



This is the accepted manuscript made available via CHORUS. The article has been published as:

# Joint population and cosmological properties inference with gravitational waves standard sirens and galaxy surveys

Simone Mastrogiovanni, Danny Laghi, Rachel Gray, Giada Caneva Santoro, Archisman Ghosh, Christos Karathanasis, Konstantin Leyde, Danièle A. Steer, Stéphane Perriès, and Grégoire Pierra

Phys. Rev. D **108**, 042002 — Published 9 August 2023

DOI: [10.1103/PhysRevD.108.042002](https://doi.org/10.1103/PhysRevD.108.042002)

# Joint Population and cosmological properties inference with gravitational waves standard sirens and galaxy surveys

Simone Mastrogiovanni <sup>1,\*</sup> Danny Laghi <sup>2</sup> Rachel Gray <sup>3,4</sup> Giada Caneva Santoro <sup>5</sup> Archisman Ghosh <sup>6</sup>  
Christos Karathanasis <sup>5</sup> Konstantin Leyde <sup>7</sup> Danièle A. Steer <sup>7</sup> Stéphane Perriès <sup>8</sup> and Grégoire Pierra <sup>8</sup>

<sup>1</sup>*INFN, Sezione di Roma, I-00185 Roma, Italy*

<sup>2</sup>*Laboratoire des 2 Infinis - Toulouse (L2IT-IN2P3), Université de Toulouse, CNRS, UPS, F-31062 Toulouse Cedex 9, France*

<sup>3</sup>*SUPA, University of Glasgow, Glasgow, G12 8QQ, United Kingdom*

<sup>4</sup>*Department of Physics & Astronomy, Queen Mary University of London, Mile End Road, London, E1 4NS, United Kingdom*

<sup>5</sup>*Institut de Física d'Altes Energies (IFAE), Barcelona Institute of Science and Technology, Barcelona, Spain*

<sup>6</sup>*Department of Physics and Astronomy, Ghent University, Proeftuinstraat 86, 9000 Ghent, Belgium*

<sup>7</sup>*Université Paris Cité, CNRS, Astroparticule et Cosmologie, F-75013 Paris, France*

<sup>8</sup>*Université Lyon, Université Claude Bernard Lyon 1, CNRS,*

*IP2I Lyon/IN2P3, UMR 5822, F-69622 Villeurbanne, France*

(Dated: July 12, 2023)

Gravitational wave (GW) sources at cosmological distances can be used to probe the expansion rate of the Universe. GWs directly provide a distance estimation of the source but no direct information on its redshift. The optimal scenario to obtain a redshift is through the direct identification of an electromagnetic (EM) counterpart and its host galaxy. With almost 100 GW sources detected without EM counterparts (dark sirens), it is becoming crucial to have statistical techniques able to perform cosmological studies in the absence of EM emission. Currently, only two techniques for dark sirens are used on GW observations: the spectral siren method, which is based on the source-frame mass distribution to estimate conjointly cosmology and the source's merger rate, and the galaxy survey method, which uses galaxy surveys to assign a probabilistic redshift to the source while fitting cosmology. It has been recognized, however, that these two methods are two sides of the same coin. In this paper, we present a novel approach to unify these two methods. We apply this approach to several observed GW events using the GLADE+ galaxy catalog discussing limiting cases. We provide estimates of the Hubble constant, modified gravity propagation effects, and population properties for binary black holes. We also estimate the binary black hole merger rate per galaxy to be  $10^{-6} - 10^{-5} \text{yr}^{-1}$  depending on the galaxy catalog hypotheses.

## I. INTRODUCTION

Compact binary coalescences (CBCs) detected via gravitational waves (GWs) are rapidly becoming a central tool to study cosmology. GWs directly provide the luminosity distance of the source, but they are unable to provide the cosmological redshift estimation that is required to measure the cosmic expansion. An estimation of the source redshift could be obtained from the observation of an electromagnetic counterpart (EM) as occurred in the case of the binary neutron star merger GW170817 [1, 2] observed by the LIGO [3] and Virgo [4] GW detectors. However, most of the sources currently detected are binary black holes (BBHs) [5] which are not accompanied by an EM counterpart. In order to exploit this population of “dark sirens” for cosmological studies several methods have been proposed. In this study, we focus on two methods which, however, as we will see, are not independent: the galaxy catalog method and the spectral siren method. Additional methods to assign an implicit redshift from the GW alone include the “cross-correlation method”, which explores the spatial clustering between GW sources and galaxies [6–10], and the “Equation of State method”, which uses the measurement of tidal deformability of neutron stars and knowledge of their Equation of State [11, 12]. For a review on these methods, see [13, 14].

The *galaxy catalog* method, also referred to as the “statistical method”, consists in using galaxy catalogs to statistically assign a host galaxy, hence a redshift, to the GW event [15–17]. This method has been applied in recent studies [18–23] using various galaxy catalogs. One of the main bottlenecks of this method is the fact that at high redshift galaxy catalogs are not complete and therefore could not bring any relevant information on the source redshift. The galaxy catalog method also makes assumptions about the source frame mass distribution that we comment on more in detail later. Another method used in the literature is the *spectral siren* method [24–33] which aims at obtaining a redshift estimate for the GW sources from the relation between source-frame masses and detector-frame masses. In fact, it is possible to obtain a constraint on cosmology by jointly fitting for the population distribution of source-frame masses. The spectral siren method comes with the price of assuming a functional form and some priors for the source-frame mass distribution. In particular, wrong modeling of the population distribution of dark sirens (in particular source-frame masses) can introduce a bias in the estimation of the cosmological parameters [28]. In analogy to the mass method, information on the possible redshift of GW events can also come from the CBC merger rate as a function of redshift. If the CBC merger rate presents local overdensities, such as a global merger rate redshift peak, then one can perform cosmological studies even in the absence of electromagnetic counterparts [34, 35].

As discussed in [28], the galaxy catalog method is in fact a particular extension of the spectral siren method in which

---

\* simone.mastrogiovanni@roma1.infn.it

additional information from galaxy surveys is encoded. The galaxy catalog method requires the definition of a source-frame mass distribution, or more generally a population model for the CBCs. Assumptions about the population of sources typically enter the galaxy catalog method in two ways: (i) they provide a systematic preference among host candidates at different redshift and (ii) they provide a redshift preference for the detected source through the *completeness correction*. The completeness correction accounts for the missing galaxies in the catalog. In fact, in the case that the galaxy catalog is 0% complete (it contains no observed galaxy), the catalog method reduces to the spectral siren method. The tight connection between the spectral siren and the galaxy catalog methods has strong implications in interpreting results from current GW detections. For instance, in [36] it is shown that the constraint on the Hubble constant ( $H_0$ ) from 42 BBHs reported in the third GW transient catalog (GWTC-3) is dominated by the population assumptions for the distribution of BBHs even when a galaxy catalog is employed.

In this paper, we present a new methodology, which we refer to as the “*galaxy density method*”, that is able to unify the galaxy catalog *and* the spectral siren methods. As a result, it is possible to marginalize over all the uncertainties describing the population model while taking into account information from a galaxy catalog. The core of this new method is to link the CBC merger rate (spectral siren) to the galaxy number density (galaxy catalog).

The paper is organized as follows. In Sec. II we review the basics of hierarchical Bayesian inference in presence of selection biases in the context of GW detections. In Sec. III we introduce a parametrization for the CBC merger rate in terms of galaxy number density, and we discuss limiting cases, including how this new approach collapses to the standard spectral siren method. In Sec. IV, we show some tests and discuss the behavior of the method with the two best localized standard sirens to date, GW170817 and GW190814. In Sec. V we perform a full population and galaxy catalog-based analysis of the 42 BBH detections from GWTC-3 analyzed in [36] and we discuss the obtained results. Finally, in Sec. VI we draw our conclusions.

## II. THE HIERARCHICAL LIKELIHOOD

The GW detection of  $N_{\text{GW}}$  events in a given observing time  $T_{\text{obs}}$ , collectively represented by some data  $\{x\}$ , can be described in terms of an inhomogeneous Poissonian process [37, 38] characterised by some parameters  $\Lambda$  by using the following hierarchical likelihood:

$$\begin{aligned} \mathcal{L}(\{x\}|\Lambda) &\propto e^{-N_{\text{exp}}(\Lambda)} \prod_i^{N_{\text{GW}}} T_{\text{obs}} \int d\theta dz \mathcal{L}_{\text{GW}}(x_i|\theta, z, \Lambda_c) \times \\ &\times \frac{1}{1+z} \frac{dN_{\text{CBC}}}{d\theta dz dt_s}(\Lambda). \end{aligned} \quad (1)$$

Here  $N_{\text{exp}}$  is the number of expected CBC detections in  $T_{\text{obs}}$ , while  $z$  and  $\theta$  indicate the redshift and a set of parameters characteristic of each GW event (such as spins and luminosity dis-

tance), respectively. The central quantity of the hierarchical inference is the CBC merger rate in source-frame time  $t_s$ :

$$\frac{dN_{\text{CBC}}}{d\theta dz dt_s}(\Lambda), \quad (2)$$

where  $\Lambda = \{\Lambda_p, \Lambda_c\}$  contains population-level parameters  $\Lambda_p$  as well as cosmological parameters  $\Lambda_c$  such as  $H_0$  and  $\Omega_m$ . The information about how precisely we are able to measure the GW parameters  $\theta$  (such as spins and source masses) from the data is given by the GW likelihood  $\mathcal{L}_{\text{GW}}(x_i|\theta, z, \Lambda_c)$ . The expected number of GW events accounts for the GW selection biases:

$$N_{\text{exp}}(\Lambda) = T_{\text{obs}} \int d\theta dz P_{\text{det}}(z, \theta, \Lambda_c) \frac{dN_{\text{CBC}}}{dz d\theta dt_s}(\Lambda) \frac{1}{1+z}, \quad (3)$$

where the detection probability  $P_{\text{det}}(z, \theta, \Lambda)$  is defined as:

$$P_{\text{det}}(z, \theta, \Lambda_c) = \int_{x \in \text{detectable}} dx \mathcal{L}_{\text{GW}}(x_i|\theta, z, \Lambda_c). \quad (4)$$

Eq. 1 is formally equivalent to:

$$\begin{aligned} \mathcal{L}(\{x\}|\Lambda) &\propto e^{-N_{\text{exp}}(\Lambda)} [N_{\text{exp}}(\Lambda)]^{N_{\text{GW}}} \times \\ &\times \prod_i^{N_{\text{GW}}} \frac{\int d\theta dz \mathcal{L}_{\text{GW}}(x_i|\theta, z, \Lambda_c) \frac{1}{1+z} \frac{dN_{\text{CBC}}}{d\theta dz dt_s}(\Lambda)}{\int d\theta dz P_{\text{det}}(\theta, z, \Lambda_c) \frac{1}{1+z} \frac{dN_{\text{CBC}}}{d\theta dz dt_s}(\Lambda)}, \end{aligned} \quad (5)$$

that can be rewritten as:

$$\begin{aligned} \mathcal{L}(\{x\}|\Lambda) &\propto e^{-N_{\text{exp}}(\Lambda)} [N_{\text{exp}}(\Lambda)]^{N_{\text{GW}}} \times \\ &\times \prod_i^{N_{\text{GW}}} \frac{\int d\theta dz \mathcal{L}_{\text{GW}}(x_i|\theta, z, \Lambda_c) \pi_{\text{pop}}(\theta, z|\Lambda)}{\int d\theta dz P_{\text{det}}(\theta, z, \Lambda_c) \pi_{\text{pop}}(\theta, z|\Lambda)}, \end{aligned} \quad (6)$$

by defining the population parameter probability as:

$$\pi_{\text{pop}}(\theta, z|\Lambda) = \frac{1}{C(\Lambda)} \frac{1}{1+z} \frac{dN_{\text{CBC}}}{d\theta dz dt_s}(\Lambda), \quad (7)$$

where  $C(\Lambda)$  is a normalization constant. Eq. 6 is the form of the hierarchical likelihood typically reported in [28, 39, 40].

However, from a numerical point of view, Eq. 1 has some advantages in comparison to Eq. 6 [41]: (i) it is more connected to a physical quantity, the CBC merger rate and (ii) it involves the numerical computation of two integrals instead of three. In order to obtain Eq. 6 from Eq. 1, recall that the observing time is given by:

$$T_{\text{obs}} = \frac{N_{\text{exp}}(\Lambda)}{\int d\theta dz P_{\text{det}}(z, \theta, \Lambda_c) \frac{dN_{\text{CBC}}}{dz d\theta dt_s}(\Lambda) \frac{1}{1+z}}, \quad (8)$$

which, after substitution into Eq. 1, gives Eq. 6.

The two forms of the hierarchical likelihood in Eq. 1 and Eq. 6 can be written in a “scale-free” version that does not depend on  $N_{\text{exp}}$  (but still accounts for GW selection biases). By choosing a prior<sup>1</sup>  $\pi(N_{\text{exp}}) \propto 1/N_{\text{exp}}$ , it is possible to marginalize analytically over  $N_{\text{exp}}$  and obtain [42]

$$\mathcal{L}(\{x\}|\Lambda) \propto \prod_i^{N_{\text{GW}}} \frac{\int d\theta dz \mathcal{L}(x_i|\theta, z, \Lambda_c) \frac{dN_{\text{CBC}}(\Lambda)}{dz d\theta dt_s} \frac{1}{1+z}}{\int d\theta dz P_{\text{det}}(\theta, z, \Lambda_c) \frac{dN_{\text{CBC}}(\Lambda)}{dz d\theta dt_s} \frac{1}{1+z}}. \quad (9)$$

<sup>1</sup> Note that a prior  $\pi(N_{\text{exp}}) \propto 1/N_{\text{exp}}$  would correspond to a prior  $\pi(R_0) \propto 1/R_0$ , where  $R_0$  is the CBC merger rate per  $\text{Gpc}^{-3}\text{yr}^{-1}$ .

The Poissonian process described by the hierarchical likelihoods in Eqs. 1-6-9 can be used to jointly measure the CBC merger rate and cosmological parameters. The fundamental aspect that opens up this possibility is the fact that GW detectors directly measure the luminosity distance  $d_L$  to the source and the detector frame masses  $\vec{m}_d$ . We recall that, for a flat  $\Lambda$ CDM model, the luminosity distance is:

$$d_L(z) = \frac{c(1+z)}{H_0} \int_0^z \frac{dz'}{E(z')}, \quad (10)$$

where  $E(z) = \sqrt{\Omega_m(1+z)^3 + (1-\Omega_m)}$ . Furthermore, detector masses are linked to source masses through  $\vec{m}_d = \vec{m}_s(1+z)$ . In [24–29, 31–33] it was shown that by modeling the CBC merger rate in terms of source-frame masses, it is possible to obtain a joint measurement of both the CBC merger rate and the cosmological parameters. Similarly, in [34, 35] it was demonstrated that if the CBC merger rate shows some features in redshift, such as peaks, this might help to measure cosmological parameters even without any information about the source-frame masses. In this work, we will apply Eq. 6 and Eq. 9 either to measure the Hubble constant only, or to jointly measure  $H_0$  and the population parameters related to the mass and redshift models used (see App. A). In the rest of this paper, we assume a flat  $\Lambda$ CDM cosmology model with free  $H_0$  and fixed  $\Omega_m = 0.308$ , with  $\Omega_\Lambda = 1 - \Omega_m$ .

### III. CBC MERGER RATE PARAMETRIZATIONS

In Sec. II we have seen that the relevant GW parameters which are crucial for GW cosmology are the source-frame masses  $\vec{m}_s$  and the redshift  $z$ . However, when working with galaxy catalogs also the sky position, which we label by  $\Omega$ , is a central quantity since it can help us to identify possible host galaxies along a given line-of-sight. Therefore, in the remainder of this paper, we will work with a CBC merger rate written as:

$$\frac{dN_{\text{CBC}}}{d\vec{m}_s d\theta d\Omega dz dt_s}(\Lambda). \quad (11)$$

Note that Eq. 11 is a CBC source-frame merger rate per source-frame mass, per GW parameters  $\theta$ , per sky position  $\Omega$ , per redshift  $z$ . To simplify the notation, we will refer to this quantity as the ‘‘CBC merger rate’’, unless we want to discuss an explicit parametrization of it. For brevity, we will also drop the  $\Lambda$  dependence.

#### A. The CBC spectral siren parametrization

In standard cosmological-population analyses such as [28, 41], which do not make explicit use of galaxy surveys, the CBC merger rate is always factorized in terms of comoving volume rather than the redshift, since it is easier to model and also easier to motivate physically. The CBC merger rate thus

becomes:

$$\begin{aligned} \frac{dN_{\text{CBC}}}{d\vec{m}_s d\theta d\Omega dz dt_s} &= \frac{dN_{\text{CBC}}}{d\vec{m}_s d\theta dV_c dt_s} \frac{dV_c}{dz d\Omega}(\Lambda_c) \\ &= R_0 \psi(z; \Lambda) p_{\text{pop}}(\vec{m}_s, \theta | z, \Lambda) \frac{dV_c}{dz d\Omega}, \end{aligned} \quad (12)$$

where  $R_0$  is the CBC merger rate per comoving volume at redshift  $z = 0$ ,  $\psi(z; \Lambda)$  is a phenomenological parametrization for the CBC merger rate as a function of redshift such that  $\psi(z = 0; \Lambda) = 1$ ,  $p_{\text{pop}}$  is a probability density function describing the distribution of GW sources in masses and other parameters  $\theta$ , and

$$\frac{dV_c}{dz d\Omega} = \left[ \frac{c}{H_0} \right]^3 \left[ \int_0^z \frac{dz'}{E(z')} \right]^2 \quad (13)$$

is the differential of the comoving volume. Typically, it is assumed that  $p_{\text{pop}}$  in Eq. 12 is independent of redshift, *i.e.* the distribution of masses and GW population parameters does not evolve with redshift. This is indeed a simplifying assumption as source mass distributions might depend on redshift [43, 44] as well as spin parameters [45, 46]. This assumption, if wrong, could introduce a bias in the estimation of the cosmological parameters. However, it has been shown in [32] that for a mild evolution one should still be able to obtain unbiased estimation of  $H_0$ .

#### B. The CBC merger rate informed by galaxy surveys

We now proceed to derive the CBC merger rate encoding information from the galaxy catalog. To construct this CBC merger rate, we will make two assumptions: (i) mergers can only occur inside galaxies and (ii) mergers can be more or less frequent according to the absolute magnitude  $M$  (or luminosity) and redshift of the galaxy. In this paper, we will mostly use galaxies observed in the K-band, but also other colors could be used, in particular if they trace the binary merger rate in redshift. We can write the overall CBC merger rate as the integral over the absolute magnitude of the CBC rate per galaxy absolute magnitude, namely:

$$\frac{dN_{\text{CBC}}}{d\vec{m}_s d\theta d\Omega dz dt_s} = \int dM \frac{dN_{\text{CBC}}}{d\vec{m}_s d\theta d\Omega dz dt_s dM}. \quad (14)$$

Similar to the spectral siren case (see Eq. 12), we parametrize the CBC rate per galaxy (rather than comoving volume), namely:

$$\frac{dN_{\text{CBC}}}{d\vec{m}_s d\theta d\Omega dz dt_s dM} = \frac{dN_{\text{CBC}}}{dN_{\text{gal}} d\theta d\vec{m}_s dt_s} \frac{dN_{\text{gal}}}{dz d\Omega dM}. \quad (15)$$

In order to write Eq. 15, we have further assumed that the number of galaxies only depends on sky position, absolute magnitude, and redshift. The first term in Eq. 15 is the CBC rate per galaxy which, similarly to Eq. 12, we parametrize as:

$$\frac{dN_{\text{CBC}}}{dN_{\text{gal}} d\vec{m}_s d\theta dt_s} = R_{\text{gal},0}^* \Psi_L(z, M; \Lambda_p) p_{\text{pop}}(\vec{m}_s, \theta | z, M, \Lambda_p), \quad (16)$$

where  $R_{\text{gal},0}^*$  is the CBC rate per year at a reference galaxy luminosity (or absolute magnitude)  $L_* = L(M_*)$ , while  $\Psi_L(z, M; \Lambda_p)$  is an evolution function for the CBC merger rate per galaxy such that  $\Psi_L(z = 0, M = M_*; \Lambda_p) = 1$ . In Eq. 16  $p_{\text{pop}}$  has been expressed in a general form assuming that the probability of having some GW parameters might be dependent on the galaxy redshift and absolute magnitude. For the analysis presented in this paper, we will not consider this possibility and we will choose  $p_{\text{pop}}(\vec{m}_s, \theta|z, M, \Lambda_p) = p_{\text{pop}}(\vec{m}_s, \theta|\Lambda_p)$ . Regarding the rate function  $\Psi_L(z, M; \Lambda_p)$  we use the following parametrization:

$$\Psi_L(z, L(M); \Lambda_p) = \psi(z; \Lambda_p) \left\| \frac{L}{L_*} \right\|^\epsilon = \psi(z; \Lambda_p) 10^{0.4\epsilon(M_* - M)}, \quad (17)$$

*i.e.* we factorize the rate evolution function into a redshift-dependent part and a factor that depends on the luminosity of the galaxy. Similarly to Eq. 12, the redshift part  $\psi(z; \Lambda_p)$  is modeled to track the merger rate as a function of redshift, while the luminosity part models the fact that the merger rate might depend on the luminosity of the galaxy (which could

be a tracer of the Star Formation Rate [47]). As an example, the parameter  $\epsilon > 0$  indicates that more luminous galaxies are more likely to host CBCs, while  $\epsilon = 0$  indicates that there is no such preference. In principle, one could also treat  $\epsilon$  as a population parameter and jointly fit also for it. However, considering  $\epsilon$  as a free parameter poses some numerical challenges because when trying different values of  $\epsilon$ , the galaxy catalog must be handled multiple times.

A good choice for the reference luminosity  $L_*$  is the ‘‘knee’’ luminosity of the Schechter function. The Schechter function describes the number density of galaxies in the comoving volume per luminosity (or absolute magnitude). In terms of absolute magnitude, it is identified by five parameters: a galaxy number density  $\phi_*$  per comoving volume, a knee absolute magnitude  $M_*$ , a faint end  $M_{\text{max}}$ , a bright end  $M_{\text{min}}$ , and a slope parameter  $\alpha$ . Note that in practice the bright end choice is not important as the Schechter function quickly drops to 0 after  $M_*(H_0)$ . However, here we write it explicitly for a pedagogic derivation. We can thus write the Schechter function as:

$$\text{Sch}(M; \phi_*, M_*, M_{\text{max}}, M_{\text{min}}, \alpha) = \begin{cases} 0.4 \ln(10) \phi_* 10^{0.4(\alpha+1)(M_* - M)} \exp[-10^{0.4(M_* - M)}], & M_{\text{min}} \leq M \leq M_{\text{max}}, \\ 0, & \text{otherwise.} \end{cases} \quad (18)$$

Usually, Schechter parameters are provided for  $H_0 = 100h \text{ km s}^{-1} \text{ Mpc}^{-1}$  with  $h = 1$ , so one typically needs to rescale them. The scaling relations are given in terms of absolute magnitude by:

$$\begin{aligned} M_*(h) &= M_*(h = 1) + 5 \log_{10} h, & (19) \\ \phi_*(h) &= \phi_*(h = 1) h^3. & (20) \end{aligned}$$

Note that in writing down the Schechter function, we have implicitly assumed that it does not evolve in redshift. This is a simplification as we can measure for some colors an evolution of the Schechter function in redshift (see *e.g.* [48]). As we will see later, the modeling of the Schechter function is crucial in the case that the galaxy catalog is not complete, as it enters the completeness correction. A wrong modelization of the Schechter function evolution might result in a wrong modelization of the CBC merger rate as a function of redshift (as mergers are expected to track galaxy density). This will translate to bias on  $H_0$ . In order to understand the magnitude of the bias on  $H_0$ , and how this interacts with other population parameters, we would require a dedicated mock data challenge which is beyond the scope of this paper.

We now discuss how to build the number density of galaxies in Eq. 15. In the limit in which we are provided with a galaxy survey containing all the galaxies in the Universe, this term can be constructed as a sum of delta functions at the redshift of each galaxy, sky position, and absolute magnitude. However, galaxy catalogs are not complete and therefore the galaxy density must be corrected to include the missing galaxies. In other words, the overall galaxy density can be calcu-

lated as the sum of a galaxy density detected by the catalog and a completeness correction, *i.e.*

$$\frac{dN_{\text{gal}}}{dz d\Omega dM} = \frac{dN_{\text{gal,cat}}}{dz d\Omega dM} + \frac{dN_{\text{gal,out}}}{dz d\Omega dM}. \quad (21)$$

The overall CBC merger rate will be then given by plugging the galaxy density distribution, Eq. 21, and the model of the CBC rate per galaxy, Eq. 17, into Eq. 16, thus obtaining:

$$\begin{aligned} \frac{dN_{\text{CBC}}}{d\vec{m}_s d\theta d\Omega dz dt_s dM} &= R_{\text{gal},0}^* p_{\text{pop}}(\vec{m}_s, \theta|\Lambda_p) \psi(z; \Lambda_p) \times \\ &\times 10^{0.4\epsilon(M_* - M)} \left[ \frac{dN_{\text{gal,cat}}}{dz d\Omega dM} + \frac{dN_{\text{gal,out}}}{dz d\Omega dM} \right]. \end{aligned} \quad (22)$$

Note that in order to compute the hierarchical likelihood in Eq. 1 we need to marginalize out the absolute magnitude from Eq. 22, as GW detectors are not able to measure this quantity. In fact, we will need to calculate:

$$\begin{aligned} \frac{dN_{\text{CBC}}}{d\vec{m}_s d\theta d\Omega dz dt_s} &= R_{\text{gal},0}^* p_{\text{pop}}(\vec{m}_s, \theta|\Lambda_p) \psi(z; \Lambda_p) \times \\ &\times \int_{M_{\text{min}}}^{M_{\text{max}}} dM 10^{0.4\epsilon(M_* - M)} \left[ \frac{dN_{\text{gal,cat}}}{dz d\Omega dM} + \frac{dN_{\text{gal,out}}}{dz d\Omega dM} \right]. \end{aligned} \quad (23)$$

### 1. The completeness correction

The second term in Eq. 21, which is the completeness correction, or ‘‘out-of-catalog’’ term [21, 22], can be written as

the background galaxy density times the probability  $p_{\text{miss}}$  to miss a galaxy at redshift  $z$ , absolute magnitude  $M$ , and sky position  $\Omega$ , namely:

$$\frac{dN_{\text{gal,out}}}{dzd\Omega dM} = \frac{dN_{\text{gal,bg}}}{dzd\Omega dM} p_{\text{miss}}(z, \Omega, M). \quad (24)$$

The background galaxy density can be defined in terms of Schechter function with slope  $\alpha$  as:

$$\frac{dN_{\text{gal,bg}}}{dzd\Omega dM} = \text{Sch}(M; \alpha) \frac{dV_c}{dzd\Omega}, \quad (25)$$

where for brevity we just write  $\text{Sch}(M; \alpha)$  to indicate the Schechter function. Moving to  $p_{\text{miss}}$ , we recall that usually galaxy catalogs are flux-limited, *i.e.* a galaxy is detected if its apparent magnitude  $m$  is brighter than a given threshold  $m_{\text{thr}}$ . Note that  $m_{\text{thr}}$  could be sky-angle dependent to indicate that the galaxy survey is not able to cover in the same way all the sky, *e.g.* a blind region would be identified by  $m_{\text{thr}} = -\infty$ . Therefore, we can model  $p_{\text{miss}}$  as an Heaviside step function,  $p_{\text{miss}} = \Theta[(M > M_{\text{thr}}(z, m_{\text{thr}}(\Omega), H_0))]$ .

We can now use the out-of-catalog part to calculate its contribution to the CBC merger rate in Eq. 23. In doing so we first note that the integral of the Schechter function multiplied by the luminosity weight can be written as:

$$\begin{aligned} & \int_{M_{\text{min}}}^{M_{\text{max}}} dM \text{Sch}(M; \alpha) 10^{0.4\epsilon(M_*(H_0)-M)} = \\ & = \int_{M_{\text{min}}}^{M_{\text{max}}} dM \text{Sch}(M; \alpha + \epsilon) \\ & = \phi_*(H_0) \Gamma_{\text{inc}}(\alpha + \epsilon + 1, x_{\text{min}}, x_{\text{max}}), \end{aligned} \quad (26)$$

where  $\Gamma_{\text{inc}}$  is the incomplete gamma function, while in the last step of the above equation, we have changed variable to  $x = 10^{0.4(M_*(H_0)-M)}$  and  $x_{\text{min}/\text{max}} = 10^{0.4(M_*(H_0)-M_{\text{min}/\text{max}})}$ . Bearing in mind the result obtained in Eq. 26 and using Eq. 24 we can define an “effective” galaxy density per steradian per redshift which are CBC emitters:

$$\begin{aligned} \frac{dN_{\text{gal,out}}^{\text{eff}}}{dzd\Omega} &= \int_{M_{\text{min}}(H_0)}^{M_{\text{max}}(H_0)} dM 10^{0.4\epsilon(M_*(H_0)-M)} \frac{dN_{\text{gal,out}}}{dzd\Omega dM} \\ &= \int_{M_{\text{thr}}(H_0, m_{\text{thr}}(\Omega), z)}^{M_{\text{max}}(H_0)} dM \text{Sch}(M; \alpha + \epsilon) \frac{dV_c}{dzd\Omega} \\ &= \phi_*(H_0) \Gamma_{\text{inc}}(\alpha + \epsilon + 1, x_{\text{max}}, x_{\text{thr}}) \frac{dV_c}{dzd\Omega}, \end{aligned} \quad (27)$$

where  $x_{\text{thr}} = 10^{0.4(M_*(H_0)-M_{\text{thr}}(H_0))}$ . The bright end of the integration is given by the absolute magnitude threshold, which is computed from the sky-dependent apparent magnitude threshold as a function of redshift and  $H_0$ . Before explaining the physical sense of the above result, let us note that Eq. 27 is not actually dependent on  $H_0$ . On one hand,  $\phi_*(H_0) \propto H_0^3$ , while  $\frac{dV_c}{dzd\Omega} \propto H_0^{-3}$ , therefore the  $H_0$ -dependence will cancel. On the other hand, the incomplete gamma function depends on the slope parameters  $\alpha$  and  $\epsilon$  (which are cosmology-independent) and on the parameter  $x$ . The parameter  $x$  is

function of  $M_*(H_0) - M_{\text{max}/\text{thr}}(H_0)$ , and this difference is  $H_0$ -independent as the two absolute magnitudes will scale in the same way with  $H_0$ .

Eq. 27 can be thought of as an “effective” density of galaxy per steradian per redshift, which are the CBC emitters that are missing from the catalog. As an example, if the catalog is 0% complete (no galaxies are present and  $M_{\text{thr}} = -\infty$ ) and no luminosity weight is applied ( $\epsilon = 0$ ), then Eq. 27 will simply reduce to the integral of the Schechter function times the differential of the comoving volume, which is by definition the number density of galaxies per redshift per steradian. If the galaxy catalog is 100% complete (all the galaxies are present and  $M_{\text{thr}} = \infty$ ), then the integral in Eq. 27 will evaluate to 0 as by definition the Schechter function is null for  $M$  fainter than  $M_{\text{max}}$ .

Using Eq. 27 we can also give a more formal definition of “incompleteness”. This can be defined as the fraction of galaxies that are CBC emitters that we think are not present in the catalog part. More formally, we can define the incompleteness  $\mathcal{I}_{\text{CBC}}(z, \Omega)$  as:

$$\begin{aligned} \mathcal{I}_{\text{CBC}}(z, \Omega) &= \frac{\int_{M_{\text{thr}}(H_0, m_{\text{thr}}(\Omega), z)}^{M_{\text{max}}(H_0)} dM \text{Sch}(M; \alpha + \epsilon) \frac{dV_c}{dzd\Omega}}{\int_{M_{\text{min}}}^{M_{\text{max}}(H_0)} dM \text{Sch}(M; \alpha + \epsilon) \frac{dV_c}{dzd\Omega}} \\ &= \frac{\Gamma_{\text{inc}}(\alpha + \epsilon + 1, x_{\text{max}}, x_{\text{thr}}(\Omega, z))}{\Gamma_{\text{inc}}(\alpha + \epsilon + 1, x_{\text{max}}, x_{\text{min}})}. \end{aligned} \quad (28)$$

In analogy with the discussion in the previous paragraph, we note that the incompleteness is also a quantity that does not depend on  $H_0$ .

## 2. The catalog part

We now calculate the “in-catalog” part of Eq. 23, which is built starting from the galaxies present in the catalog. The discussion in this section follows the same logic as presented in [49]. In analogy to what we discussed in the previous section, even for the catalog part, we can define an effective number of CBC emitters:

$$\begin{aligned} \frac{dN_{\text{gal,cat}}^{\text{eff}}}{dzd\Omega} &= \\ &= \int_{M_{\text{min}}}^{M_{\text{thr}}(z, m_{\text{thr}}(\Omega), H_0)} dM 10^{0.4\epsilon(M_*(H_0)-M)} \frac{dN_{\text{gal,cat}}}{dzd\Omega dM}. \end{aligned} \quad (29)$$

The integral in Eq. 29 is limited up to a faint end identified by  $M_{\text{thr}}(z, m_{\text{thr}}(\Omega), H_0)$ , since by definition we can not detect galaxies that are fainter than a given threshold value (which are accounted for by the completeness correction). Since galaxy surveys usually provide lists of galaxy redshifts, sky locations, and apparent magnitudes, it is more convenient to express the above integral in terms of apparent magnitudes  $m$  as:

$$\begin{aligned} \frac{dN_{\text{gal,cat}}^{\text{eff}}}{dzd\Omega} &= \\ &= \int_{m_{\text{min}}}^{m_{\text{thr}}(\Omega)} dm 10^{0.4\epsilon(M_*(H_0)-M(H_0, m, z))} \frac{dN_{\text{gal,cat}}}{dzd\Omega dm}. \end{aligned} \quad (30)$$

The galaxy number density can be built from the observed list of galaxies as the product of the total number of galaxies observed in the catalog and the probability of finding a galaxy at redshift  $z$ , sky position  $\Omega$ , and apparent magnitude  $m$ , given our observation of the catalog (labeling by  $C$  the observed quantities):

$$\frac{dN_{\text{gal,cat}}}{dzd\Omega dm} = N_{\text{gal}} p_{\text{gal}}(z, \Omega, m|C). \quad (31)$$

If the galaxy catalog provides us with perfectly measured redshift, sky position, and apparent magnitudes,  $p_{\text{gal}}(z, \Omega, m|C)$  would simply be a sum of delta functions centered at the galaxy redshift, sky position, and apparent magnitude. However, we are not usually able to measure perfectly some of these quantities, in particular redshift, so we need to account for this.

Following [49], the probability of finding a galaxy at redshift  $z$ , sky position  $\Omega$ , and apparent magnitude  $m$  given the observed survey can be written as:

$$\begin{aligned} p_{\text{gal}}(z, \Omega, m|C) &= \int d\{z\}_{\text{gal}} d\{\Omega\}_{\text{gal}} d\{m\}_{\text{gal}} \times \\ &\times p_{\text{loc}}(z, \Omega, m|\{z\}_{\text{gal}}, \{\Omega\}_{\text{gal}}, \{m\}_{\text{gal}}, C) \times \\ &\times p_{\text{cat}}(\{z\}_{\text{gal}}, \{\Omega\}_{\text{gal}}, \{m\}_{\text{gal}}|C). \end{aligned} \quad (32)$$

In Eq. 32  $\{z_{\text{gal}}\}$ ,  $\{\Omega_{\text{gal}}\}$ , and  $\{m_{\text{gal}}\}$  indicate a set of *true* redshifts, sky positions, and apparent magnitudes for all the  $N_{\text{gal}}$  galaxies reported in the survey. The term  $p_{\text{loc}}$  is the probability of having a galaxy at  $z, \Omega, m$  once we know the true  $z_{\text{gal}}, \Omega_{\text{gal}},$  and  $m_{\text{gal}}$  of all the other galaxies. This probability does not actually depend on the galaxy catalog and is simply:

$$\begin{aligned} p_{\text{loc}}(z, \Omega, m|\{z\}_{\text{gal}}, \{\Omega\}_{\text{gal}}, \{m\}_{\text{gal}}) &= \\ &= \frac{1}{N_{\text{gal}}} \sum_{j=1}^{N_{\text{gal}}} \delta(z - z_{\text{gal}}^j) \delta(\Omega - \Omega_{\text{gal}}^j) \delta(m - m_{\text{gal}}^j). \end{aligned} \quad (33)$$

The term  $p_{\text{cat}}$  represents the probability that the galaxies have their true values of redshift, magnitude, and sky location of the galaxies  $\{z_{\text{gal}}\}$ ,  $\{\Omega_{\text{gal}}\}$ , and  $\{m_{\text{gal}}\}$ , given the *observed* galaxy survey. Generally, the galaxy catalog is made of a collection of galaxies, each of them observed *independently* and reported with *observed* apparent magnitude, redshift, and sky location  $\{m\}_{\text{obs}}, \{\Omega\}_{\text{obs}}, \{z\}_{\text{obs}}$  and typical uncertainties  $\{\sigma_m\}_{\text{obs}}, \{\sigma_\Omega\}_{\text{obs}}, \{\sigma_z\}_{\text{obs}}$ . Given the *observed* quantities  $C$ , the probability that a given galaxy  $j$  has a *true* redshift  $z_{\text{gal}}^j$ , sky location  $\Omega_{\text{gal}}^j$ , and apparent magnitude  $m_{\text{gal}}^j$  can be written as:

$$\begin{aligned} p(z_{\text{gal}}^j, \Omega_{\text{gal}}^j, m_{\text{gal}}^j|C) &= p(z_{\text{gal}}^j|z_{\text{obs}}^j, \sigma_{z,\text{obs}}^j) \times \\ &\times p(\Omega_{\text{gal}}^j|\Omega_{\text{obs}}^j, \sigma_{\Omega,\text{obs}}^j) \times \\ &\times p(m_{\text{gal}}^j|m_{\text{obs}}^j, \sigma_{m,\text{obs}}^j), \end{aligned} \quad (34)$$

and since galaxies are observed independently from each other, we can write:

$$\begin{aligned} p_{\text{cat}}(\{z\}_{\text{gal}}, \{\Omega\}_{\text{gal}}, \{m\}_{\text{gal}}|C) &= \prod_{j=0}^{N_{\text{gal}}} p(z_{\text{gal}}^j|z_{\text{obs}}^j, \sigma_{z,\text{obs}}^j) \times \\ &\times p(\Omega_{\text{gal}}^j|\Omega_{\text{obs}}^j, \sigma_{\Omega,\text{obs}}^j) p(m_{\text{gal}}^j|m_{\text{obs}}^j, \sigma_{m,\text{obs}}^j). \end{aligned} \quad (35)$$

By substituting Eqs. 33 and 35 into Eq. 32, we find that the reconstructed galaxy number density is:

$$\begin{aligned} \frac{dN_{\text{gal,cat}}}{dzd\Omega dm} &= \sum_j^{N_{\text{gal}}(\Omega)} p(z|z_{\text{obs}}^j, \sigma_{z,\text{obs}}^j) p(\Omega|\Omega_{\text{obs}}^j, \sigma_{\Omega,\text{obs}}^j) \times \\ &\times p(m|m_{\text{obs}}^j, \sigma_{m,\text{obs}}^j). \end{aligned} \quad (36)$$

We now use Eq. 36 to calculate the effective CBC emitters hosts in Eq. 30. If we now assume that the apparent magnitude and sky position of each galaxy are measured perfectly (this is a reasonable assumption for galaxy surveys), and that the sky is divided in equal-size pixels of area  $\Delta\Omega$ , we can write the effective CBC emitters hosts in a given sky direction as:

$$\begin{aligned} \frac{dN_{\text{gal,cat}}^{\text{eff}}}{dzd\Omega} &\approx \frac{1}{\Delta\Omega} \sum_j^{N_{\text{gal}}(\Omega)} 10^{0.4\epsilon(M_*(H_0) - M(H_0, m_j, z))} \times \\ &\times p(z|z_{\text{obs}}^j, \sigma_{z,\text{obs}}^j), \end{aligned} \quad (37)$$

where the sum is only on galaxies in a given sky direction. Eq. 37 is the relation needed to calculate the effective number of CBC hosts from the galaxy survey. The quantity  $p(z|z_{\text{obs}}^j, \sigma_{z,\text{obs}}^j)$  is the distribution of what true redshift we expect the galaxy to have, namely

$$p(z|z_{\text{obs}}^j, \sigma_{z,\text{obs}}^j) \propto \mathcal{N}(z_{\text{obs}}^j, \sigma_{z,\text{obs}}^j) \frac{dV_c}{dz}. \quad (38)$$

where  $\mathcal{N}$  is a Gaussian. In Eq.38, we have assumed that the galaxy redshift is measured from the catalog with gaussian uncertainty and that we apply a prior proportional to  $\frac{dV_c}{dz}$  [49].

We now make the following remarks. Eq. 37 does not depend on  $H_0$  for the same motivations explained for the out-of-catalog part, see the discussion below Eq. 27. Therefore, Eq. 37 can be computed only once for each observed galaxy catalog, thus decreasing significantly the computational cost of the analysis. Strictly speaking, as discussed in [49], in principle, the *true* redshift of each galaxy should be treated as a ‘‘population’’ parameter for the hierarchical inference. In other words, in the limit that the galaxy redshifts are not observed perfectly, the effective density of CBC hosts can not be reconstructed as a sum of redshift probability density distributions centered around the redshift values provided by the galaxy catalog. However, in the limit that the number density of observed CBC events is much lower than the number density of CBC hosts, Eq. 37 should still be valid to perform the hierarchical inference. Finally, we adopt also the same color corrections used in [36] to calculate the absolute magnitude for the apparent magnitude.

#### IV. TESTS WITH WELL-LOCALIZED DARK SIRENS

In this section, we provide some working examples of the methodology discussed in the previous sections using the well-localized standard sirens GW170817 [1] and GW190814 [50]. We use the GLADE+ galaxy catalog [51] and the galaxy luminosities reported for the infrared K-band. GLADE+ already

reports redshift values which are corrected for galaxies peculiar motion. Additionally, we consider the peculiar motion uncertainties addition by summing it in quadrature to the reported spectroscopic or photometric redshift uncertainty. In this section we apply Eq. 9 to infer  $H_0$  only, while we fix all the other population parameters (see App. A for more details).

To compute the completeness correction in Eq. 27, we adopt a Schechter function specified by the parameters  $M_{\min} = -27.85$ ,  $M_{\max} = -19.84$ ,  $\alpha = -1.09$ ,  $\phi_* = 0.03 \text{ Mpc}^{-3}$  (for  $H_0 = 67.7 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ). To calculate the effective hosts' number from the galaxy catalog in Eq. 37, we subdivide the sky into 49152 pixels<sup>2</sup> and a choice of equal area of  $0.8 \text{ deg}^2$ . The choice of the pixel size is driven by the fact that it should be smaller than the localization areas of GW190814, which is  $18.5 \text{ deg}^2$  [50]. While for GW170817, since we will fix the sky position, we should take a pixel size large enough to contain the galaxies' localization errors and angular size.

To calculate the apparent magnitude threshold, we follow the same procedure as in [36]. The apparent magnitude threshold is computed all over the sky using equal-sized pixels of  $3.35 \text{ deg}^2$ . On one hand, the pixel size for the calculation of the apparent magnitude threshold should be small enough to appreciate the different characteristics of the surveys composing GLADE+. On the other hand, the pixel size should be high enough to contain a reasonable number of galaxies to calculate the apparent magnitude threshold. The apparent magnitude threshold is defined as the median of the apparent magnitude of the galaxies reported in each pixel. This is an over-conservative threshold that is applied to avoid any systematic that could be potentially introduced by galaxy catalogs selection biases. We find that in a pixel of  $3.35 \text{ deg}^2$  the median number of galaxies present is 80, which is enough to calculate an apparent magnitude threshold. All the galaxies fainter than the apparent magnitude threshold are removed from the calculation of the CBC host redshift density.

#### A. GW170817 as quasi-dark siren: the pencil beam approach

So far, GW170817 is the only GW source for which it was possible to detect an associated electromagnetic counterpart and identify its host galaxy: NGC4993 [52]. GW170817 is also the only standard siren observed with EM counterpart. Using the GW luminosity distance estimation and the redshift of NGC4993 it was possible to estimate  $H_0 = 70_{-8}^{+19} \text{ km s}^{-1} \text{ Mpc}^{-1}$  [53, 54]. An analysis of GW170817 as “dark siren” was also performed in [55] with the GLADE [56] catalog using B-band luminosities, obtaining  $H_0 = 77_{-18}^{+37} \text{ km s}^{-1} \text{ Mpc}^{-1}$  [55].

Here, we employ GW170817 as a “quasi-dark siren”, namely, we assume the sky position to be perfectly known but the host galaxy not identified. This is type of methodology is also referred to as the “pencil beam” approach [19].

We made this choice in order to show the connection between the framework presented in Sec. III and the various

TABLE I. Galaxies with K-band luminosities from the GLADE+ galaxy catalog reported in the  $0.8 \text{ deg}^2$  pixel of GW170817. From left to right columns, the reported values are observed redshift, redshift uncertainty (including peculiar motion), apparent magnitude in the K-band, and absolute magnitude calculated for a Planck15 [58] cosmology. The first entry is NGC4993, the host of GW170817. The apparent magnitude threshold for this particular pixel is  $m_{\text{thr}} = 13.68$ .

$z_{\text{obs}}$	$\sigma_{z,\text{obs}}$	$m_K$	$M_K$
0.0092	0.0005	9.3	-23.7
0.0301	0.0148	13.0	-22.6
0.0423	0.0014	12.9	-23.4
0.0613	0.0153	13.5	-23.7
0.0709	0.0152	13.0	-24.4
0.1207	0.0151	13.4	-25.1
0.1505	0.0151	13.4	-25.6

physical quantities involved. For this study, we use the parameter estimation (PE) samples from the IMRPHENOMPv2NRT waveform model with low spin priors<sup>3</sup> released for GWTC-1 [57]. The PE samples are generated by fixing GW170817 sky location to the one its EM counterpart. We take all the galaxies falling within the  $0.8 \text{ deg}^2$  pixel of GW170817's counterpart, for which we obtain an apparent magnitude threshold of  $m_{\text{thr}} = 13.68$  (defined as the median of the apparent magnitude threshold of all the galaxies reported in the pixel). After removing all the galaxies fainter than  $m_{\text{thr}}$  from GLADE+, we obtain only 7 galaxies with redshift spanning from 0.0092 to 0.15 in the pixel of GW170817. The observed redshifts and luminosities are reported in Tab. I, where the first row of the table is NGC4993, the host of GW170817. We can immediately notice that farther galaxies display a lower (brighter) absolute magnitude. This is due to the presence of a selection bias, for which at high redshift we can observe only the brightest galaxies, and which is accounted using the completeness correction.

Using the galaxies reported in Tab. I and the apparent magnitude threshold previously computed, we calculate the effective CBC hosts density present in the catalog (Eq. 37) and their completeness correction (Eq. 27). The top panels of Fig. 1 show a comparison between the effective number density of CBC hosts “in” and “out” of the galaxy catalog. The figure displays two cases: one where we assume that the CBC merger rate does not depend on the intrinsic galaxy luminosity ( $\epsilon = 0$ , left column), and one where we assume the CBC merger rate is proportional to the galaxy luminosity ( $\epsilon = 1$ , right column). Importantly, we note that in the  $\epsilon = 1$  case, the effective number of CBC hosts in the catalog is lower. This is because in this model galaxies fainter than  $M_*$ , which are the majority of galaxies according to the Schechter function, host fewer GW sources. We also note that in the  $\epsilon = 1$  model, the CBC completeness of the galaxy catalog (second row in Fig. 1) is higher since the brightest galaxies included in the catalog are the majority fraction of the CBC emitters. Finally, from Fig. 1 we notice that from the in-catalog part, only two

<sup>2</sup> calculated with healpy and a choice of nside=64

<sup>3</sup> <https://dcc.ligo.org/LIGO-P1800370/public>

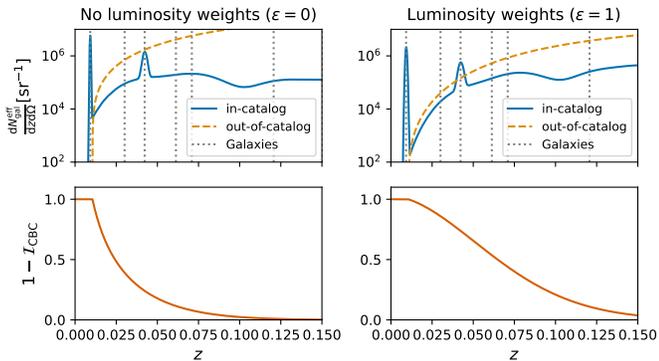


FIG. 1. *Top row*: Effective number of in-catalog (solid lines) and out-of-catalog (dashed line) CBC emitters per redshift per steradian for GW170817 as a quasi-dark siren. The blue solid line is the contribution from the galaxy catalog, while the orange dashed line is the completeness correction. The vertical dotted lines indicate the observed redshift values reported in Tab. I. *Bottom row*: CBC completeness  $1 - \mathcal{I}_{\text{CBC}}$ , with  $\mathcal{I}_{\text{CBC}}$  given by Eq. 28. The luminosity weights are assigned using a probability  $p(L|\epsilon) \propto L^\epsilon$ . The first column is generated using a CBC emission rate independent of galaxy luminosity ( $\epsilon = 0$ ), while the second is generated using a rate linearly proportional to the galaxy luminosity ( $\epsilon = 1$ ).

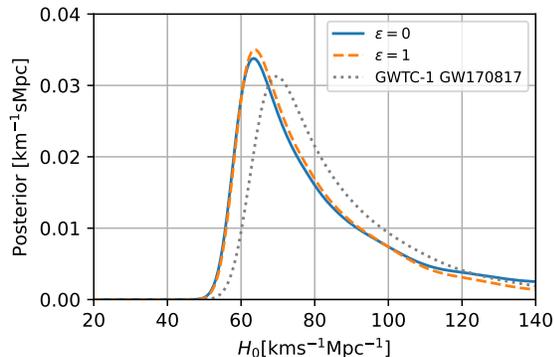


FIG. 2. Hubble constant posterior distribution obtained for GW170817 as quasi-dark siren. The blue solid and orange dashed lines indicate the posterior obtained using a CBC merger rate independent of the galaxies’ luminosity and proportional to it, respectively. The luminosity weights are assigned using a probability  $p(L|\epsilon) \propto L^\epsilon$ . The gray dotted line is the posterior of GW170817 obtained assuming the identification of the host galaxy with  $z = 0.0100 \pm 0.0005$  [59].

galaxies can be well localized in redshift and provide a redshift scale (the first peak is NGC4993). The other galaxies have large redshift uncertainties (cf. Tab. I) and will not significantly contribute to the measure.

We use the CBC host densities shown in Fig. 1 to calculate the CBC rate in Eq. 23, which is then used to evaluate the scale-free hierarchical likelihood in Eq. 9 and obtain a posterior distribution on the value of  $H_0$ . We consider a single population model for CBC rates in terms of masses and redshift. In App. A we provide more details about the CBC rate

models for masses and redshift used, as well as the corrections of selection biases. Fig. 2 shows the  $H_0$  posterior obtained for the different cases for which we computed the in-catalog and out-of-catalog CBC hosts densities. We observe that, for GW170817, the  $H_0$  measure is nearly independent of the choice of using or not luminosity weights. This is a consequence of the fact that the only structure in the CBC host density contributing to the  $H_0$  estimation is the one associated with NGC4993. Compared to the posterior obtained by the LIGO-Virgo Collaboration using GW posterior samples from GWTC-1 [59], we obtain an  $H_0$  posterior slightly narrower with a maximum a posteriori shifted to smaller values. This is due to the fact that in GLADE+, NGC4993 redshift is slightly lower ( $z = 0.0092 \pm 0.0005$ ) than the value of  $z = 0.0100 \pm 0.0005$  used for the analysis in [59]. Moreover, in the pencil-beam approach, the  $H_0$  inference gets some contribution from galaxies not well localized at high redshifts.

Moreover, in our analysis, we also have the contribution from galaxies at higher redshifts which are not well localized (see Fig. 1): this introduces a slight preference for lower values of  $H_0$ .

## B. GW190814 as a typical dark siren

GW190814 [50] is one of the best localized GW events so far, second only to GW170817. The 90% credible interval sky area of GW190814 covers about  $20 \text{ deg}^2$  and its localization volume (assuming the Planck15 cosmology [58]) is about  $900 \text{ Mpc}^3$  [36]. Previous  $H_0$  studies for this event using the galaxy catalog method with the GLADE (B-band) and the Dark Energy Survey [60] found a posterior with a maximum a posteriori in the Hubble constant tension region. However, it was not possible to constrain the  $H_0$  value in the prior range of  $[20, 140] \text{ km s}^{-1} \text{ Mpc}^{-1}$ . One of the peculiar aspects of this event is that the mass of the lighter object falls between the expected gap between neutron stars and black hole masses.

In our test case, we classify GW190814 as a massive neutron star black hole merger. Note that, when considering cosmology unknown we are intrinsically changing the source frame mass of the secondary object, so one should use a population model that includes neutron stars and black holes altogether. This is due to the fact that the secondary object might end up in the neutron star or black hole mass region depending on the cosmological model. However, we do not expect results for this event to strongly depend on the population model of the neutron star masses. For the prior extremes of  $H_0$  assumed, this event would change its source frame mass from a median value of  $2.6 M_\odot$  (for Planck cosmology),<sup>4</sup> to a range of  $2.46 M_\odot - 2.67 M_\odot$  (depending on which  $H_0$  extreme we are using). This range is well between the “massive” neutron star region and it is also smaller than the typical statistical uncertainties.

<sup>4</sup> value taken from <https://gwosc.org/eventapi/html/all/events/>

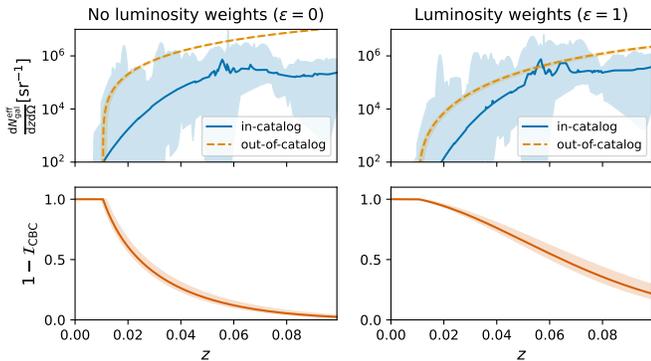


FIG. 3. *Top row*: Effective number of in-catalog (solid lines) and out-of-catalog (dashed line) CBC emitters per redshift per steradian in the 90% credible interval sky area of GW190814. *Bottom row*: CBC completeness  $1 - I_{\text{CBC}}$ , with  $I_{\text{CBC}}$  given by Eq. 28. The shaded areas indicate the contours identified by the 90% credible sky area of GW190814, while the lines correspond to the median values. The luminosity weights are assigned using a probability  $p(L|\epsilon) \propto L^\epsilon$ . The first column is generated using a CBC emission rate independent of galaxy luminosity ( $\epsilon = 0$ ), while the second is generated using a rate linearly proportional to the galaxy luminosity ( $\epsilon = 1$ ).

The population model we fix for neutron star black hole mergers is described in App. A. For the GW190814 PE samples, we use the “Publication Samples”<sup>5</sup> that are a combination of the samples obtained with the two IMRPHENOM and EOB waveform families released with GWTC-2 [61].

Consistently with [36], we find that GW190814’s 90% credible sky localization contains 90 galaxies with K-band luminosities reported in GLADE+, while we find a median apparent magnitude threshold equal to  $m_{\text{thr}} = 13.55$ . Fig. 3 displays the effective number density of in-catalog and out-of-catalog CBC hosts calculated for different models. Given the larger sky area covered by GW190814, in Fig. 3 we show the 90% credible intervals of the different quantities. We observe that there is an overdensity of galaxies in the redshift region from 0.05 to 0.06, which is the redshift expected for GW190814 if one assumes a cosmological model with  $H_0$  priors in the Hubble tension region.

In Fig. 4 we show the posteriors obtained for  $H_0$  using different prescriptions for the galaxy catalog. We note that when we assume a CBC rate proportional to the galaxies’ intrinsic luminosity ( $\epsilon = 1$ ), the Hubble constant posterior displays a local maximum in the  $H_0$ -tension region. The local maximum is still present, but less pronounced when a constant CBC rate with respect to the galaxies’ luminosity is assumed ( $\epsilon = 0$ ), the reason being that the CBC completeness is lower for the uniform rate, so that the out-of-catalog term washes out the impact of redshift features introduced by the in-catalog term.

Regarding the comparison with the GWTC-2 estimate, the analysis in [36] also uses a sky pixel of  $0.8 \text{ deg}^2$ , but it makes the choice of considering pixels with less than 10 galaxies as

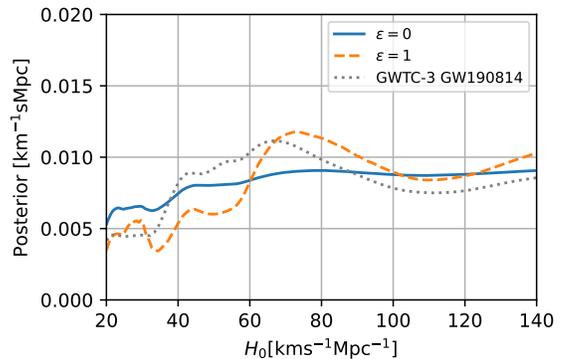


FIG. 4. Hubble constant posterior distribution obtained for the dark siren GW190814. The blue solid and orange dashed lines indicate the posterior obtained using a CBC merger rate independent of the galaxies’ luminosity and proportional to it. The luminosity weights are assigned using a probability  $p(L|\epsilon) \propto L^\epsilon$ . The gray dotted line is the posterior of GW190814 obtained by the LVK collaboration in [36].

“empty”. This choice translates to a higher incompleteness of the galaxy catalog for GW190814, resulting in a less informative  $H_0$  posterior.

## V. APPLICATION TO BINARY BLACK HOLES IN GWTC-3

We now apply our framework to probe the current expansion rate of the Universe using BBHs reported in GWTC-3. In [36, 41], it has been shown that BBHs can be used to probe cosmology either by using the standard galaxy catalog approach, or knowledge from the source-frame mass distribution. Although in [36, 41] these two aspects are considered separately, it is important to consider them conjointly [28], in order to obtain an unbiased estimate of the cosmological parameters. For the analyses presented in this section, and consistently with [36], we use galaxies with K-band luminosities reported in the GLADE+ catalog. To calculate the apparent magnitude threshold, we divide the sky into pixels of equal area<sup>6</sup> of  $13 \text{ deg}^2$ . This resolution is suitable for the best localized BBHs detections, which are usually constrained in a sky region larger than  $50 - 100 \text{ deg}^2$  [36]. Smaller pixel areas will not affect the results also because, as we will see later, most of the  $H_0$  information will actually come from the completeness correction. In this section, we apply Eq. 6 to infer  $H_0$  jointly with the population parameters.

<sup>5</sup> <https://dcc.ligo.org/LIGO-P2000223/public>

<sup>6</sup> calculated with `healpy` and `nside=16`.

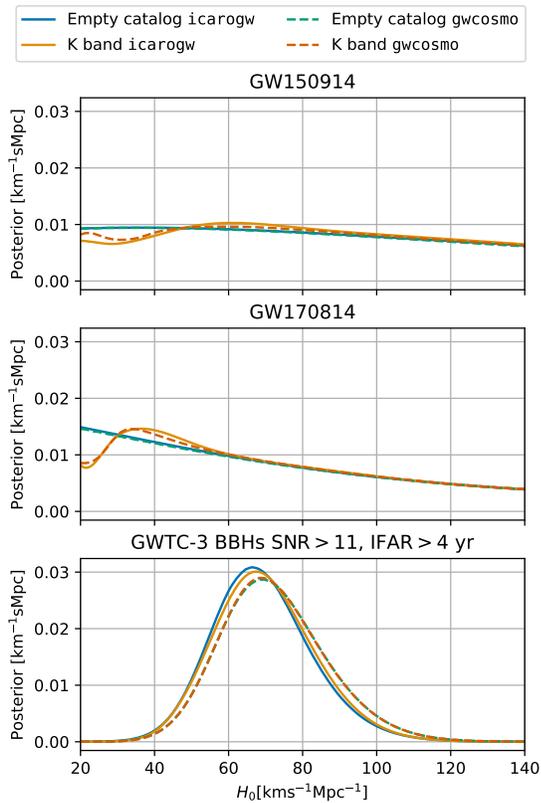


FIG. 5. Hubble constant posteriors distributions for an analysis of the 42 BBHs from GWTC-3 with SNR above 11 and IFAR higher than 4, fixing the BBH population model in the inference. The dashed lines are the  $H_0$  posterior obtained by gwcosmo in [36], while the solid lines are the ones obtained in this work. The “empty catalog” assumes a 0% complete galaxy catalog, where the inference is driven mostly by BBH population assumptions. *Top panel*:  $H_0$  posterior for GW150914. *Middle panel*:  $H_0$  posterior for GW170814. *Bottom panel*:  $H_0$  posterior for the combined 42 BBHs considered in the analysis.

### A. Catalog-only benchmark analysis

We first perform a benchmark analysis. We fix the BBH population model (source mass spectrum) to the one used by the galaxy catalog analysis in [36] (see App. A for the population models employed). We select 42 BBHs from GWTC-3 with a signal-to-noise ratio (SNR) above 11 and inverse false alarm rate (IFAR) higher than 4 yr. In Fig. 5 we show the  $H_0$  posterior obtained in this analysis for two events (GW150914 and GW170814) and the combination of all 42 dark sirens. In general, we find a good agreement between the results in [36] generated by gwcosmo and our approach. The plots show two test cases: the “empty catalog” case and the “full catalog” case. The former is generated by artificially assuming that the galaxy catalog contains no galaxies, while the latter uses the entire galaxy catalog. The empty catalog case is used to quantify the impact of population assumptions as the redshift information is dominated by the completeness correction. Compar-

ing the empty and full catalog curves in Fig. 5, we see that, for the combined posterior, the  $H_0$  inference is completely dominated by the BBH population prescription, while on a single-event level, adding redshift information from the galaxy catalog we are able to disfavor lower values of  $H_0$ . This is explained by the fact that for low  $H_0$  values, the redshift of the GW events is lower and is located in a region of the catalog that is more complete.

### B. Full catalog and BBH population analysis

Next, we perform a joint fit of  $H_0$  and BBH population properties, this time using information also from the galaxy catalog. Following [36], for the BBH population model, we adopt a POWER LAW + PEAK phenomenological mass model and a Madau-Dickinson [62] model  $\psi(z; \Lambda_p) = \psi_{\text{MD}}(z; \gamma, k, z_p)$  for the CBC merger rate as a function of redshift. Also for this analysis, we selected 42 BBHs from GWTC-3 with SNR above 11 and IFAR higher than 4 yr. More details about the redshift-related parameters as well as the BBH population models with their prior ranges can be found in App. A. In this analysis, we model the CBC merger rate per galaxy as:

$$\frac{dN_{\text{CBC}}}{dN_{\text{gal}}dt_s} = R_{\text{gal},0}^* \left( \frac{L}{L_*} \right)^\epsilon \psi_{\text{MD}}(z; \gamma, k, z_p). \quad (39)$$

We consider three scenarios characterized by different assumptions on the luminosity weight:  $\epsilon = 1$ ,  $\epsilon = 0$ , and  $\epsilon = 1$  but with the further assumption that the galaxy catalog is 0% complete (empty catalog case). note that for the “empty catalog” case, changing  $\epsilon$  would only result in a different  $R_{\text{gal},0}^*$ , the GW population parameters are only determined by the completeness correction.

In all the cases, we find that the population parameters related to the mass distribution do not strongly depend on the prescription used for the galaxy catalog (see Tab. II). This is a consequence of the fact that our CBC mass rate model in Eq. 16 does not explicitly depend on galaxies’ redshift and absolute magnitude.

We also note that the parameters related to the shape of CBC rate as a function of redshift  $\psi_{\text{MD}}$  do not strongly depend on the catalog prescription (see Tab. III). Redshift-related population parameters are independent of the catalog prescription since (i) GW events are in general not precisely localised in the redshift space, thus they are associated to thousands of galaxies and (ii) the galaxy catalog is highly incomplete at  $z > 0.1$ , and the CBC redshift rate contribution is mostly calculated with the completeness correction (dependent only on the comoving volume).

Let us now focus on the estimation of  $R_{\text{gal},0}^*$ , the CBC merger rate per galaxy with a reference luminosity  $L_*$ , jointly with  $H_0$ . In Fig. 6 we report the joint posterior distribution for the 42 BBHs selected in the analysis and the various galaxy catalog and rate prescriptions. Firstly, we find that the value of  $R_{\text{gal},0}^*$  is strongly dependent on the prescription used for the CBC merger rate as a function of the galaxy luminosity ( $\epsilon$ ). In particular, the CBC merger rate per galaxy should be

TABLE II. Maximum a posteriori and symmetric 68.3% credible intervals of the marginal distributions for the mass-related population parameters using 3 galaxy catalog prescriptions used in the analysis of the 42 BBHs from GWTC-3 having SNR above 12 and IFAR higher than 4 yr. The reported population parameters are used to build the BBH mass spectrum, see App. A for more details. Parameters with no entry are not constrained within their prior ranges.

Model	$\alpha$	$\beta$	$m_{\min}$	$m_{\max}$	$\delta_m [M_\odot]$	$\mu_g [M_\odot]$	$\sigma_g [M_\odot]$	$\lambda$
$\epsilon = 1$	$3.85^{+0.50}_{-0.52}$	$0.7^{+1.3}_{-1.1}$	$5.17^{+0.69}_{-0.95}$	-	$4.8 \pm 2.4$	$32.3^{+2.8}_{-4.3}$	$1.9^{+3.4}_{-1.5}$	$0.019^{+0.030}_{-0.017}$
$\epsilon = 0$	$3.81^{+0.49}_{-0.48}$	$0.7 \pm 1.2$	$5.16^{+0.66}_{-0.96}$	-	$4.7^{+2.9}_{-1.8}$	$33.2^{+2.8}_{-4.3}$	$2.8^{+3.0}_{-2.3}$	$0.018^{+0.029}_{-0.016}$
Empty ( $\epsilon = 1$ )	$3.82^{+0.46}_{-0.51}$	$0.56^{+1.47}_{-0.97}$	$4.99^{+0.95}_{-0.68}$	-	$5.0^{+2.4}_{-2.2}$	$33.0^{+3.0}_{-4.0}$	$2.8^{+3.0}_{-2.4}$	$0.019^{+0.030}_{-0.017}$

TABLE III. Maximum a posteriori and symmetric 68.3% credible intervals of the marginal distributions of the cosmology and redshift-related parameters for the 3 galaxy catalog prescriptions used in the analysis of the 42 BBHs from GWTC-3 having SNR above 12 and IFAR higher than 4 yr. The reported population parameters are used to build the BBH rate as function of redshift, see App. A for more details. Parameters with no entry are not constrained within their prior ranges.

Model	$H_0$ [km s <sup>-1</sup> Mpc <sup>-1</sup> ]	$\gamma$	$k$	$z_p$	$\log_{10} \left[ \frac{R_{\text{gal},0}^*}{\text{yr}^{-1}} \right]$
$\epsilon = 1$	$71^{+35}_{-30}$	$4.5^{+2.0}_{-1.7}$	-	-	$-5.67^{+0.63}_{-0.50}$
$\epsilon = 0$	$43^{+48}_{-18}$	$4.9^{+2.2}_{-2.1}$	-	-	$-6.10^{+0.81}_{-0.51}$
Empty ( $\epsilon = 1$ )	$52^{+36}_{-25}$	$4.9^{+2.3}_{-2.1}$	-	-	$-5.48^{+0.80}_{-0.50}$

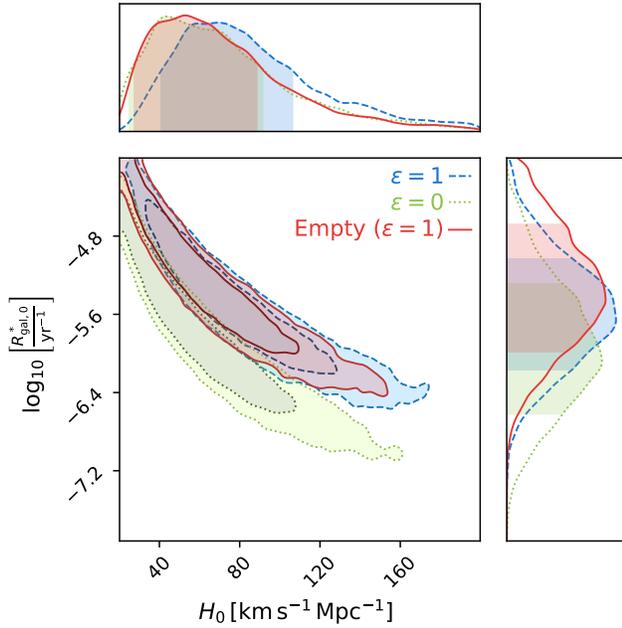


FIG. 6. Two dimensional and one-dimensional marginal distributions on the logarithm of the CBC merger rate per galaxy today  $R_{\text{gal},0}^*$  and  $H_0$ . The inner contour of the two-dimensional posterior indicates the 68.3% credible intervals while the outer contour is the 90%. The luminosity weights are assigned using a probability  $p(L|\epsilon) \propto L^\epsilon$ . The marginal 1-dimensional distribution reports the 68.3% credible intervals. Different colors correspond to the three galaxy catalog prescriptions.

higher when we assume that brighter galaxies are more likely to host GW events. This is expected as brighter galaxies are significantly lower in number than fainter galaxies, therefore to keep constant the expected number of CBC detections per year  $R_{\text{gal},0}^*$  must significantly increase. Secondly, we find that there is a strong anti-correlation between  $R_{\text{gal},0}^*$  and  $H_0$ . Cor-

relations are introduced when the variation of a certain combination of population-level parameters keeps the hierarchical likelihood in Eq. 1 (or Eq. 6) constant. For the particular case of  $R_{\text{gal},0}^*$  and  $H_0$ , the two parameters are degenerate for the calculation of the expected number of CBC detections  $N_{\text{exp}}$ . The expected number of CBC detections  $N_{\text{exp}}$  is linearly dependent on  $R_{\text{gal},0}^*$  and is also dependent on  $H_0$ , as for lower values of  $H_0$  the GW detection range can include galaxies at higher redshift. In other words:

$$N_{\text{exp}} \approx R_{\text{gal},0}^* T_{\text{obs}} N_{\text{gal}}(z < z(d_L^H, H_0)), \quad (40)$$

where  $d_L^H$  is a typical luminosity distance horizon for GW detection and  $N_{\text{gal}}(z < z(d_L^H, H_0))$  the number of galaxies enclosed in a redshift sphere with radius  $z(d_L^H, H_0)$ . The number of galaxies enclosed in the GW detection horizon will roughly scale as  $N_{\text{gal}} \propto [d_L^H H_0]^3$ . This implies that, to keep constant the value of  $N_{\text{exp}}$  we can either increase  $R_{\text{gal},0}^*$  and decrease  $H_0$  or viceversa. Note that this type of correlation is not usually observed between  $H_0$  and the CBC merger rate per comoving volume  $R_0$ , which is the parametrization typically used in population studies. This is due to the fact that the  $R_0$ -parameterization depends on the differential of the comoving volume (which scales as  $H_0^{-3}$ ):

$$N_{\text{exp}} \approx R_0 T_{\text{obs}} V_c(z < z(d_L^H, H_0)), \quad (41)$$

where  $V_c(z < z(d_L^H, H_0))$  is the comoving volume within a redshift shell with radius  $z(d_L^H, H_0)$ . As the comoving volume scales as  $V_c \propto [z(d_L^H, H_0)]^3 / H_0^3$  and  $z(d_L^H, H_0) \propto d_L^H H_0$ , then  $V_c$  will not strongly scale with  $H_0$ . As a consequence  $R_0$  is not correlated with  $H_0$  during the inference. In App. B we discuss in more detail how  $R_0$  and  $R_{\text{gal},0}^*$  are related to each other.

Let us finally comment about  $H_0$ . On one hand, the marginal posterior for  $H_0$  is almost identical in the empty catalog and  $\epsilon = 0$  case. In fact, as we have shown in Sec. IV, assuming  $\epsilon = 0$  corresponds to a galaxy catalog that is significantly less complete from the CBC hosts' point of view

(see Fig. 3 as an example). On the other hand, assuming  $\epsilon = 1$  with the galaxy catalog we obtain a marginal  $H_0$  posterior which disfavors low values of  $H_0$ . This is due to the fact that for small  $H_0$  values, GW events are located at lower redshifts, where the galaxy catalog is more complete (especially for  $\epsilon = 1$ ). Galaxy catalogs, if complete, can provide additional constraints to lower values of  $H_0$ , therefore it is crucial to use them to improve current constraints.

### C. GWTC-3 and modified gravity

Proposed modifications to general relativity (GR) that include a friction term for the GW propagation can result in different expressions for the GW luminosity distance  $d_L^{\text{GW}}$  and the classical (EM) luminosity distance  $d_L^{\text{EM}}$ . Different parametrizations of  $d_L^{\text{GW}}$  have been proposed. Here we focus on three models: the  $\Xi_0$ -model [63], the extra-dimensions model [64], and the running Planck mass model [65]. The  $\Xi_0$ -model adds the two phenomenological parameters  $\Xi_0$  and  $n_{\Xi_0}$  to describe the distance relation. The extra-dimensions model adds three parameters: the number of spacetime dimensions  $D$ , the screening distance  $R_c$ , and a scaling factor  $n_D$ . The running Planck mass adds only one extra parameter  $c_M$ . These models parametrize the relation between  $d_L^{\text{GW}}$  and  $d_L^{\text{EM}}$  as:

$$d_L^{\text{GW}} = d_L^{\text{EM}} \left[ \Xi_0 + \frac{1 - \Xi_0}{(1+z)^{n_{\Xi_0}}} \right], \quad (42)$$

$$d_L^{\text{GW}} = d_L^{\text{EM}} \left[ 1 + \left( \frac{d_L^{\text{EM}}}{(1+z)R_c} \right)^n \right]^{\frac{D-4}{2n_D}}, \quad (43)$$

$$d_L^{\text{GW}} = d_L^{\text{EM}} \exp \left[ \frac{c_M}{2} \int_0^z \frac{1}{(1+z')E^2(z')} dz' \right]. \quad (44)$$

Bright sirens are often used to put constraints on these models (see [31] and references therein). However, even the galaxy catalog method [66] and the spectral siren method [30, 31, 41, 67] can be used.

Here we repeat the full spectral siren and spectral+catalog analyses for these three models. To do so, we use the same selection of events used in Sec. V A and Sec. V B. This time, we fix the value of  $H_0$  to  $67.7 \text{ km s}^{-1} \text{ Mpc}^{-1}$  because the determination of  $H_0$  and the modified gravity parameters is strongly degenerate and with the current number of events we are unable to jointly constrain both. For the  $\Xi_0$ -model, we use a uniform prior in  $[0.1, 50]$  for  $\Xi_0$  and a uniform prior in  $[1, 10]$  for  $n_{\Xi_0}$ . For the extra dimension model, we use a uniform prior in  $[3.6, 8]$  for  $D$ , a log uniform prior in  $[0.1, 100]$  for  $n_D$ , and a log uniform prior in  $[10, 10^5] \text{ Mpc}$  for  $R_c$ . For the running Planck mass model, we use a uniform prior in  $[-10, 50]$ . The rest of the priors assumed for the population model are equal to the ones used in Sec. V B.

The marginal posterior distributions that we obtain on the modified gravity parameters are displayed in Figs. 7-8-9. As we can see from the marginal posterior distributions, the constraints on the modified gravity parameters are equal when only using source mass information (spectral) and the galaxy catalog information. This is due to the fact that the galaxy

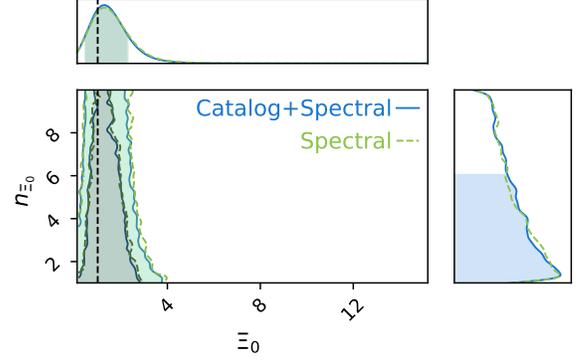


FIG. 7. Marginal posterior distribution for the parameters of the  $\Xi_0$  model for the catalog+spectral (blue solid line) and spectral (green dashed line) siren analyses. The black dashed line is the GR value ( $\Xi_0 = 0$ ). The plots report the 68.3% and 90% credible intervals. The marginal 1-dimensional distribution reports the 68.3% credible intervals.

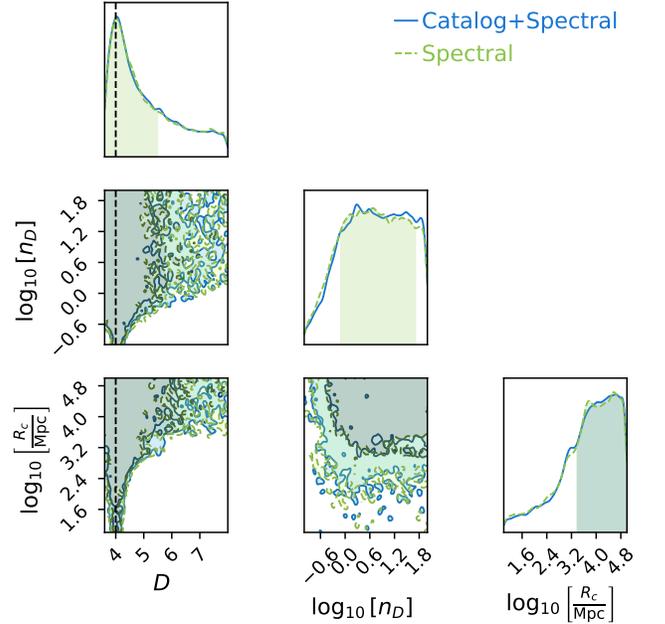


FIG. 8. Marginal posterior distribution for the parameters of the extra-dimensions model for the catalog+spectral (blue solid line) and spectral (green dashed line) siren analyses. The black dashed line is the GR value ( $D = 4$ ). The plots report the 68.3% and 90% credible intervals. The marginal 1-dimensional distribution reports the 68.3% credible intervals.

catalog is strongly incomplete and all the information on the redshift distribution is obtained from the source frame mass distribution. In all the cases, we obtain marginal posterior distributions consistent with no deviations from GR.

For the running Planck mass model, we obtain a constraint of  $1.0^{+2.6}_{-3.4}$  and  $1.6^{+2.2}_{-4.0}$  on  $c_M$  for the catalog+spectral and spectral siren analyses respectively (credible intervals reported at

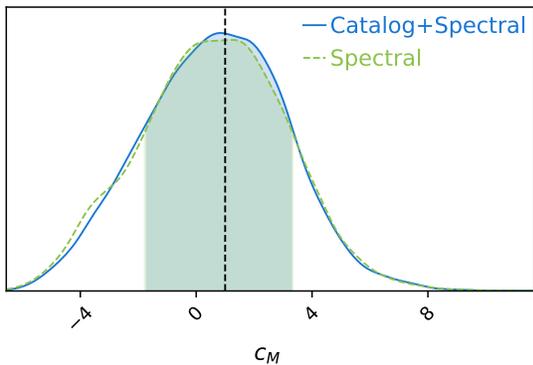


FIG. 9. Marginal posterior distribution for the parameter of the running Planck mass model for the catalog+spectral (blue solid line) and spectral (green dashed line) siren analyses. The black dashed line is the GR value ( $c_M = 1$ ). The plots report the 68.3% and 90% credible intervals. The marginal 1-dimensional distribution reports the 68.3% credible intervals.

68.3% credible interval unless stated otherwise). For the  $\Xi_0$  model we obtain a bound of  $1.44^{+1.17}_{-0.93}$  and  $1.37^{+1.36}_{-0.90}$  on  $\Xi_0$  for the catalog+spectral and spectral siren cases, but we are unable to constrain  $n_{\Xi_0}$ . For the extra dimensions model, we are not able to place strong bounds on any of the parameters as the number of space time dimensions and screening distance are strongly degenerate. If the screening distance is smaller than  $\sim 100$  Mpc, then the number of spacetime dimensions is strongly constrained to 4. The constraint on the rest of the population parameters is reported in Tab. IV

## VI. CONCLUSION

In this paper, we have presented a new methodology to jointly estimate the CBC merger rate and cosmological parameters using GW data and galaxies surveys.

In Sec. III, we presented a new parametrization for the CBC merger rate as a function of galaxies' number density. We have discussed how the galaxies' number density can be constructed from a real survey also accounting for a completeness correction and errors in the redshift evaluation of each galaxy. The proposed CBC merger rate parameterization introduces new astrophysical quantities such as  $R_{\text{gal},0}^*$ , the CBC merger rate per galaxy. The parametrization allows us to fit the number density of galaxies once for the entire analysis, thus making the study computationally feasible, while at the same time linking CBC physics to galaxy physics.

In Sec. IV and Sec. V, we have shown how this method could be applied to well-localized dark sirens and to a population of GW events, inferring both cosmology and population properties. We have discussed extensively the relation between  $H_0$  and the CBC merger rate per galaxy  $R_{\text{gal},0}^*$ . We constrained the value of  $\log_{10} R_{\text{gal},0}^*$  to  $-5.67^{+0.63}_{-0.50}$  and  $-6.10^{+0.81}_{-0.51}$  at 90% CI in the two scenarios where more luminous galaxies are more likely to emit GW events and where there is no such

dependency, respectively. We have also shown that, if more luminous galaxies are more likely to host GW sources, then we can exclude lower values of  $H_0$  just by using the 42 BBHs used in the LVK analysis in [36].

## ACKNOWLEDGMENTS

The authors are grateful for computational resources provided by the LIGO Laboratory and supported by National Science Foundation Grants PHY-0757058 and PHY-0823459. This research has made use of data or software obtained from the Gravitational Wave Open Science Center (gwosc.org), a service of LIGO Laboratory, the LIGO Scientific Collaboration, the Virgo Collaboration, and KAGRA. LIGO Laboratory and Advanced LIGO are funded by the United States National Science Foundation (NSF) as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, Spain. KAGRA is supported by Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan Society for the Promotion of Science (JSPS) in Japan; National Research Foundation (NRF) and Ministry of Science and ICT (MSIT) in Korea; Academia Sinica (AS) and National Science and Technology Council (NSTC) in Taiwan.

This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation.

The research of AG is supported by the Ghent University BOF project BOF/STA/202009/040 and the Fonds Wetenschappelijk Onderzoek (FWO) iBOF project BOF20/IBF/124. RG was supported by ERC starting grant SHADE 949572 and STFC grant ST/V005634/1.

TABLE IV. Constraints on the population parameters for the spectral and catalog+spectral analyses performed with 42 BBHs and various modified gravity models. The values are reported as the median with symmetric 68.3% credible intervals.

Model	$\alpha$	$\beta$	$m_{\min}$	$m_{\max}$	$\delta_m [M_\odot]$	$\mu_g [M_\odot]$	$\sigma_g [M_\odot]$	$\lambda$	$\gamma$	$k$	$z_p$
$\Xi_0$ (Catalog+Spectral)	$3.74^{+0.49}_{-0.41}$	$0.8^{+1.2}_{-1.1}$	$4.98^{+0.95}_{-0.64}$	-	$5.0^{+2.6}_{-2.2}$	$32.9^{+3.2}_{-2.2}$	-	$0.019^{+0.032}_{-0.018}$	-	-	-
Extra dimensions (Catalog+Spectral)	$3.80^{+0.45}_{-0.48}$	$0.6^{+1.2}_{-1.1}$	$5.37^{+0.55}_{-1.02}$	-	$4.6^{+2.7}_{-2.0}$	$33.1^{+2.1}_{-2.4}$	$2.3^{+3.0}_{-1.9}$	$0.017^{+0.031}_{-0.015}$	-	-	-
$c_M$ (Catalog+Spectral)	$3.72^{+0.50}_{-0.45}$	$0.41^{+1.49}_{-0.82}$	$5.08^{+0.78}_{-0.81}$	-	-	$32.9^{+2.7}_{-2.8}$	-	$0.018^{+0.031}_{-0.017}$	-	-	-
$\Xi_0$ (Spectral)	$3.71^{+0.49}_{-0.40}$	$0.7^{+1.4}_{-1.1}$	$5.14^{+0.77}_{-0.79}$	-	$5.2 \pm 2.4$	$33.3^{+2.7}_{-3.1}$	$3.7^{+3.2}_{-2.5}$	$0.018^{+0.035}_{-0.015}$	-	-	-
Extra dimensions (Spectral)	$3.86^{+0.42}_{-0.50}$	$0.7^{+1.2}_{-1.1}$	$5.13^{+0.84}_{-0.70}$	-	$4.9^{+2.5}_{-2.1}$	$33.0 \pm 2.3$	$2.4^{+3.0}_{-2.0}$	$0.016^{+0.032}_{-0.014}$	-	-	-
$c_M$ (Spectral)	$3.79^{+0.46}_{-0.50}$	$0.7^{+1.2}_{-1.1}$	$5.22^{+0.64}_{-0.90}$	-	$5.3^{+2.3}_{-2.4}$	$32.7^{+2.6}_{-2.7}$	$3.0^{+2.4}_{-2.6}$	$0.017^{+0.032}_{-0.015}$	-	-	-

## Appendix A: CBC mass and redshift models

### 1. Redshift model

In this paper, we employ a CBC rate model as a function of redshift following the Madau-Dickinson star formation rate [62]. We parametrize the CBC merger rate evolution  $\psi(z; \Lambda_p)$  as:

$$\psi_{\text{MD}}(z; \gamma, k, z_p) = [1 + (1 + z_p)^{-\gamma-k}] \frac{(1+z)^\gamma}{1 + \left(\frac{1+z}{1+z_p}\right)^{\gamma+k}}. \quad (\text{A1})$$

For the analysis with GW170817 in Sec. IV A we fix  $\gamma = 0, k = 0, z_p = 2.47$  a uniform in comoving volume rate. For the analysis with G190814 in Sec. IV B and the catalog analysis in Sec. V A we fix  $\gamma = 4.59, k = 2.86, z_p = 2.47$  [36], while for the full catalog and population analysis presented in Sec. V B we assume priors equal to the ones used in [36], see Tab. V.

Parameter	Description	Prior
$R_0$	BBH merger rate today in $\text{Gpc}^{-3} \text{yr}^{-1}$	$\mathcal{U}(0.0, 100.0)$
$\gamma$	Slope of the power law regime for the rate evolution before the point $z_p$	$\mathcal{U}(0.0, 12.0)$
$k$	Slope of the power law regime for the rate evolution after the point $z_p$	$\mathcal{U}(0.0, 6.0)$
$z_p$	Redshift turning point between the power law regimes with $\gamma$ and $k$	$\mathcal{U}(0.0, 4.0)$

TABLE V. Summary of the prior hyper-parameters used for the merger rate evolution models adopted in this paper.

### 2. Mass models

The source-frame mass models we use are composed of two statistical distributions: a truncated power law distribution:

$$\mathcal{P}(x|x_{\min}, x_{\max}, \alpha) \propto \begin{cases} x^\alpha, & (x_{\min} \leq x \leq x_{\max}), \\ 0, & \text{otherwise,} \end{cases} \quad (\text{A2})$$

and a Gaussian distribution with mean  $\mu$  and standard deviation  $\sigma$ :

$$\mathcal{G}(x|\mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]. \quad (\text{A3})$$

The distribution of the source frame masses  $m_1, m_2$  is factorized as:

$$\pi(m_1, m_2|\Phi_m) = \pi(m_1|\Lambda_p)\pi(m_2|m_1, \Lambda_p), \quad (\text{A4})$$

where  $\pi(m_1|\Phi_m)$  is the primary mass distribution (different for the various cases of the paper, as specified below). The secondary mass distribution  $\pi(m_2|m_1, \Phi_m)$  is a truncated power law distribution. The secondary mass is conditioned to the primary mass as  $m_2 < m_1$ :

$$\pi(m_2|m_1, m_{\min}, \alpha) = \mathcal{P}(m_2|m_{\min}, m_1, \beta). \quad (\text{A5})$$

Below we list the primary mass models and priors used in this paper:

- Sec. IV A (GW170817 as a quasi-dark siren): The primary mass model is a truncated power law with  $\alpha = 0, m_{\max} = 3M_\odot, m_{\min} = 1M_\odot$ ; the secondary mass model is a truncated power law with  $\beta = 0.81, m_{\max} = m_1, m_{\min} = 1M_\odot$  [36].
- Sec. IV B (GW190814 as dark siren): For the primary mass we use the POWER LAW+PEAK model from [36] to describe the source frame distribution. The POWER LAW+PEAK distribution is:

$$\pi(m_1|m_{\min}, m_{\max}, \alpha, \lambda_g, \mu_g, \sigma_g) = (1 - \lambda_g)\mathcal{P}(m_1|m_{\min}, m_{\max}, -\alpha) + \lambda_g\mathcal{G}(m_1|\mu_g, \sigma_g), \quad (\text{A6})$$

where the power law part has slope  $-\alpha$  between  $m_{\min}$  and  $m_{\max}$ , while the Gaussian component has mean  $\mu_g$  and standard deviation  $\sigma_g$  and accounts  $\lambda_g$  of the total fraction of the distribution. We also apply an additional smoothing at the lower edge of the distribution:

$$\pi(m_1, m_2|\Phi_m) = [\pi(m_1|\Phi_m)S(m_1|\delta_m, m_{\min})][\pi(m_2|m_1, \Phi_m)S(m_2|\delta_m, m_{\min})], \quad (\text{A7})$$

where  $S$  is a sigmoid-like window function as described in [39]. The population parameters fixed for the analysis are  $\alpha = 3.78, \beta = 0.81, m_{\max} = 112.5M_{\odot}, m_{\min} = 4.98M_{\odot}, \delta_m = 4.8M_{\odot}, \sigma_g = 3.88M_{\odot}, \mu_g = 342.27M_{\odot}, \lambda_g = 0.03$ . For the secondary mass (the neutron star) we use a truncated power law with  $m_{\min, \text{NS}} = 1M_{\odot}, m_{\max, \text{NS}} = 3M_{\odot}$  [36].

- Sec. V A (catalog-only BBH analysis): For the primary mass we use the same POWER LAW + PEAK population model used in Sec. IV B. For the secondary mass, we use a truncated power law with  $\beta = 0.81$  and  $m_{\min} = 4.88M_{\odot}, m_{\max} = m_1$ .
- Sec. V B (full catalog and population BBH analysis): the baseline model of the primary mass is always the POWER LAW + PEAK, but this time the population parameters of the model are also sampled using priors equal to the ones used in [36].

Parameter	Description	Prior
$\alpha$	Spectral index for the PL of the primary mass distribution.	$\mathcal{U}(1.5, 12.0)$
$\beta$	Spectral index for the PL of the mass ratio distribution.	$\mathcal{U}(-4.0, 12.0)$
$m_{\min}$	Minimum mass of the PL component of the primary mass distribution.	$\mathcal{U}(2.0M_{\odot}, 10.0M_{\odot})$
$m_{\max}$	Maximum mass of the PL component of the primary mass distribution.	$\mathcal{U}(50.0M_{\odot}, 200.0M_{\odot})$
$\lambda_g$	Fraction of the model in the Gaussian component.	$\mathcal{U}(0.0, 1.0)$
$\mu_g$	Mean of the Gaussian component in the primary mass distribution.	$\mathcal{U}(20.0M_{\odot}, 50.0M_{\odot})$
$\sigma_g$	Width of the Gaussian component in the primary mass distribution.	$\mathcal{U}(0.4M_{\odot}, 10.0M_{\odot})$
$\delta_m$	Range of mass tapering at the lower end of the mass distribution.	$\mathcal{U}(0.0M_{\odot}, 10.0M_{\odot})$

TABLE VI. Summary of the priors used for the population hyper-parameters for the three phenomenological mass models.

### 3. Correction of selection biases

To correct the selection biases, we need to calculate the expected number of detections

$$N_{\text{exp}}(\Lambda) = T_{\text{obs}} \int d\theta dz P_{\text{det}}(z, \theta, \Lambda_c) \frac{dN_{\text{CBC}}}{dzd\theta dt_s}(\Lambda) \frac{1}{1+z}. \quad (\text{A8})$$

The integral above can be approximated as a Monte Carlo sum over a set of simulated injections

$$N_{\text{exp}} \approx \frac{T_{\text{obs}}}{N_{\text{gen}}} \sum_{j=1}^{N_{\text{det}}} \frac{1}{\pi_{\text{inj}}(z_j, \theta_j)} \frac{1}{1+z_j} \frac{dN_{\text{CBC}}}{dzd\theta dt_s} \Big|_j \equiv \frac{T_{\text{obs}}}{N_{\text{gen}}} \sum_{j=1}^{N_{\text{det}}} s_j, \quad (\text{A9})$$

where  $T_{\text{obs}}$  is the observing time,  $N_{\text{gen}}$  the number of simulated signals (also the non-detected ones),  $N_{\text{det}}$  the number of detected signals, and  $\pi_{\text{inj}}(z_j, \theta_j)$  the prior used to generate the injections. Following Farr [68], we also define a numerical stability estimator for the selection bias. We require that the effective number of injections used to evaluate the selection bias  $N_{\text{eff, inj}}$  is 4 times higher than the observed number of GW signals. The effective number of injections is given by

$$N_{\text{eff, inj}} = \frac{\left[ \sum_j^{N_{\text{det}}} s_j \right]^2}{\left[ \sum_j^{N_{\text{det}}} s_j^2 - N_{\text{gen}}^{-1} \left( \sum_j^{N_{\text{det}}} s_j \right)^2 \right]}. \quad (\text{A10})$$

In this work, we used different injection sets. For the GW170817 analysis, we used the BNSs and NSBHs injections released after O3<sup>7</sup>. For the BBH analysis we use the injection set released with [36].

### Appendix B: The relation between $R_0$ and $R_{\text{gal},0}^*$ for the empty catalog case

In this Appendix we show how the CBC merger rate based on galaxy number density in Sec. III B relates to the spectral siren parameterization in Sec. III A. The CBC merger rate written in terms of galaxy number density can be written as

$$\frac{dN_{\text{CBC}}}{d\vec{m}_s d\theta d\Omega dz dt_s} = R_{\text{gal},0}^* P_{\text{pop}}(\vec{m}_s, \theta | \Lambda_p) \psi(z; \Lambda_p) \int dM 10^{0.4\epsilon(M_s - M)} \left[ \frac{dN_{\text{gal,cat}}}{dzd\Omega dM} + \frac{dN_{\text{gal,out}}}{dzd\Omega dM} \right]. \quad (\text{B1})$$

<sup>7</sup> <https://zenodo.org/record/5546676>

Let us assume that the catalog is completely empty and our rate is determined by the completeness correction; then we have:

$$\frac{dN_{\text{CBC}}}{d\vec{m}_s d\theta d\Omega dz dt_s} = R_{\text{gal},0}^* p_{\text{pop}}(\vec{m}_s, \theta | \Lambda_p) \psi(z; \Lambda_p) \phi_*(H_0) \Gamma_{\text{inc}}(\alpha + \epsilon + 1, x_{\text{max}}, x_{\text{min}}) \frac{dV_c}{dz d\Omega}, \quad (\text{B2})$$

where  $\phi_*(H_0) \Gamma_{\text{inc}}(\alpha + \epsilon + 1, x_{\text{max}}, x_{\text{min}})$  is simply the overall integral of the entire Schechter function (with the addition of the luminosity weight). The spectral siren rate parameterization is:

$$\frac{dN_{\text{CBC}}}{d\vec{m}_s d\theta d\Omega dz dt_s} = R_0 p_{\text{pop}}(\vec{m}_s, \theta | \Lambda_p) \psi(z; \Lambda_p) \frac{dV_c}{dz d\Omega}. \quad (\text{B3})$$

By equating Eq. B2 and Eq. B3, we obtain the relation between the spectral siren and galaxy rate, namely:

$$R_0 = R_{\text{gal},0}^* \phi_*(H_0) \Gamma_{\text{inc}}(\alpha + \epsilon + 1, x_{\text{max}}, x_{\text{min}}) \propto R_{\text{gal},0}^* H_0^3. \quad (\text{B4})$$

The above relation also shows why in the full population parameter estimation of Sec. V B  $R_{\text{gal},0}^*$  and  $H_0$  are correlated.

- [1] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Gw170817: Observation of gravitational waves from a binary neutron star inspiral, *Phys. Rev. Lett.* **119**, 161101 (2017).
- [2] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, *et al.*, Multi-messenger Observations of a Binary Neutron Star Merger, *ApJ* **848**, L12 (2017), arXiv:1710.05833 [astro-ph.HE].
- [3] J. Aasi *et al.* (LIGO Scientific), Advanced LIGO, *Class. Quant. Grav.* **32**, 074001 (2015), arXiv:1411.4547 [gr-qc].
- [4] F. Acernese *et al.* (VIRGO), Advanced Virgo: a second-generation interferometric gravitational wave detector, *Class. Quant. Grav.* **32**, 024001 (2015), arXiv:1408.3978 [gr-qc].
- [5] LIGO Scientific Collaboration, Virgo Collaboration, and KAGRA Collaboration, Gwtc-3: Compact binary coalescences observed by ligo and virgo during the second part of the third observing run (2021), arXiv:2111.03606 [gr-qc].
- [6] M. Oguri, Measuring the distance-redshift relation with the cross-correlation of gravitational wave standard sirens and galaxies, *Phys. Rev. D* **93**, 083511 (2016).
- [7] 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**, 1956 (2020), arXiv:1908.08951 [astro-ph.CO].
- [8] S. Mukherjee, B. D. Wandelt, S. M. Nissanke, and A. Silvestri, Accurate precision Cosmology with redshift unknown gravitational wave sources, *Phys. Rev. D* **103**, 043520 (2021), arXiv:2007.02943 [astro-ph.CO].
- [9] S. Bera, D. Rana, S. More, and S. Bose, Incompleteness Matters Not: Inference of  $H_0$  from Binary Black Hole–Galaxy Cross-correlations, *Astrophys. J.* **902**, 79 (2020), arXiv:2007.04271 [astro-ph.CO].
- [10] 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**, 2782 (2022), arXiv:2107.12787 [astro-ph.CO].
- [11] D. Chatterjee, A. Hegade K. R., G. Holder, D. E. Holz, S. Perkins, K. Yagi, and N. Yunes, Cosmology with Love: Measuring the Hubble constant using neutron star universal relations, *PhRvD* **104**, 083528 (2021), arXiv:2106.06589 [gr-qc].
- [12] T. Ghosh, B. Biswas, and S. Bose, Simultaneous inference of neutron star equation of state and the Hubble constant with a population of merging neutron stars, *Phys. Rev. D* **106**, 123529 (2022), arXiv:2203.11756 [astro-ph.CO].
- [13] C. Bambi, S. Katsanevas, and K. D. Kokkotas, eds., *Handbook of Gravitational Wave Astronomy* (Springer Singapore, 2021).
- [14] M. Moresco, L. Amati, L. Amendola, S. Birrer, J. P. Blakeslee, M. Cantiello, A. Cimatti, J. Darling, M. Della Valle, M. Fishbach, and *et al.*, Unveiling the Universe with emerging cosmological probes, *LRR* **25**, 6 (2022), arXiv:2201.07241 [astro-ph.CO].
- [15] B. F. Schutz, Determining the Hubble constant from gravitational wave observations, *Nature* **323**, 310 (1986).
- [16] D. E. Holz and S. A. Hughes, Using Gravitational-Wave Standard Sirens, *ApJ* **629**, 15 (2005), astro-ph/0504616.
- [17] N. Dalal, D. E. Holz, S. A. Hughes, and B. Jain, Short grb and binary black hole standard sirens as a probe of dark energy, *Phys. Rev. D* **74**, 063006 (2006), arXiv:astro-ph/0601275.
- [18] W. Del Pozzo, Inference of cosmological parameters from gravitational waves: Applications to second generation interferometers, *Phys. Rev. D* **86**, 043011 (2012).
- [19] H.-Y. Chen, M. Fishbach, and D. E. Holz, A two per cent Hubble constant measurement from standard sirens within five years, *Nature* **562**, 545 (2018), arXiv:1712.06531 [astro-ph.CO].
- [20] M. Fishbach *et al.* (LIGO Scientific, Virgo), A Standard Siren Measurement of the Hubble Constant from GW170817 without the Electromagnetic Counterpart, *Astrophys. J. Lett.* **871**, L13 (2019), arXiv:1807.05667 [astro-ph.CO].
- [21] R. Gray *et al.*, Cosmological inference using gravitational wave standard sirens: A mock data analysis, *Phys. Rev. D* **101**, 122001 (2020), arXiv:1908.06050 [gr-qc].
- [22] R. Gray, C. Messenger, and J. Veitch, A pixelated approach to galaxy catalogue incompleteness: improving the dark siren measurement of the Hubble constant, *Monthly Notices of the Royal Astronomical Society* **512**, 1127 (2022), <https://academic.oup.com/mnras/article-pdf/512/1/1127/45303118/stac366.pdf>.
- [23] H. Leandro, V. Marra, and R. Sturani, Measuring the hubble constant with black sirens, *Phys. Rev. D* **105**, 023523 (2022).
- [24] D. F. Chernoff and L. S. Finn, Gravitational Radiation, Inspiral Binaries, and Cosmology, *ApJ* **411**, L5 (1993), arXiv:gr-qc/9304020 [gr-qc].
- [25] S. R. Taylor, J. R. Gair, and I. Mandel, Cosmology using advanced gravitational-wave detectors alone, *Physical Review D* **85**, 10.1103/physrevd.85.023535 (2012).
- [26] W. M. Farr, M. Fishbach, J. Ye, and D. E. Holz, A future percent-level measurement of the hubble expansion at redshift 0.8 with advanced LIGO, *The Astrophysical Journal* **883**, L42 (2019).
- [27] J. María Ezquiaga and D. E. Holz, Jumping the gap: searching for LIGO’s biggest black holes, arXiv , arXiv:2006.02211 (2020), arXiv:2006.02211 [astro-ph.HE].
- [28] S. Mastroianni, K. Leyde, C. Karathanasis, E. Chassande-Mottin, D. A. Steer, J. Gair, A. Ghosh, R. Gray, S. Mukherjee, and S. Rinaldi, On the importance of source population models for gravitational-wave cosmology, *Phys. Rev. D* **104**, 062009 (2021), arXiv:2103.14663 [gr-qc].
- [29] S. Mukherjee, The redshift dependence of black hole mass distribution: is it reliable for standard sirens cosmology?, *Mon. Not. Roy. Astron. Soc.* **515**, 5495 (2022), arXiv:2112.10256 [astro-ph.CO].
- [30] J. M. Ezquiaga, Hearing gravity from the cosmos: GWTC-2 probes general relativity at cosmological scales, *Physics Letters B* **822**, 136665 (2021), arXiv:2104.05139 [astro-ph.CO].
- [31] K. Leyde, S. Mastroianni, D. A. Steer, E. Chassande-Mottin, and C. Karathanasis, Current and future constraints on cosmology and modified gravitational wave friction from binary black holes, *J. Cosmology Astropart. Phys.* **2022**, 012 (2022), arXiv:2203.11680 [gr-qc].
- [32] J. M. Ezquiaga and D. E. Holz, Spectral Sirens: Cosmology from the Full Mass Distribution of Compact Binaries, *Phys. Rev. Lett.* **129**, 061102 (2022), arXiv:2202.08240 [astro-ph.CO].
- [33] C. Karathanasis, S. Mukherjee, and S. Mastroianni, Binary black holes population and cosmology in new lights: Signature of PISN mass and formation channel in GWTC-3, arXiv , arXiv:2204.13495 (2022), arXiv:2204.13495 [astro-ph.CO].
- [34] X. Ding, M. Biesiada, X. Zheng, K. Liao, Z. Li, and Z.-H. Zhu, Cosmological inference from standard sirens without redshift measurements, *JCAP* **2019** (4), 033, arXiv:1801.05073 [astro-ph.CO].

- [35] C. Ye and M. Fishbach, Cosmology with standard sirens at cosmic noon, *PhRvD* **104**, 043507 (2021), arXiv:2103.14038 [astro-ph.CO].
- [36] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, *et al.*, Constraints on the cosmic expansion history from gwtc-3 (2021), arXiv:2111.03604 [astro-ph.CO].
- [37] I. Mandel, W. M. Farr, and J. R. Gair, Extracting distribution parameters from multiple uncertain observations with selection biases, *MNRAS* **486**, 1086 (2019), arXiv:1809.02063 [physics.data-an].
- [38] S. Vitale, D. Gerosa, W. M. Farr, and S. R. Taylor, Inferring the Properties of a Population of Compact Binaries in Presence of Selection Effects, in *Handbook of Gravitational Wave Astronomy. Edited by C. Bambi* (2022) p. 45.
- [39] R. Abbott *et al.* (LIGO Scientific, Virgo), Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog, *Astrophys. J. Lett.* **913**, L7 (2021), arXiv:2010.14533 [astro-ph.HE].
- [40] R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, N. Adhikari, R. X. Adhikari, V. B. Adya, C. Affeldt, D. Agarwal, and *et al.*, Population of Merging Compact Binaries Inferred Using Gravitational Waves through GWTC-3, *PhRvX* **13**, 011048 (2023), arXiv:2111.03634 [astro-ph.HE].
- [41] M. Mancarella, E. Genoud-Prachex, and M. Maggiore, Cosmology and modified gravitational wave propagation from binary black hole population models, *PhRvD* **105**, 064030 (2022), arXiv:2112.05728 [gr-qc].
- [42] M. Fishbach and D. E. Holz, Where Are LIGO's Big Black Holes?, *ApJ* **851**, L25 (2017), arXiv:1709.08584 [astro-ph.HE].
- [43] M. Fishbach, Z. Doctor, T. Callister, B. Edelman, J. Ye, R. Essick, W. M. Farr, B. Farr, and D. E. Holz, When Are LIGO/Virgo's Big Black Hole Mergers?, *ApJ* **912**, 98 (2021), arXiv:2101.07699 [astro-ph.HE].
- [44] L. A. C. van Son, S. E. de Mink, T. Callister, S. Justham, M. Renzo, T. Wagg, F. S. Broekgaarden, F. Kummer, R. Pakmor, and I. Mandel, The Redshift Evolution of the Binary Black Hole Merger Rate: A Weighty Matter, *ApJ* **931**, 17 (2022), arXiv:2110.01634 [astro-ph.HE].
- [45] S. Biscoveanu, T. A. Callister, C.-J. Haster, K. K. Y. Ng, S. Vitale, and W. M. Farr, The Binary Black Hole Spin Distribution Likely Broadens with Redshift, *ApJL* **932**, L19 (2022), arXiv:2204.01578 [astro-ph.HE].
- [46] S. S. Bavera, M. Fishbach, M. Zevin, E. Zapartas, and T. Fragos, The  $\chi_{eff} - z$  correlation of field binary black hole mergers and how 3G gravitational-wave detectors can constrain it, *A&A* **665**, A59 (2022), arXiv:2204.02619 [astro-ph.HE].
- [47] L. P. Singer *et al.*, Going the Distance: Mapping Host Galaxies of LIGO and Virgo Sources in Three Dimensions Using Local Cosmography and Targeted Follow-up, *Astrophys. J. Lett.* **829**, L15 (2016), arXiv:1603.07333 [astro-ph.HE].
- [48] M. R. Blanton, J. Dalcanton, D. Eisenstein, J. Loveday, M. A. Strauss, M. SubbaRao, D. H. Weinberg, J. Anderson, John E., J. Annis, N. A. Bahcall, and *et al.*, The Luminosity Function of Galaxies in SDSS Commissioning Data, *AJ* **121**, 2358 (2001), arXiv:astro-ph/0012085 [astro-ph].
- [49] J. R. Gair, A. Ghosh, R. Gray, D. E. Holz, S. Mastrogiovanni, S. Mukherjee, A. Palmese, N. Tamanini, T. Baker, F. Beirnaert, and *et al.*, The Hitchhiker's guide to the galaxy catalog approach for gravitational wave cosmology, arXiv, arXiv:2212.08694 (2022), arXiv:2212.08694 [gr-qc].
- [50] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, M. Agathos, and *et al.*, GW190814: Gravitational Waves from the Coalescence of a 23 Solar Mass Black Hole with a 2.6 Solar Mass Compact Object, *ApJL* **896**, L44 (2020), arXiv:2006.12611 [astro-ph.HE].
- [51] G. Dálya, R. Díaz, F. R. Bouchet, Z. Frei, J. Jasche, G. Lavaux, R. Macas, S. Mukherjee, M. Pálfi, R. S. de Souza, B. D. Wandelt, M. Bilicki, and P. Raffai, Glade+: An extended galaxy catalogue for multimessenger searches with advanced gravitational-wave detectors (2021), arXiv:2110.06184 [astro-ph.CO].
- [52] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and *et al.*, Multi-messenger Observations of a Binary Neutron Star Merger, *ApJL* **848**, L12 (2017), arXiv:1710.05833 [astro-ph.HE].
- [53] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and *et al.*, A gravitational-wave standard siren measurement of the Hubble constant, *Natur* **551**, 85 (2017), arXiv:1710.05835 [astro-ph.CO].
- [54] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, and *et al.*, Properties of the Binary Neutron Star Merger GW170817, *PhRvX* **9**, 011001 (2019), arXiv:1805.11579 [gr-qc].
- [55] M. Fishbach *et al.* (LIGO Scientific, Virgo), A Standard Siren Measurement of the Hubble Constant from GW170817 without the Electromagnetic Counterpart, *Astrophys. J. Lett.* **871**, L13 (2019), arXiv:1807.05667 [astro-ph.CO].
- [56] G. Dálya, G. Galgóczi, L. Dobos, Z. Frei, I. S. Heng, R. Macas, C. Messenger, P. Raffai, and R. S. de Souza, GLADE: A galaxy catalogue for multimessenger searches in the advanced gravitational-wave detector era, *MNRAS* **479**, 2374 (2018), arXiv:1804.05709 [astro-ph.HE].
- [57] B. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt, and *et al.*, Gwtc-1: A gravitational-wave transient catalog of compact binary mergers observed by ligo and virgo during the first and second observing runs, *Physical Review X* **9**, 10.1103/physrevx.9.031040 (2019).
- [58] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and *et al.*, Planck 2015 results. XIII. Cosmological parameters, *A&A* **594**, A13 (2016), arXiv:1502.01589 [astro-ph.CO].
- [59] LIGO Scientific Collaboration and Virgo Collaboration, A gravitational-wave measurement of the hubble constant following the second observing run of advanced ligo and virgo (2019), arXiv:1908.06060 [astro-ph.CO].
- [60] A. Palmese, J. deVicente, M. E. S. Pereira, J. Annis, and DES Collaboration, A Statistical Standard Siren Measurement of the Hubble Constant from the LIGO/Virgo Gravitational Wave Compact Object Merger GW190814 and Dark Energy Survey Galaxies, *ApJ* **900**, L33 (2020), arXiv:2006.14961 [astro-ph.CO].
- [61] R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, A. Adams, C. Adams, R. Adhikari, V. Adya, C. Affeldt, and *et al.*, Gwtc-2: Compact binary coalescences observed by ligo and virgo during the first half of the third observing run, *Physical Review X* **11**, 10.1103/physrevx.11.021053 (2021).
- [62] P. Madau and M. Dickinson, Cosmic Star-Formation History, *ARA&A* **52**, 415 (2014), arXiv:1403.0007 [astro-ph.CO].
- [63] E. Belgacem, Y. Dirian, S. Foffa, and M. Maggiore, Modified gravitational-wave propagation and standard sirens, *Phys. Rev. D* **98**, 023510 (2018).

- [64] M. Corman, A. Ghosh, C. Escamilla-Rivera, M. A. Hendry, S. Marsat, and N. Tamanini, Constraining cosmological extra dimensions with gravitational wave standard sirens: From theory to current and future multimessenger observations, *Phys. Rev. D* **105**, 064061 (2022), arXiv:2109.08748 [gr-qc].
- [65] M. Lagos, M. Fishbach, P. Landry, and D. E. Holz, Standard sirens with a running Planck mass, *Phys. Rev. D* **99**, 083504 (2019), arXiv:1901.03321 [astro-ph.CO].
- [66] A. Finke, S. Foffa, F. Iacovelli, M. Maggiore, and M. Mancarella, Cosmology with ligo/virgo dark sirens: Hubble parameter and modified gravitational wave propagation (2021), arXiv:2101.12660 [astro-ph.CO].
- [67] I. Magaña Hernandez, Constraining the number of spacetime dimensions from gwtc-3 binary black hole mergers, *Phys. Rev. D* **107**, 084033 (2023).
- [68] W. M. Farr, Accuracy Requirements for Empirically Measured Selection Functions, *Research Notes of the American Astronomical Society* **3**, 66 (2019), arXiv:1904.10879 [astro-ph.IM].