Fellowship of the Netherlands Organization for Scientific Research (NWO). AG acknowledges support from the NWO and the Dutch Ministry of Education. M.M. acknowledges funding by the Agenzia Spaziale Italiana (ASI) under agreement no. 2018-23-HH.0 and support from INFN/Euclid Sezione di Roma. The work of SM is a part of the `<data|theory>` **Universe-Lab** which is supported by the TIFR and the Department of Atomic Energy, Government of India. AS acknowledges support from the NWO and the Dutch Ministry of Education, Culture and Science (OCW) (grant VI.Vidi.192.069).

A Impact of GW binning

We address here the impact that our tomographic binning choice for the GW sources has on the final results, investigating if the additional constraining power brought by GW-WL and its cross correlation with galaxy survey is affected changing the number of bins.

We choose the figures of merit (FoM) defined in section 6 as indicators of such constraining power, and we compute the FoM for several observational configurations, characterized by the total number of detected GW events with a spectroscopic counterpart ($N_{\text{gw}} \in [5 \times 10^3, 10^6]$) and by the average percentage uncertainty on the GW luminosity distance ($\sigma_{d_L} \in [1\%, 5\%]$). We repeat our computations binning the GW sources respectively in 6, 8 and 10 equipopulated tomographic bins. In figure 10, we plot the ratio $\text{FoM}_n/\text{FoM}_{10}$ for all observational scenarios, with subscripts indicating the number of bins and $n = (6, 8)$ in the left and right panel respectively. We use the 10 bins case as a reference, as it is the most optimistic scenario and also the one we chose to present our results in the main body of this work. In figure 10 we present the FoM for Model I. We checked that similar results still hold for Λ CDM and Model II.

We observe that, reducing the numbers of bins in all cases implies lower FoM, which in turns signifies a deterioration of the constraints on cosmological parameters. This happens because decreasing the bins number, and thus increasing the bins size, has the effect of smoothing out the correlation signal in any given bin. However, this effect is mostly compensated by the augmented number of events falling in each bin, that attenuates the correlation noise of eq. (5.9). Indeed, we observe that in all scenarios the FoM depends only very mildly on the number of bins. Higher deviations occur for higher numbers of detected events and better accuracy on the luminosity distance measurement, though we notice that even in the most extreme case of 6 tomographic bins and 10^6 GW events with d_L detected at 1% accuracy, FoM_6 is lower than FoM_{10} only of $\sim 1\%$.

Consequently, we can conclude that our choice of the number of bins for the GW sources impacts only very mildly the overall constraining power on cosmological parameters gained through GW-WL, and does not affect significantly the results presented here or our conclusions.

B Impact of GW sources redshift distribution

In section 5, we made a choice about the assumed redshift distribution of sources of bright gravitational waves events. This redshift distribution results, in general, from a convolution of the underlying distribution of merging binaries in the Universe below $z = 2.5$, with the selection function of the GW detector considered for their observation and an additional selection function dependent on the redshift identification method. This last function accounts for the fact that it will in practice be impossible to associate a redshift to all of the observed binaries, and the number of bright sources will differ if the redshift inference is based on a direct counterpart rather than on the identification of a host within a galaxy catalogue. Moreover, in the case of direct counterparts, the final redshift distribution of bright sources also depends on the considered channels for the EM counterpart emission.

Various combinations of all different possibilities generate a mosaic of viable observational scenarios. To preserve a certain degree of generality, in the main body of this work we decided not to make any direct assumption on the observational setup concerning the redshift. Rather, we took a phenomenological approach, modelling the number count of bright GW sources as a function of redshift as

$$\frac{dn_{\text{gw}}}{dz} \propto \left(\frac{z}{z_0}\right)^2 \exp \left[- \left(\frac{z}{z_0}\right)^{3/2} \right]. \quad (\text{B.1})$$

with $z_0 = 1.5$. This distribution was chosen to match the population of injected BNS to be observed by a network of ET and two CE detectors in [92]. Observationally, this choice is closest to a particularly favorable scenario: it considers BNS - which are expected to be the most numerous binary population observable in future GW detectors - and it assumes that almost all merging binaries can be (i) detected by the GW observatory and (ii) matched with a redshift measurement. Condition (i) is close to what happens in a 4th generation GW detector with very low signal-to-noise threshold, such as e.g. BBO [95], while (ii) is likely to be satisfied observationally for high numbers of sources only through host identification, again in principle possible thanks to the high accuracy of 4th generation GW detectors.

In the rest of this appendix, we characterize the impact of these assumptions on the derived cosmological constraints. We do so, by looking at how the constraints change if the underlying GW source distribution is chosen to peak at progressively lower redshifts, i.e. we considered two alternative scenarios with $z_0 = 1.0$ and $z_0 = 0.5$. The former is observationally closest to a more realistic 4th-generation detector situation: the distribution again assumes that the low SNR of the GW events in the detector is such that a consistent statistic of sources can be detected up to relatively high redshifts. However, it accounts for the fact that it is more difficult to associate hosts to high redshift events ($z \geq 1.5$), both because the sky-localization of the GW event partially deteriorates at high z and because the galaxy catalogues become progressively less complete. Finally, the case with $z_0 = 0.5$ is closest to what could realistically be expected for 3rd-generation detectors, such as ET and CE: the SNR threshold, especially if a network of detectors is considered, is low enough that many events can be detected up to $z \sim 2.5$. In turn, the sky-location is so poor that individual host identification is impossible for most of the events. Still, BNS are expected to emit direct EM counterparts, and for 3G experiments that is likely to be the primary channel of redshift measurement, with direct counterparts being detectable only at relatively low redshifts. Astrophysical models concerning EM emission and detection vary significantly, though many current forecasts place the peak of measurable bright events around $z \sim 0.5 - 0.7$, with the total number of events dropping quickly for $z > 1.0$ [10, 92, 93].

Our three different choices for the source distributions are compared in the left panel of figure 11. The right panel shows instead the evolution of the figures of merit (FoM) of our Model I defined in section 6.2, as a function of the total number of observed events and for different choices of z_0 (solid, dashed and dotted lines) and different average uncertainty on the luminosity distance ($\sigma_{d_L} = 1\%$ for dark blue lines and $\sigma_{d_L} = 10\%$ for light blue lines). The FoM for Model II are not shown, as we find no evident qualitative difference with respect to Model I.

We notice that the FoMs for the cases $z_0 = 1.5$ and $z_0 = 1.0$ are still very similar, and there is only a mild deterioration of the overall bounds in the latter case. This can be interpreted in the light of our findings of section 6.2, i.e. that most of the increase in constraining power obtained when adding GW to the LSS does not come from the self-correlation of GW-WL, but from the cross-correlation of GWs with the galaxies. These cross-correlations are the highest when there is a significant overlap between the redshift distributions of GW binaries and galaxies, which is the case for both $z_0 = 1.5$ and $z_0 = 1.0$. The case $z_0 = 1.5$ remains however slightly more performing, as it presents higher statistics of GW sources at high z , which enhances the constraining power of the GW-WL self-correlations terms. On the contrary, the distribution with $z_0 = 0.5$ has very few sources at $z \geq 1.0$, thus reducing the overlap with the galaxy distribution and affecting the constraining power of the cross-correlations terms. The overall cosmological constraints are looser as a result.

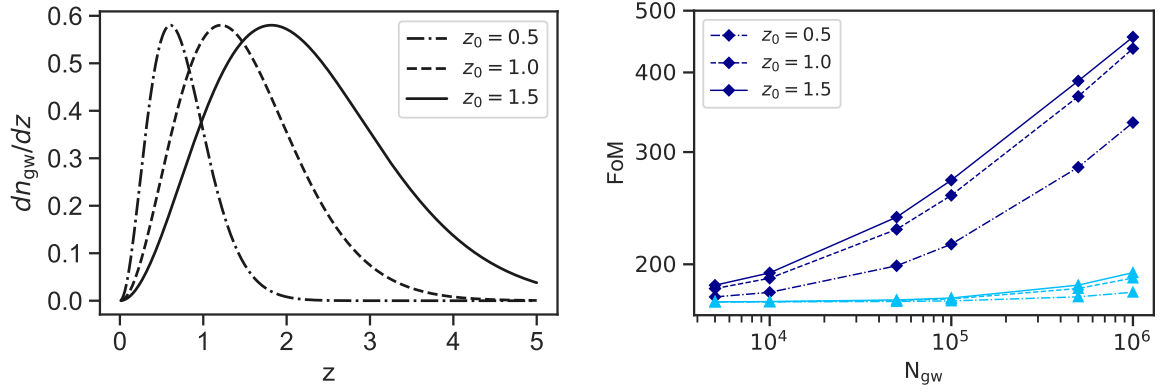


Figure 11. Left panel: comparison between three different choices for the normalized redshift distribution of gravitational wave sources. All distributions assume eq. (B.1) with $z_0 = 1.5$ (solid line), $z_0 = 1.0$ (dashed line) and $z_0 = 0.5$ (dash-dotted line). Right panel: FoM as a function of the total number of bright GW events, plotted for different choices of z_0 and assuming $\sigma_{d_L} = 1\%$ (dark blue diamonds) or $\sigma_{d_L} = 10\%$ (light blue triangles).

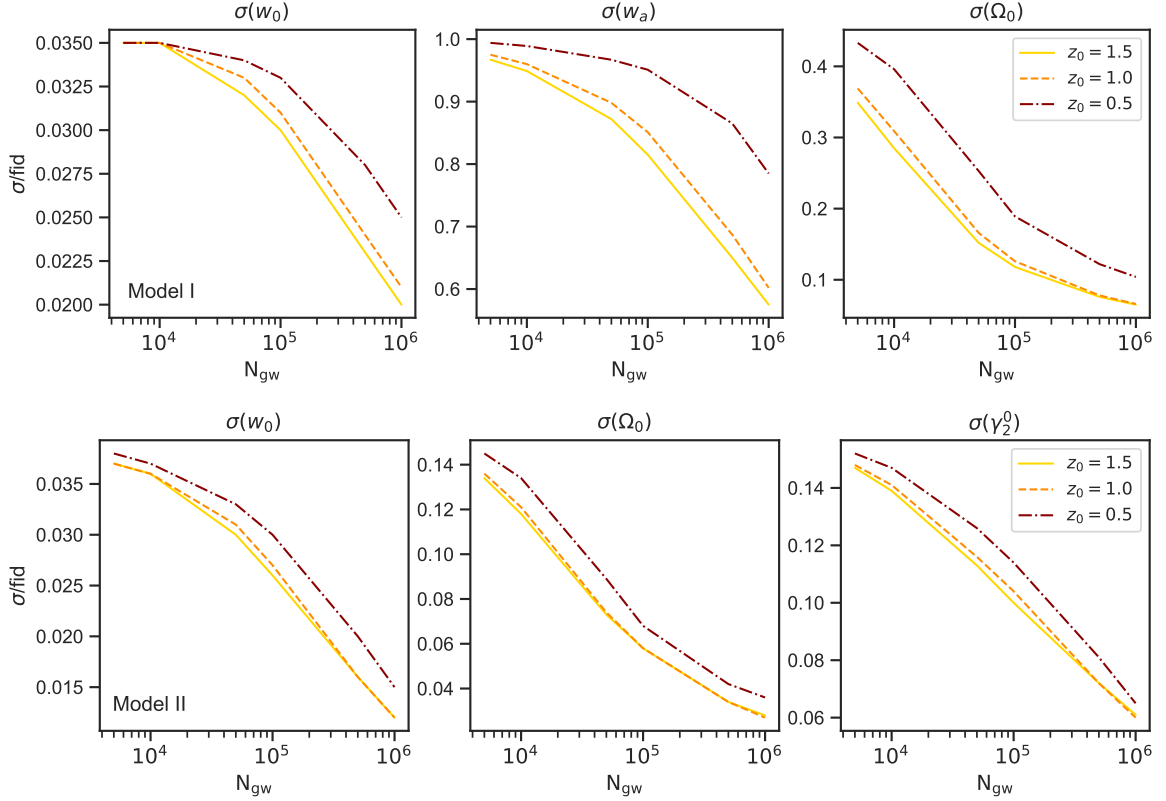


Figure 12. Relative 1σ bounds on different MG parameters θ_i computed as $\sigma(\theta_i)/\theta_{i,\text{fid}}$ as a function of the total number of bright GW sources, for Model I (top row) and Model II (bottom row). Solid, dashed and dash-dotted lines correspond to $z_0 = 1.5$, 1.0 and 0.5 respectively.

In figure 12, we show the evolution of the relative 1σ bounds ($\sigma(\theta_\alpha)/\theta_{\alpha,\text{fid}}$) on the individual MG parameters of Model I (top row) and Model II (bottom row) as a function of

the total number of GW events. Bounds are plotted here assuming $\sigma_{dL} = 1\%$ and we plot in different colours and line styles the results for different assumptions on the source distribution. We notice that, in general, the impact of the source distribution is more prominent for Model I than for Model II. However, in both cases, there are only minor differences in the individual bounds between the choices $z_0 = 1.5$ and $z_0 = 1.0$, while the case $z_0 = 0.5$ differs more significantly, according to the specific parameter considered. For example, in Model I, choosing a certain source distribution can lead to a variation in the bounds on w_0 of the order of less than 1%, while bounds on w_a and Ω_0 deteriorate faster, with differences in the relative bounds of the order of 10%.

We also notice a variation in the trend of these curves depending on the parameters considered. For example, in the case of w_0 and w_a for both Model I and II the difference between the different curves is way smaller at lower N_{gw} than it is at higher N_{gw} . This is likely due to the fact that, as we showed in section 6, for a low number of GW sources cosmological bounds on the DE equation of state are still highly dominated by the galaxies, while the GW contribution becomes more and more important as the number of sources increases. However, for Ω_0 and γ_2^0 we notice instead that the difference among the curve is the same approximately at any N_{gw} , even diminishing slightly as N_{gw} gets bigger. This is consistent with figure 5 and 6, as we notice that when $\sigma_{dL} = 1\%$, GW contribute significantly to constraining those two parameters at all N_{gw} . The curves in figure 12 seems to hint that, for the bounds on these parameters in particular, the loss of high redshift sources has not a very severe impact, and is on the contrary partially compensated by the increased statistics at low redshifts.

In conclusion, we find that the choice of the underlying GW source distribution can have a non-negligible effect on cosmological bounds, as could be expected. While some of the results presented in the main body of this work correspond to an optimistic observational scenario, we find that considering a more realistic source distribution in the context of future 4th-generation detectors ($z_0 = 1.0$) does not affect our results in a relevant way. On the contrary, considering a redshift distribution of bright GW events closer to realistic expectations for 3rd-generation GW detectors leads to a non-negligible deterioration of the cosmological constraints. This further validates our previous conclusion that GW-WL is unlikely to become a probe competitive with LSS with the next generation of GW detectors, while it remains potentially a promising cosmological probe in the farther future.

References

- [1] LIGO SCIENTIFIC and VIRGO collaborations, *Observation of gravitational waves from a binary black hole merger*, *Phys. Rev. Lett.* **116** (2016) 061102 [[arXiv:1602.03837](#)] [[INSPIRE](#)].
- [2] LIGO SCIENTIFIC and VIRGO collaborations, *GWTC-1: a gravitational-wave transient catalog of compact binary mergers observed by LIGO and Virgo during the first and second observing runs*, *Phys. Rev. X* **9** (2019) 031040 [[arXiv:1811.12907](#)] [[INSPIRE](#)].
- [3] LIGO SCIENTIFIC and VIRGO collaborations, *GWTC-2: compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run*, *Phys. Rev. X* **11** (2021) 021053 [[arXiv:2010.14527](#)] [[INSPIRE](#)].
- [4] LIGO SCIENTIFIC et al. collaborations, *GWTC-3: compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run*, [arXiv:2111.03606](#) [[INSPIRE](#)].

- [5] LIGO SCIENTIFIC and VIRGO collaborations, *GW170817: observation of gravitational waves from a binary neutron star inspiral*, *Phys. Rev. Lett.* **119** (2017) 161101 [[arXiv:1710.05832](#)] [[INSPIRE](#)].
- [6] LIGO SCIENTIFIC et al. collaborations, *Gravitational waves and gamma-rays from a binary neutron star merger: GW170817 and GRB 170817A*, *Astrophys. J. Lett.* **848** (2017) L13 [[arXiv:1710.05834](#)] [[INSPIRE](#)].
- [7] LIGO SCIENTIFIC et al. collaborations, *Multi-messenger observations of a binary neutron star merger*, *Astrophys. J. Lett.* **848** (2017) L12 [[arXiv:1710.05833](#)] [[INSPIRE](#)].
- [8] KAGRA collaboration, *Overview of KAGRA: KAGRA science*, [arXiv:2008.02921](#) [[DOI:10.1093/ptep/ptaa120](#)] [[INSPIRE](#)].
- [9] LISA collaboration, *Laser Interferometer Space Antenna*, [arXiv:1702.00786](#) [[INSPIRE](#)].
- [10] M. Maggiore et al., *Science case for the Einstein Telescope*, *JCAP* **03** (2020) 050 [[arXiv:1912.02622](#)] [[INSPIRE](#)].
- [11] D. Reitze et al., *Cosmic explorer: the U.S. contribution to gravitational-wave astronomy beyond LIGO*, *Bull. Am. Astron. Soc.* **51** (2019) 035 [[arXiv:1907.04833](#)] [[INSPIRE](#)].
- [12] LIGO SCIENTIFIC et al. collaborations, *A gravitational-wave standard siren measurement of the Hubble constant*, *Nature* **551** (2017) 85 [[arXiv:1710.05835](#)] [[INSPIRE](#)].
- [13] ASTRO-WISE and KiDS collaborations, *The Kilo-Degree Survey*, *Exper. Astron.* **35** (2013) 25 [[arXiv:1206.1254](#)] [[INSPIRE](#)].
- [14] E. van Uitert et al., *KiDS+GAMA: cosmology constraints from a joint analysis of cosmic shear, galaxy-galaxy lensing, and angular clustering*, *Mon. Not. Roy. Astron. Soc.* **476** (2018) 4662 [[arXiv:1706.05004](#)] [[INSPIRE](#)].
- [15] DES collaboration, *The Dark Energy Survey: more than dark energy — an overview*, *Mon. Not. Roy. Astron. Soc.* **460** (2016) 1270 [[arXiv:1601.00329](#)] [[INSPIRE](#)].
- [16] DES collaboration, *Dark Energy Survey year 3 results: cosmological constraints from galaxy clustering and weak lensing*, *Phys. Rev. D* **105** (2022) 023520 [[arXiv:2105.13549](#)] [[INSPIRE](#)].
- [17] EUCLID collaboration, *Euclid definition study report*, [arXiv:1110.3193](#) [[INSPIRE](#)].
- [18] LSST SCIENCE and LSST PROJECT collaborations, *LSST science book, version 2.0*, [arXiv:0912.0201](#) [[INSPIRE](#)].
- [19] LSST DARK ENERGY SCIENCE collaboration, *The LSST Dark Energy Science Collaboration (DESC) science requirements document*, [arXiv:1809.01669](#) [[INSPIRE](#)].
- [20] LSST collaboration, *LSST: from science drivers to reference design and anticipated data products*, *Astrophys. J.* **873** (2019) 111 [[arXiv:0805.2366](#)] [[INSPIRE](#)].
- [21] D. Spergel et al., *Wide-Field InfrarRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 report*, [arXiv:1503.03757](#) [[INSPIRE](#)].
- [22] E. Abdalla et al., *Cosmology intertwined: a review of the particle physics, astrophysics, and cosmology associated with the cosmological tensions and anomalies*, *JHEAp* **34** (2022) 49 [[arXiv:2203.06142](#)] [[INSPIRE](#)].
- [23] L. Perivolaropoulos and F. Skara, *Challenges for Λ CDM: an update*, *New Astron. Rev.* **95** (2022) 101659 [[arXiv:2105.05208](#)] [[INSPIRE](#)].
- [24] E. Di Valentino et al., *In the realm of the Hubble tension — a review of solutions*, *Class. Quant. Grav.* **38** (2021) 153001 [[arXiv:2103.01183](#)] [[INSPIRE](#)].
- [25] A. Silvestri and M. Trodden, *Approaches to understanding cosmic acceleration*, *Rept. Prog. Phys.* **72** (2009) 096901 [[arXiv:0904.0024](#)] [[INSPIRE](#)].

- [26] T. Clifton, P.G. Ferreira, A. Padilla and C. Skordis, *Modified gravity and cosmology*, *Phys. Rept.* **513** (2012) 1 [[arXiv:1106.2476](#)] [[INSPIRE](#)].
- [27] C. Cheung et al., *The effective field theory of inflation*, *JHEP* **03** (2008) 014 [[arXiv:0709.0293](#)] [[INSPIRE](#)].
- [28] P. Creminelli, G. D’Amico, J. Norena and F. Vernizzi, *The effective theory of quintessence: the $w < -1$ side unveiled*, *JCAP* **02** (2009) 018 [[arXiv:0811.0827](#)] [[INSPIRE](#)].
- [29] M. Park, K.M. Zurek and S. Watson, *A unified approach to cosmic acceleration*, *Phys. Rev. D* **81** (2010) 124008 [[arXiv:1003.1722](#)] [[INSPIRE](#)].
- [30] J.K. Bloomfield and E.E. Flanagan, *A class of effective field theory models of cosmic acceleration*, *JCAP* **10** (2012) 039 [[arXiv:1112.0303](#)] [[INSPIRE](#)].
- [31] G. Gubitosi, F. Piazza and F. Vernizzi, *The effective field theory of dark energy*, *JCAP* **02** (2013) 032 [[arXiv:1210.0201](#)] [[INSPIRE](#)].
- [32] F. Piazza and F. Vernizzi, *Effective field theory of cosmological perturbations*, *Class. Quant. Grav.* **30** (2013) 214007 [[arXiv:1307.4350](#)] [[INSPIRE](#)].
- [33] G.W. Horndeski, *Second-order scalar-tensor field equations in a four-dimensional space*, *Int. J. Theor. Phys.* **10** (1974) 363 [[INSPIRE](#)].
- [34] *Euclid webpage*, <https://www.cosmos.esa.int/web/euclid>.
- [35] *Vera C. Rubin observatory webpage*, <https://www.lsst.org>.
- [36] L. Amendola et al., *Cosmology and fundamental physics with the Euclid satellite*, *Living Rev. Rel.* **21** (2018) 2 [[arXiv:1606.00180](#)] [[INSPIRE](#)].
- [37] E. Belgacem, Y. Dirian, S. Foffa and M. Maggiore, *Gravitational-wave luminosity distance in modified gravity theories*, *Phys. Rev. D* **97** (2018) 104066 [[arXiv:1712.08108](#)] [[INSPIRE](#)].
- [38] L. Amendola, I. Sawicki, M. Kunz and I.D. Saltas, *Direct detection of gravitational waves can measure the time variation of the Planck mass*, *JCAP* **08** (2018) 030 [[arXiv:1712.08623](#)] [[INSPIRE](#)].
- [39] A. Nishizawa, *Generalized framework for testing gravity with gravitational-wave propagation. I. Formulation*, *Phys. Rev. D* **97** (2018) 104037 [[arXiv:1710.04825](#)] [[INSPIRE](#)].
- [40] E. Belgacem, Y. Dirian, S. Foffa and M. Maggiore, *Modified gravitational-wave propagation and standard sirens*, *Phys. Rev. D* **98** (2018) 023510 [[arXiv:1805.08731](#)] [[INSPIRE](#)].
- [41] LISA COSMOLOGY WORKING GROUP collaboration, *Testing modified gravity at cosmological distances with LISA standard sirens*, *JCAP* **07** (2019) 024 [[arXiv:1906.01593](#)] [[INSPIRE](#)].
- [42] S. Mukherjee, B.D. Wandelt and J. Silk, *Testing the general theory of relativity using gravitational wave propagation from dark standard sirens*, *Mon. Not. Roy. Astron. Soc.* **502** (2021) 1136 [[arXiv:2012.15316](#)] [[INSPIRE](#)].
- [43] T. Baker and I. Harrison, *Constraining scalar-tensor modified gravity with gravitational waves and large scale structure surveys*, *JCAP* **01** (2021) 068 [[arXiv:2007.13791](#)] [[INSPIRE](#)].
- [44] K. Leyde et al., *Current and future constraints on cosmology and modified gravitational wave friction from binary black holes*, *JCAP* **09** (2022) 012 [[arXiv:2202.00025](#)] [[INSPIRE](#)].
- [45] A. Finke et al., *Cosmology with LIGO/Virgo dark sirens: Hubble parameter and modified gravitational wave propagation*, *JCAP* **08** (2021) 026 [[arXiv:2101.12660](#)] [[INSPIRE](#)].
- [46] A. Finke et al., *Modified gravitational wave propagation and the binary neutron star mass function*, *Phys. Dark Univ.* **36** (2022) 100994 [[arXiv:2108.04065](#)] [[INSPIRE](#)].
- [47] A. Finke et al., *Probing modified gravitational wave propagation with strongly lensed coalescing binaries*, *Phys. Rev. D* **104** (2021) 084057 [[arXiv:2107.05046](#)] [[INSPIRE](#)].

- [48] D.E. Holz and S.A. Hughes, *Using gravitational-wave standard sirens*, *Astrophys. J.* **629** (2005) 15 [[astro-ph/0504616](#)] [[INSPIRE](#)].
- [49] 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** (2010) 124046 [[arXiv:1004.3988](#)] [[INSPIRE](#)].
- [50] S. Camera and A. Nishizawa, *Beyond concordance cosmology with magnification of gravitational-wave standard sirens*, *Phys. Rev. Lett.* **110** (2013) 151103 [[arXiv:1303.5446](#)] [[INSPIRE](#)].
- [51] S. Mukherjee, B.D. Wandelt and J. Silk, *Probing the theory of gravity with gravitational lensing of gravitational waves and galaxy surveys*, *Mon. Not. Roy. Astron. Soc.* **494** (2020) 1956 [[arXiv:1908.08951](#)] [[INSPIRE](#)].
- [52] S. Libanore et al., *Clustering of gravitational wave and supernovae events: a multitracers analysis in luminosity distance space*, *JCAP* **02** (2022) 003 [[arXiv:2109.10857](#)] [[INSPIRE](#)].
- [53] LISA COSMOLOGY WORKING GROUP collaboration, *Cosmology with the Laser Interferometer Space Antenna*, [arXiv:2204.05434](#) [[INSPIRE](#)].
- [54] S. Mukherjee, B.D. Wandelt and J. Silk, *Multimessenger tests of gravity with weakly lensed gravitational waves*, *Phys. Rev. D* **101** (2020) 103509 [[arXiv:1908.08950](#)] [[INSPIRE](#)].
- [55] G. Scelfo et al., *Gravitational waves $\times$ HI intensity mapping: cosmological and astrophysical applications*, *JCAP* **01** (2022) 004 [[arXiv:2106.09786](#)] [[INSPIRE](#)].
- [56] G. Scelfo, M. Berti, A. Silvestri and M. Viel, *Testing gravity with gravitational waves $\times$ electromagnetic probes cross-correlations*, *JCAP* **02** (2023) 010 [[arXiv:2210.02460](#)] [[INSPIRE](#)].
- [57] G. Congedo and A. Taylor, *Joint cosmological inference of standard sirens and gravitational wave weak lensing*, *Phys. Rev. D* **99** (2019) 083526 [[arXiv:1812.02730](#)] [[INSPIRE](#)].
- [58] C.T. Mpeha, G. Congedo and A. Taylor, *Future prospects on testing extensions to Λ CDM through the weak lensing of gravitational waves*, *Phys. Rev. D* **107** (2023) 103518 [[arXiv:2208.05959](#)] [[INSPIRE](#)].
- [59] M. Oguri, *Measuring the distance-redshift relation with the cross-correlation of gravitational wave standard sirens and galaxies*, *Phys. Rev. D* **93** (2016) 083511 [[arXiv:1603.02356](#)] [[INSPIRE](#)].
- [60] F. Calore et al., *Cross-correlating galaxy catalogs and gravitational waves: a tomographic approach*, *Phys. Rev. Res.* **2** (2020) 023314 [[arXiv:2002.02466](#)] [[INSPIRE](#)].
- [61] S. Mukherjee, B.D. Wandelt, S.M. Nissanke and A. Silvestri, *Accurate precision cosmology with redshift unknown gravitational wave sources*, *Phys. Rev. D* **103** (2021) 043520 [[arXiv:2007.02943](#)] [[INSPIRE](#)].
- [62] G. Scelfo, L. Boco, A. Lapi and M. Viel, *Exploring galaxies-gravitational waves cross-correlations as an astrophysical probe*, *JCAP* **10** (2020) 045 [[arXiv:2007.08534](#)] [[INSPIRE](#)].
- [63] C.C. Diaz and S. Mukherjee, *Mapping the cosmic expansion history from LIGO-Virgo-KAGRA in synergy with DESI and SPHEREx*, *Mon. Not. Roy. Astron. Soc.* **511** (2022) 2782 [[arXiv:2107.12787](#)] [[INSPIRE](#)].
- [64] S. Mukherjee, A. Krolewski, B.D. Wandelt and J. Silk, *Cross-correlating dark sirens and galaxies: measurement of H_0 from GWTC-3 of LIGO-Virgo-KAGRA*, [arXiv:2203.03643](#) [[INSPIRE](#)].
- [65] B.F. Schutz, *Determining the Hubble constant from gravitational wave observations*, *Nature* **323** (1986) 310 [[INSPIRE](#)].

- [66] V. Varma and P. Ajith, *Effects of nonquadrupole modes in the detection and parameter estimation of black hole binaries with nonprecessing spins*, *Phys. Rev. D* **96** (2017) 124024 [[arXiv:1612.05608](#)] [[INSPIRE](#)].
- [67] V. Varma et al., *Gravitational-wave observations of binary black holes: effect of nonquadrupole modes*, *Phys. Rev. D* **90** (2014) 124004 [[arXiv:1409.2349](#)] [[INSPIRE](#)].
- [68] C. Guidorzi et al., *Improved constraints on H_0 from a combined analysis of gravitational-wave and electromagnetic emission from GW170817*, *Astrophys. J. Lett.* **851** (2017) L36 [[arXiv:1710.06426](#)] [[INSPIRE](#)].
- [69] K.P. Mooley et al., *Superluminal motion of a relativistic jet in the neutron-star merger GW170817*, *Nature* **561** (2018) 355 [[arXiv:1806.09693](#)] [[INSPIRE](#)].
- [70] K. Hotokezaka et al., *A Hubble constant measurement from superluminal motion of the jet in GW170817*, *Nature Astron.* **3** (2019) 940 [[arXiv:1806.10596](#)] [[INSPIRE](#)].
- [71] H. Wang and D. Giannios, *Multimessenger parameter estimation of GW170817: from jet structure to the Hubble constant*, *Astrophys. J.* **908** (2021) 200 [[arXiv:2009.04427](#)] [[INSPIRE](#)].
- [72] S.R. Taylor and J.R. Gair, *Cosmology with the lights off: standard sirens in the Einstein Telescope era*, *Phys. Rev. D* **86** (2012) 023502 [[arXiv:1204.6739](#)] [[INSPIRE](#)].
- [73] W.M. Farr, M. Fishbach, J. Ye and D. Holz, *A future percent-level measurement of the Hubble expansion at redshift 0.8 with advanced LIGO*, *Astrophys. J. Lett.* **883** (2019) L42 [[arXiv:1908.09084](#)] [[INSPIRE](#)].
- [74] S. Mastrogiovanni et al., *On the importance of source population models for gravitational-wave cosmology*, *Phys. Rev. D* **104** (2021) 062009 [[arXiv:2103.14663](#)] [[INSPIRE](#)].
- [75] S. Mukherjee, *The redshift dependence of black hole mass distribution: is it reliable for standard sirens cosmology?*, *Mon. Not. Roy. Astron. Soc.* **515** (2022) 5495 [[arXiv:2112.10256](#)] [[INSPIRE](#)].
- [76] J.M. Ezquiaga and D.E. Holz, *Spectral sirens: cosmology from the full mass distribution of compact binaries*, *Phys. Rev. Lett.* **129** (2022) 061102 [[arXiv:2202.08240](#)] [[INSPIRE](#)].
- [77] E. Troja et al., *The X-ray counterpart to the gravitational wave event GW170817*, *Nature* **551** (2017) 71 [[arXiv:1710.05433](#)] [[INSPIRE](#)].
- [78] G. Hallinan et al., *A radio counterpart to a neutron star merger*, *Science* **358** (2017) 1579 [[arXiv:1710.05435](#)] [[INSPIRE](#)].
- [79] G. Ghirlanda et al., *Compact radio emission indicates a structured jet was produced by a binary neutron star merger*, *Science* **363** (2019) 968 [[arXiv:1808.00469](#)] [[INSPIRE](#)].
- [80] E. Pian et al., *Spectroscopic identification of r-process nucleosynthesis in a double neutron star merger*, *Nature* **551** (2017) 67 [[arXiv:1710.05858](#)] [[INSPIRE](#)].
- [81] S.J. Smartt et al., *A kilonova as the electromagnetic counterpart to a gravitational-wave source*, *Nature* **551** (2017) 75 [[arXiv:1710.05841](#)] [[INSPIRE](#)].
- [82] D. Kasen et al., *Origin of the heavy elements in binary neutron-star mergers from a gravitational wave event*, *Nature* **551** (2017) 80 [[arXiv:1710.05463](#)] [[INSPIRE](#)].
- [83] H.T. Janka, T. Eberl, M. Ruffert and C.L. Fryer, *Black hole: neutron star mergers as central engines of gamma-ray bursts*, *Astrophys. J. Lett.* **527** (1999) L39 [[astro-ph/9908290](#)] [[INSPIRE](#)].
- [84] J.M. Miller, A.C. Fabian and M.C. Miller, *A comparison of intermediate mass black hole candidate ULXs and stellar-mass black holes*, *Astrophys. J. Lett.* **614** (2004) L117 [[astro-ph/0406656](#)] [[INSPIRE](#)].

- [85] A. Caputo et al., *Gravitational-wave detection and parameter estimation for accreting black-hole binaries and their electromagnetic counterpart*, *Astrophys. J.* **892** (2020) 90 [[arXiv:2001.03620](#)] [[INSPIRE](#)].
- [86] J.D. Schnittman and J.H. Krolik, *The infrared afterglow of supermassive black hole mergers*, *Astrophys. J.* **684** (2008) 835 [[arXiv:0802.3556](#)] [[INSPIRE](#)].
- [87] E.M. Rossi et al., *Black hole mergers: the first light*, *Mon. Not. Roy. Astron. Soc.* **401** (2010) 2021 [[arXiv:0910.0002](#)] [[INSPIRE](#)].
- [88] C. Palenzuela, L. Lehner and S.L. Liebling, *Dual jets from binary black holes*, *Science* **329** (2010) 927 [[arXiv:1005.1067](#)] [[INSPIRE](#)].
- [89] P. Moesta et al., *On the detectability of dual jets from binary black holes*, *Astrophys. J. Lett.* **749** (2012) L32 [[arXiv:1109.1177](#)] [[INSPIRE](#)].
- [90] R. Gold et al., *Accretion disks around binary black holes of unequal mass: general relativistic MHD simulations of postdecoupling and merger*, *Phys. Rev. D* **90** (2014) 104030 [[arXiv:1410.1543](#)] [[INSPIRE](#)].
- [91] A. Khan, V. Paschalidis, M. Ruiz and S.L. Shapiro, *Disks around merging binary black holes: from GW150914 to supermassive black holes*, *Phys. Rev. D* **97** (2018) 044036 [[arXiv:1801.02624](#)] [[INSPIRE](#)].
- [92] S. Ronchini et al., *Perspectives for multimessenger astronomy with the next generation of gravitational-wave detectors and high-energy satellites*, *Astron. Astrophys.* **665** (2022) A97 [[arXiv:2204.01746](#)] [[INSPIRE](#)].
- [93] M. Pieroni, A. Ricciardone and E. Barausse, *Detectability and parameter estimation of stellar origin black hole binaries with next generation gravitational wave detectors*, *Sci. Rep.* **12** (2022) 17940 [[arXiv:2203.12586](#)] [[INSPIRE](#)].
- [94] S. Kawamura et al., *Current status of space gravitational wave antenna DECIGO and B-DECIGO*, *PTEP* **2021** (2021) 05A105 [[arXiv:2006.13545](#)] [[INSPIRE](#)].
- [95] J. Crowder and N.J. Cornish, *Beyond LISA: exploring future gravitational wave missions*, *Phys. Rev. D* **72** (2005) 083005 [[gr-qc/0506015](#)] [[INSPIRE](#)].
- [96] J. Baker et al., *High angular resolution gravitational wave astronomy*, *Exper. Astron.* **51** (2021) 1441 [[arXiv:1908.11410](#)] [[INSPIRE](#)].
- [97] W.-H. Ruan, Z.-K. Guo, R.-G. Cai and Y.-Z. Zhang, *Taiji program: gravitational-wave sources*, *Int. J. Mod. Phys. A* **35** (2020) 2050075 [[arXiv:1807.09495](#)] [[INSPIRE](#)].
- [98] TIANQIN collaboration, *The TianQin project: current progress on science and technology*, *PTEP* **2021** (2021) 05A107 [[arXiv:2008.10332](#)] [[INSPIRE](#)].
- [99] KAGRA et al. collaborations, *Prospects for observing and localizing gravitational-wave transients with advanced LIGO, advanced Virgo and KAGRA*, *Living Rev. Rel.* **21** (2018) 3 [[arXiv:1304.0670](#)] [[INSPIRE](#)].
- [100] E.D. Hall and M. Evans, *Metrics for next-generation gravitational-wave detectors*, *Class. Quant. Grav.* **36** (2019) 225002 [[arXiv:1902.09485](#)] [[INSPIRE](#)].
- [101] M. Punturo et al., *The Einstein Telescope: a third-generation gravitational wave observatory*, *Class. Quant. Grav.* **27** (2010) 194002 [[INSPIRE](#)].
- [102] A. Klein et al., *The last three years: multiband gravitational-wave observations of stellar-mass binary black holes*, [arXiv:2204.03423](#) [[INSPIRE](#)].
- [103] N. Seto and K. Kyutoku, *How many extragalactic stellar mass binary black holes will be detected by space gravitational-wave interferometers?*, *Mon. Not. Roy. Astron. Soc.* **514** (2022) 4669 [[arXiv:2201.02766](#)] [[INSPIRE](#)].

- [104] C.J. Moore, D. Gerosa and A. Klein, *Are stellar-mass black-hole binaries too quiet for LISA?*, *Mon. Not. Roy. Astron. Soc.* **488** (2019) L94 [[arXiv:1905.11998](#)] [[INSPIRE](#)].
- [105] D. Gerosa et al., *Multiband gravitational-wave event rates and stellar physics*, *Phys. Rev. D* **99** (2019) 103004 [[arXiv:1902.00021](#)] [[INSPIRE](#)].
- [106] J.M. Ezquiaga and M. Zumalacárregui, *Dark energy after GW170817: dead ends and the road ahead*, *Phys. Rev. Lett.* **119** (2017) 251304 [[arXiv:1710.05901](#)] [[INSPIRE](#)].
- [107] P. Creminelli and F. Vernizzi, *Dark energy after GW170817 and GRB170817A*, *Phys. Rev. Lett.* **119** (2017) 251302 [[arXiv:1710.05877](#)] [[INSPIRE](#)].
- [108] T. Baker et al., *Strong constraints on cosmological gravity from GW170817 and GRB170817A*, *Phys. Rev. Lett.* **119** (2017) 251301 [[arXiv:1710.06394](#)] [[INSPIRE](#)].
- [109] T. Baker et al., *Testing gravitational wave propagation with multiband detections*, *JCAP* **03** (2023) 044 [[arXiv:2209.14398](#)] [[INSPIRE](#)].
- [110] N. Frusciante and L. Perenon, *Effective field theory of dark energy: a review*, *Phys. Rept.* **857** (2020) 1 [[arXiv:1907.03150](#)] [[INSPIRE](#)].
- [111] J. Gleyzes, D. Langlois, F. Piazza and F. Vernizzi, *Essential building blocks of dark energy*, *JCAP* **08** (2013) 025 [[arXiv:1304.4840](#)] [[INSPIRE](#)].
- [112] N. Frusciante, G. Papadomanolakis and A. Silvestri, *An extended action for the effective field theory of dark energy: a stability analysis and a complete guide to the mapping at the basis of EFTCAMB*, *JCAP* **07** (2016) 018 [[arXiv:1601.04064](#)] [[INSPIRE](#)].
- [113] N. Frusciante, G. Papadomanolakis, S. Peirone and A. Silvestri, *The role of the tachyonic instability in Horndeski gravity*, *JCAP* **02** (2019) 029 [[arXiv:1810.03461](#)] [[INSPIRE](#)].
- [114] M. Raveri, B. Hu, N. Frusciante and A. Silvestri, *Effective field theory of cosmic acceleration: constraining dark energy with CMB data*, *Phys. Rev. D* **90** (2014) 043513 [[arXiv:1405.1022](#)] [[INSPIRE](#)].
- [115] S. Peirone, M. Martinelli, M. Raveri and A. Silvestri, *Impact of theoretical priors in cosmological analyses: the case of single field quintessence*, *Phys. Rev. D* **96** (2017) 063524 [[arXiv:1702.06526](#)] [[INSPIRE](#)].
- [116] B. Hu, M. Raveri, N. Frusciante and A. Silvestri, *Effective field theory of cosmic acceleration: an implementation in CAMB*, *Phys. Rev. D* **89** (2014) 103530 [[arXiv:1312.5742](#)] [[INSPIRE](#)].
- [117] B. Hu, M. Raveri, N. Frusciante and A. Silvestri, *EFTCAMB/EFTCosmoMC: numerical notes v3.0*, [arXiv:1405.3590](#) [[INSPIRE](#)].
- [118] E. Bellini and I. Sawicki, *Maximal freedom at minimum cost: linear large-scale structure in general modifications of gravity*, *JCAP* **07** (2014) 050 [[arXiv:1404.3713](#)] [[INSPIRE](#)].
- [119] D. Bertacca, A. Raccañelli, N. Bartolo and S. Matarrese, *Cosmological perturbation effects on gravitational-wave luminosity distance estimates*, *Phys. Dark Univ.* **20** (2018) 32 [[arXiv:1702.01750](#)] [[INSPIRE](#)].
- [120] P. Laguna, S.L. Larson, D. Spergel and N. Yunes, *Integrated Sachs-Wolfe effect for gravitational radiation*, *Astrophys. J. Lett.* **715** (2010) L12 [[arXiv:0905.1908](#)] [[INSPIRE](#)].
- [121] A. Garoffolo et al., *Gravitational waves and geometrical optics in scalar-tensor theories*, *JCAP* **11** (2020) 040 [[arXiv:1912.08093](#)] [[INSPIRE](#)].
- [122] A. Garoffolo et al., *Detecting dark energy fluctuations with gravitational waves*, *Phys. Rev. D* **103** (2021) 083506 [[arXiv:2007.13722](#)] [[INSPIRE](#)].
- [123] N. Dalal, D.E. Holz, X.-L. Chen and J.A. Frieman, *Corrective lenses for high redshift supernovae*, *Astrophys. J. Lett.* **585** (2003) L11 [[astro-ph/0206339](#)] [[INSPIRE](#)].

- [124] T. Namikawa, A. Nishizawa and A. Taruya, *Anisotropies of gravitational-wave standard sirens as a new cosmological probe without redshift information*, *Phys. Rev. Lett.* **116** (2016) 121302 [[arXiv:1511.04638](#)] [[INSPIRE](#)].
- [125] EUCLID collaboration, *Euclid preparation: VII. Forecast validation for Euclid cosmological probes*, *Astron. Astrophys.* **642** (2020) A191 [[arXiv:1910.09273](#)] [[INSPIRE](#)].
- [126] N. Kaiser, *Weak gravitational lensing of distant galaxies*, *Astrophys. J.* **388** (1992) 272 [[INSPIRE](#)].
- [127] M. LoVerde and N. Afshordi, *Extended Limber approximation*, *Phys. Rev. D* **78** (2008) 123506 [[arXiv:0809.5112](#)] [[INSPIRE](#)].
- [128] T.D. Kitching et al., *The limits of cosmic shear*, *Mon. Not. Roy. Astron. Soc.* **469** (2017) 2737 [[arXiv:1611.04954](#)] [[INSPIRE](#)].
- [129] M. Kilbinger et al., *Precision calculations of the cosmic shear power spectrum projection*, *Mon. Not. Roy. Astron. Soc.* **472** (2017) 2126 [[arXiv:1702.05301](#)] [[INSPIRE](#)].
- [130] P. Lemos, A. Challinor and G. Efstathiou, *The effect of Limber and flat-sky approximations on galaxy weak lensing*, *JCAP* **05** (2017) 014 [[arXiv:1704.01054](#)] [[INSPIRE](#)].
- [131] W.L. Matthewson and R. Durrer, *The flat sky approximation to galaxy number counts*, *JCAP* **02** (2021) 027 [[arXiv:2006.13525](#)] [[INSPIRE](#)].
- [132] A. Lewis, A. Challinor and A. Lasenby, *Efficient computation of CMB anisotropies in closed FRW models*, *Astrophys. J.* **538** (2000) 473 [[astro-ph/9911177](#)] [[INSPIRE](#)].
- [133] C. Howlett, A. Lewis, A. Hall and A. Challinor, *CMB power spectrum parameter degeneracies in the era of precision cosmology*, *JCAP* **04** (2012) 027 [[arXiv:1201.3654](#)] [[INSPIRE](#)].
- [134] L. Pogosian and A. Silvestri, *What can cosmology tell us about gravity? Constraining Horndeski gravity with Σ and μ* , *Phys. Rev. D* **94** (2016) 104014 [[arXiv:1606.05339](#)] [[INSPIRE](#)].
- [135] N. Frusciante, S. Peirone, S. Casas and N.A. Lima, *Cosmology of surviving Horndeski theory: the road ahead*, *Phys. Rev. D* **99** (2019) 063538 [[arXiv:1810.10521](#)] [[INSPIRE](#)].
- [136] E. Bellini, A.J. Cuesta, R. Jimenez and L. Verde, *Constraints on deviations from Λ CDM within Horndeski gravity*, *JCAP* **02** (2016) 053 [Erratum *ibid.* **06** (2016) E01] [[arXiv:1509.07816](#)] [[INSPIRE](#)].
- [137] M. Chevallier and D. Polarski, *Accelerating universes with scaling dark matter*, *Int. J. Mod. Phys. D* **10** (2001) 213 [[gr-qc/0009008](#)] [[INSPIRE](#)].
- [138] E.V. Linder, *Exploring the expansion history of the universe*, *Phys. Rev. Lett.* **90** (2003) 091301 [[astro-ph/0208512](#)] [[INSPIRE](#)].
- [139] C. Deffayet, O. Pujolas, I. Sawicki and A. Vikman, *Imperfect dark energy from kinetic gravity braiding*, *JCAP* **10** (2010) 026 [[arXiv:1008.0048](#)] [[INSPIRE](#)].
- [140] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [Erratum *ibid.* **652** (2021) C4] [[arXiv:1807.06209](#)] [[INSPIRE](#)].
- [141] M. Mancarella, E. Genoud-Prachex and M. Maggiore, *Cosmology and modified gravitational wave propagation from binary black hole population models*, *Phys. Rev. D* **105** (2022) 064030 [[arXiv:2112.05728](#)] [[INSPIRE](#)].
- [142] J.M. Ezquiaga, *Hearing gravity from the cosmos: GWTC-2 probes general relativity at cosmological scales*, *Phys. Lett. B* **822** (2021) 136665 [[arXiv:2104.05139](#)] [[INSPIRE](#)].
- [143] J.-P. Uzan, *Varying constants, gravitation and cosmology*, *Living Rev. Rel.* **14** (2011) 2 [[arXiv:1009.5514](#)] [[INSPIRE](#)].

- [144] P. Brax, A.-C. Davis, B. Li and H.A. Winther, *A unified description of screened modified gravity*, *Phys. Rev. D* **86** (2012) 044015 [[arXiv:1203.4812](#)] [[INSPIRE](#)].
- [145] A. Joyce, B. Jain, J. Khoury and M. Trodden, *Beyond the cosmological standard model*, *Phys. Rept.* **568** (2015) 1 [[arXiv:1407.0059](#)] [[INSPIRE](#)].
- [146] L. Perenon, F. Piazza, C. Marinoni and L. Hui, *Phenomenology of dark energy: general features of large-scale perturbations*, *JCAP* **11** (2015) 029 [[arXiv:1506.03047](#)] [[INSPIRE](#)].
- [147] Z. Huang, *Observational effects of a running Planck mass*, *Phys. Rev. D* **93** (2016) 043538 [[arXiv:1511.02808](#)] [[INSPIRE](#)].
- [148] A. Spurio Mancini et al., *KiDS + GAMA: constraints on Horndeski gravity from combined large-scale structure probes*, *Mon. Not. Roy. Astron. Soc.* **490** (2019) 2155 [[arXiv:1901.03686](#)] [[INSPIRE](#)].
- [149] J. Noller and A. Nicola, *Cosmological parameter constraints for Horndeski scalar-tensor gravity*, *Phys. Rev. D* **99** (2019) 103502 [[arXiv:1811.12928](#)] [[INSPIRE](#)].
- [150] D. Traykova, E. Bellini and P.G. Ferreira, *The phenomenology of beyond Horndeski gravity*, *JCAP* **08** (2019) 035 [[arXiv:1902.10687](#)] [[INSPIRE](#)].
- [151] L. Amendola, M. Kunz and D. Sapone, *Measuring the dark side (with weak lensing)*, *JCAP* **04** (2008) 013 [[arXiv:0704.2421](#)] [[INSPIRE](#)].
- [152] F. Simpson et al., *CFHTLenS: testing the laws of gravity with tomographic weak lensing and redshift space distortions*, *Mon. Not. Roy. Astron. Soc.* **429** (2013) 2249 [[arXiv:1212.3339](#)] [[INSPIRE](#)].
- [153] A. Ferté, D. Kirk, A.R. Liddle and J. Zuntz, *Testing gravity on cosmological scales with cosmic shear, cosmic microwave background anisotropies, and redshift-space distortions*, *Phys. Rev. D* **99** (2019) 083512 [[arXiv:1712.01846](#)] [[INSPIRE](#)].
- [154] M. Tegmark, A. Taylor and A. Heavens, *Karhunen-Loeve eigenvalue problems in cosmology: how should we tackle large data sets?*, *Astrophys. J.* **480** (1997) 22 [[astro-ph/9603021](#)] [[INSPIRE](#)].
- [155] M. Raveri, M. Martinelli, G. Zhao and Y. Wang, *Information gain in cosmology: from the discovery of expansion to future surveys*, [arXiv:1606.06273](#) [[INSPIRE](#)].
- [156] M. Raveri, M. Martinelli, G. Zhao and Y. Wang, *CosmicFish implementation notes V1.0*, [arXiv:1606.06268](#) [[INSPIRE](#)].
- [157] M. Schmittfull and U. Seljak, *Parameter constraints from cross-correlation of CMB lensing with galaxy clustering*, *Phys. Rev. D* **97** (2018) 123540 [[arXiv:1710.09465](#)] [[INSPIRE](#)].
- [158] M. Martinelli et al., *Ultralarge-scale approximations and galaxy clustering: debiasing constraints on cosmological parameters*, *Mon. Not. Roy. Astron. Soc.* **510** (2022) 1964 [[arXiv:2106.15604](#)] [[INSPIRE](#)].
- [159] N. Tamanini, *Late time cosmology with LISA: probing the cosmic expansion with massive black hole binary mergers as standard sirens*, *J. Phys. Conf. Ser.* **840** (2017) 012029 [[arXiv:1612.02634](#)] [[INSPIRE](#)].
- [160] LISA collaboration, *Astrophysics with the Laser Interferometer Space Antenna*, *Living Rev. Rel.* **26** (2023) 2 [[arXiv:2203.06016](#)] [[INSPIRE](#)].
- [161] S. Mukherjee et al., *Velocity correction for Hubble constant measurements from standard sirens*, *Astron. Astrophys.* **646** (2021) A65 [[arXiv:1909.08627](#)] [[INSPIRE](#)].
- [162] DESI collaboration, *The DESI experiment part I: science, targeting, and survey design*, [arXiv:1611.00036](#) [[INSPIRE](#)].
- [163] M. Salvato, O. Ilbert and B. Hoyle, *The many flavours of photometric redshifts*, *Nature Astron.* **3** (2018) 212 [[arXiv:1805.12574](#)].

- [164] A. Amara and T. Kitching, *Figures of merit for testing standard models: application to dark energy experiments in cosmology*, *Mon. Not. Roy. Astron. Soc.* **413** (2011) 1505 [[arXiv:1009.3274](#)] [[INSPIRE](#)].
- [165] N. Tamanini et al., *Science with the space-based interferometer eLISA. III: probing the expansion of the universe using gravitational wave standard sirens*, *JCAP* **04** (2016) 002 [[arXiv:1601.07112](#)] [[INSPIRE](#)].
- [166] F. Foucart, *Black hole-neutron star mergers: disk mass predictions*, *Phys. Rev. D* **86** (2012) 124007 [[arXiv:1207.6304](#)] [[INSPIRE](#)].
- [167] F. Foucart, T. Hinderer and S. Nissanke, *Remnant baryon mass in neutron star-black hole mergers: predictions for binary neutron star mimickers and rapidly spinning black holes*, *Phys. Rev. D* **98** (2018) 081501 [[arXiv:1807.00011](#)] [[INSPIRE](#)].
- [168] S. Vitale and H.-Y. Chen, *Measuring the Hubble constant with neutron star black hole mergers*, *Phys. Rev. Lett.* **121** (2018) 021303 [[arXiv:1804.07337](#)] [[INSPIRE](#)].
- [169] S.M. Feeney, H.V. Peiris, S.M. Nissanke and D.J. Mortlock, *Prospects for measuring the Hubble constant with neutron-star-black-hole mergers*, *Phys. Rev. Lett.* **126** (2021) 171102 [[arXiv:2012.06593](#)] [[INSPIRE](#)].
- [170] C. Caprini and N. Tamanini, *Constraining early and interacting dark energy with gravitational wave standard sirens: the potential of the eLISA mission*, *JCAP* **10** (2016) 006 [[arXiv:1607.08755](#)] [[INSPIRE](#)].
- [171] E. Belgacem et al., *Cosmology and dark energy from joint gravitational wave-GRB observations*, *JCAP* **08** (2019) 015 [[arXiv:1907.01487](#)] [[INSPIRE](#)].
- [172] D. Laghi et al., *Gravitational-wave cosmology with extreme mass-ratio inspirals*, *Mon. Not. Roy. Astron. Soc.* **508** (2021) 4512 [[arXiv:2102.01708](#)] [[INSPIRE](#)].
- [173] A. Balaudo, A. Garoffolo, M. Martinelli, S. Mukherjee and A. Silvestri, *Impact of standard sirens on gravity constraints in combination with GW weak lensing and galaxy surveys*, in preparation (2022).