

Enhancing dark siren cosmology via Gaussian process reconstruction of incomplete galaxy catalogs

Matteo Tagliazucchi^{1,2,*}, Jonathan Gair³, Riccardo Barbieri³, and Michele Moresco^{1,4,2}

¹*Dipartimento di Fisica e Astronomia “Augusto Righi”—Università di Bologna,
Viale Berti Pichat 6/2, I-40127 Bologna, Italy*

²*INFN—Sezione di Bologna, Viale Berti Pichat 6/2, I-40127 Bologna, Italy*

³*Max Planck Institute for Gravitational Physics (Albert Einstein Institute),
Am Mühlenberg 1, Potsdam D-14476, Germany*

⁴*INAF—Osservatorio di Astrofisica e Scienza dello Spazio di Bologna,
via Piero Gobetti 93/3, I-40129 Bologna, Italy*



(Received 20 August 2026; accepted 29 September 2026; published 13 October 2026)

We present a novel framework for improving dark siren cosmology by applying a Gaussian process (GP) to the line-of-sight (LOS) reconstruction of incomplete galaxy catalogs. In the standard galaxy catalog method for inferring the Hubble constant H_0 from gravitational-wave (GW) dark sirens, missing galaxies are typically assumed to follow a uniform distribution in comoving volume, an assumption that discards galaxy redshift clustering information crucial for cosmological inference. We propose instead to model the LOS galaxy redshift distribution as a nonparametric function drawn from a GP realization, which is fitted to the observed incomplete catalog via a hierarchical Bayesian likelihood that explicitly accounts for the survey selection function. Applied to mock GW and galaxy catalogs extending up to redshift $z \leq 0.4$, our method yields H_0 constraints that are on average 23% more precise than the standard homogeneous completion when using a 24%-complete galaxy catalog, and 37% more precise for an 8%-complete catalog. The largest improvement, reaching 66%, is obtained in configurations where the GP most effectively reconstructs the redshift over- and underdensity features that the homogeneous completion fails to capture.

DOI: [10.1103/vmzf-1wxq](https://doi.org/10.1103/vmzf-1wxq)

I. INTRODUCTION

The current era of precision cosmology is characterized by a mounting tension in measurements of the Hubble constant H_0 , which quantifies the present-day expansion rate of the Universe. On one side, measurements based on the cosmic microwave background under the assumption of the standard Λ CDM model yield $H_0 = 67.4 \pm 0.5 \text{ km s}^{-1} \text{ Mpc}^{-1}$ [1], while local distance-ladder measurements using type-Ia supernovae calibrated with Cepheids report $H_0 = 73.0 \pm 1.0 \text{ km s}^{-1} \text{ Mpc}^{-1}$ [2]. This discrepancy, now reaching statistical significance exceeding 5σ , is referred to as the Hubble tension and may hint at physics beyond the standard cosmological model [3,4]. To address such tension, novel and independent cosmological probes, which may eventually be combined with each other to reduce systematic effects, are needed. Gravitational-wave

(GW) standard sirens have recently emerged as a new and independent cosmological probe, since their signals encode direct information about the luminosity distance to the source, without requiring any calibration. By combining these distance measurements with an estimate of the source redshift, the Hubble constant can be inferred using the standard distance-redshift relation.

The redshift information, however, cannot be obtained directly from the GW strain, since the latter is sensitive to the product of the source redshift and chirp mass, making these two perfectly degenerate. When an unambiguous electromagnetic (EM) counterpart is identified, the host galaxy provides a direct spectroscopic redshift and allows the mass-redshift degeneracy to be broken, as in the bright siren approach [5–9]. However, for most detected compact binary coalescences (CBCs)—particularly binary black hole (BBH) mergers—no EM counterpart is expected (and these are therefore called “dark sirens”). In the dark siren case, the redshift information may be extracted statistically from the distribution of potential host galaxies within the GW localization volume, as inferred from galaxy catalogs [the “galaxy catalog method,” [10–15]]. The galaxy catalog method relies on the assumption that CBCs occur in galaxies, so that the galaxy redshift distribution along the

*Contact author: matteo.tagliazucchi2@unibo.it

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

GW line-of-sight (LOS) provides a prior on the source redshift. In practice, this proceeds as follows: for each GW event, the galaxies within the three-dimensional GW localization volume are identified from a galaxy catalog; their redshift distributions are combined into a redshift prior $p_{\text{cbc}}(z)$; and the H_0 posterior is obtained by marginalizing the GW likelihood over this prior within a hierarchical Bayesian framework.

To correctly infer the Hubble constant using the galaxy catalog method, it is necessary to properly model the LOS redshift prior, including all possible host galaxies. However, all-sky galaxy surveys are affected by selection effects—for example, being magnitude-limited—and therefore cannot detect all galaxies. It is thus necessary to correct for galaxy catalog incompleteness to infer H_0 without introducing systematic biases [16–22]. Indeed, if an incomplete catalog is used directly without accounting for missing galaxies, the cosmological inference effectively downweights redshifts where the catalog is sparse, even if those redshifts are genuinely populated by potential host galaxies. The standard approach in current analyses models the contribution from missing galaxies by assuming they are uniformly distributed in comoving volume [13,23–29]. While this correctly accounts for the overall normalization of the galaxy number density, it implicitly washes out any clustering information present in the data: the over- and under-densities of the galaxy redshift distribution—which are precisely what make the dark siren method informative beyond a prior uniform in comoving volume [30]—are diluted in proportion to the catalog incompleteness. As a result, at high redshifts, where galaxy catalogs are inevitably highly incomplete, this approach may severely limit the cosmological constraining power of the dark siren method. Moreover, in the limit of a large number of GW events, modeling the galaxy catalog incompleteness in this way would inevitably induce a bias, since the assumed LOS redshift prior does not match the true underlying host distribution.

Several alternative approaches have been proposed to improve upon the uniform comoving volume completion. For example, [23] proposes a “multiplicative completion,” in which missing galaxies are assumed to trace the distribution of those present in the catalog. Another strategy incorporates clustering information directly into the incompleteness correction [31–35]. In particular, Leyde *et al.* [32,34] reconstruct the full galaxy density field—jointly with the galaxy magnitude distribution and detection probability—from a magnitude-limited catalog using Gaussian random fields as nonparametric priors, with a dark matter bias prescription connecting the reconstructed field to the observed galaxies. An alternative approach proposes to recover missing galaxies using the mean properties of the observed ones to keep clustering information at high redshifts [36]. Another strategy avoids the incompleteness problem partly by replacing standard

galaxy catalogs with more complete alternatives, such as galaxy cluster catalogs or considering only the brightest galaxy subset [37,38]. Alternatively, GW events can be combined with tracers of the large-scale matter distribution other than galaxies [39,40]. A distinct class of methods in which galaxy catalog incompleteness is less critical relies on the angular cross-correlation between GW events and galaxy catalogs [41–51].

In this work, we apply Gaussian processes (GPs) to model the LOS redshift prior distribution used in dark siren H_0 inference. Rather than assuming a parametric form for the galaxy distribution or a physically motivated model of galaxy clustering, we represent the unknown host distribution as a nonparametric function, corresponding to the realization of a GP, which specifies a prior distribution over smooth functions. Given an incomplete galaxy catalog, the GP is fitted to the observed galaxy redshift distribution via a hierarchical Bayesian likelihood that explicitly accounts for the survey selection function, and the resulting posterior over the galaxy distribution is subsequently used as the redshift prior in the H_0 inference. A key advantage of this approach is that the GP framework does not require any modeling of galaxy physics, such as halo occupation distributions, bias models, or the galaxy-matter cross-correlation, all of which carry significant uncertainties at the relevant scales. Instead, it learns the correlation structure of the galaxy distribution directly from the data, with the GP kernel parameters marginalizing over the degree of smoothness and the characteristic clustering scale. This makes the method robust to astrophysical uncertainties while remaining flexible enough to capture the clustering features that are essential for dark siren cosmology.

The paper is organized as follows. In Sec. II we describe the statistical framework adopted for H_0 inference. We first introduce the hierarchical Bayesian likelihood for GW events and the simplified GW toy model used throughout the analysis. We then describe three approaches to modeling the galaxy redshift distribution considered in this work: the idealized complete catalog case, the standard homogeneous incompleteness correction assuming missing galaxies are uniformly distributed in comoving volume, and the novel GP-based method. For the latter, we detail the construction of the galaxy likelihood, the GP prior on the redshift distribution, and the transformation applied to ensure positivity and physical boundedness of the reconstructed distribution. In Sec. III we present the results of the analysis applied to simulated datasets. We begin by describing the construction of the mock galaxy catalogs and the corresponding mock GW catalogs, generated for a range of detector and survey configurations. We then present the GP reconstruction of the galaxy redshift distribution and compare it to the true underlying distribution and to the homogeneous completion, assessing the impact of catalog incompleteness and redshift uncertainties. Finally, we apply

the reconstructed redshift priors to dark siren cosmology, presenting H_0 posterior distributions for all configurations studied. We conclude in Sec. IV with a summary of our findings and an outlook on future extensions of this work.

II. STATISTICAL FRAMEWORK

In the galaxy catalog method, the likelihood of observing a set of N_{obs} GW events given the Hubble constant H_0 can be described by an inhomogeneous Poisson process in the presence of selection effects [15,52,53]

$$\mathcal{L}(\{\mathbf{d}_i\}|H_0) = \prod_{i=1}^{N_{\text{obs}}} \frac{\int dz \mathcal{L}_{\text{gw}}(\mathbf{d}_i|d_L(z; H_0)) p_{\text{cbc}}(z)}{\int dz P_{\text{det}}^{\text{gw}}(z, H_0) p_{\text{cbc}}(z)}. \quad (1)$$

Here, $\{\mathbf{d}_i\}$ denotes the data from the observed events. The likelihood is evaluated by marginalizing each event's redshift using the CBC redshift prior $p_{\text{cbc}}(z)$, while the denominator accounts for GW selection effects. The expression in Eq. (1) represents a simplified version of the full hierarchical model [15,52,53], since it considers only the luminosity distances inferred from the GW signals, rather than incorporating masses and other waveform parameters, and restricts the inference to H_0 alone, treating all remaining population-level hyperparameters as fixed. While these simplifications reduce the generality of the model, they allow us to isolate the cosmological information encoded in the distance measurements and the redshift prior, thereby providing a simple and clear framework to introduce the novel methodology presented in the following sections. The quantity $p_{\text{cbc}}(z)$ entering Eq. (1) encodes the redshift distribution of GW sources and may be constructed from galaxy observations, as described in Sec. II B. We first specify the toy model adopted for the GW likelihood and selection function before turning to the construction of $p_{\text{cbc}}(z)$.

A. Gravitational-wave toy model

Since we use a simplified hierarchical framework for the inference of H_0 , we consider a toy model for the single-event GW likelihood that depends only on the observed luminosity distance $d_{L,i}^{\text{obs}}$ as

$$\mathcal{L}_{\text{gw}}(d_{L,i}^{\text{obs}}|d_L(z; H_0)) = \frac{1}{\sqrt{2\pi}\sigma_{d_L}} e^{-\frac{(d_{L,i}^{\text{obs}} - d_L(z; H_0))^2}{2\sigma_{d_L}^2}}, \quad (2)$$

where

$$\sigma_{d_L} = A d_L(z; H_0). \quad (3)$$

The factor A represents a constant fractional error.

The term $P_{\text{det}}^{\text{gw}}(z, H_0)$ appearing in Eq. (1) defines the probability of detecting a GW event at redshift z given H_0 . It is obtained by integrating the GW likelihood over the

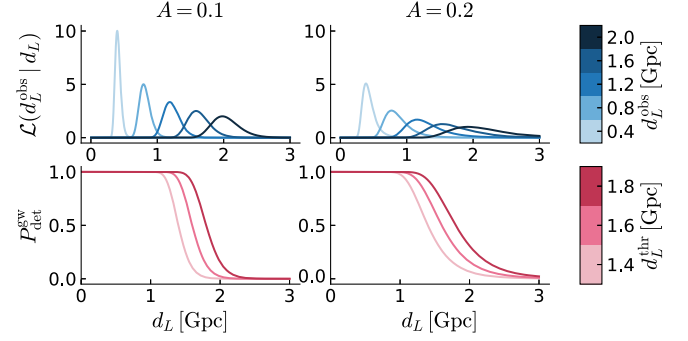


FIG. 1. Top: examples of the single-event Gaussian GW likelihood Eq. (2) as a function of d_L and two choices of the fractional error parameter A (0.1 and 0.2) and for different observed luminosity distances. Bottom: the corresponding detection probability Eq. (4) for the same two values of A and for three different thresholds d_L^{thr} .

space of detectable datasets. In this simple model, the “data” is the observed luminosity distance. We assume that a GW event is detected if this observed luminosity distance falls below a threshold value d_L^{thr} —representing the distance horizon of the network of detectors considered:

$$\begin{aligned} P_{\text{det}}^{\text{gw}}(z, H_0) &= \int_{-\infty}^{\infty} d\hat{d}_L \Theta(\hat{d}_L - d_L^{\text{thr}}) \mathcal{L}_{\text{gw}}(\hat{d}_L|d_L(z; H_0)) \\ &= \int_{-\infty}^{d_L^{\text{thr}}} d\hat{d}_L \mathcal{L}_{\text{gw}}(\hat{d}_L|d_L(z; H_0)) \\ &= \frac{1}{2} \left[1 + \text{erf} \left(\frac{d_L(z; H_0) - d_L^{\text{thr}}}{\sqrt{2} A d_L(z; H_0)} \right) \right]. \end{aligned} \quad (4)$$

In the above expression, $\Theta(\cdot)$ is the Heaviside step function and $\text{erf}(\cdot)$ is the error function.

Fig. 1 shows the GW likelihood and the GW detection probability for different values of A and d_L^{thr} .

B. Galaxy catalog modeling and incompleteness correction

The quantity $p_{\text{cbc}}(z)$ entering Eq. (1) represents the redshift distribution of the GW sources. This distribution is not known *a priori*, but can be constructed from galaxy observations, effectively using the posterior from electromagnetic (EM) observations as a prior for the GW data. Throughout the following analyses, we assume that CBC events occur in galaxies, so that $p_{\text{cbc}}(z)$ is given by the product of the probability of finding a galaxy at redshift z , $p_{\text{gal}}(z)$, and the probability that a galaxy at that redshift hosts a GW merger, $p_{\text{rate}}(z)$:

$$p_{\text{cbc}}(z) = \frac{p_{\text{gal}}(z) p_{\text{rate}}(z)}{\int dz p_{\text{gal}}(z) p_{\text{rate}}(z)}. \quad (5)$$

For simplicity, the merger rate is modeled as a uniform distribution with support in the redshift range $[0, z_{\text{cut}}]$,

where z_{cut} represents the edge of the mock galaxy catalogs used in the following analyses. We consider three different approaches to modeling the p_{gal} term. The first represents an idealized case in which the galaxy catalog is complete. In the second approach, we account for catalog incompleteness in the way commonly adopted in current analyses [25,26,29], by assuming that missing galaxies are uniformly distributed in comoving volume. Both of these approaches are described in this subsection. The third approach, introduced in Sec. II C, is a novel nonparametric method based on Gaussian processes (GP) that reconstructs the full galaxy distribution from an incomplete catalog while preserving clustering information, which is washed out by the uniform completion but is crucial for dark siren cosmology.

Complete galaxy catalog case. We first consider the case in which the redshift prior is built using a complete galaxy catalog. In this case, $p_{\text{gal}}(z)$ is simply given by the sum of the redshift posterior distributions for all N_g galaxies that fall within the localization volume of the GW event¹

$$p_{\text{gal}}(z) \equiv p_{\text{cat}}(z) = \frac{1}{N_g} \sum_{i=1}^{N_g} p(z|\tilde{z}_i, \sigma_{z,i}), \quad (6)$$

where $\tilde{z}_i$ is the observed galaxy redshift and $\sigma_{z,i}$ is its standard deviation error. The redshift posterior appearing in the above equation is the product of the galaxy likelihood—assumed to be a Gaussian with mean $\tilde{z}_i$ and standard deviation $\sigma_{z,i}$ —multiplied by a prior distribution and properly normalized. In the absence of any other galaxy information, the most conservative choice is a prior uniform in comoving volume for the galaxies [15]. The galaxy redshift posterior is then

$$p(z|\tilde{z}_i, \sigma_{z,i}) = \frac{\mathcal{N}(z; \mu = \tilde{z}_i, \sigma = \sigma_{z,i}) \frac{dV_c}{dz}(z)}{\int dz, \mathcal{N}(z; \mu = \tilde{z}_i, \sigma = \sigma_{z,i}) \frac{dV_c}{dz}(z)}. \quad (7)$$

Note that thanks to the proper normalization, this posterior is independent of H_0 .

Homogeneous incompleteness correction. We next consider the case in which the galaxy survey is affected by selection effects and cannot observe all galaxies. In this case, the galaxy catalog is incomplete and $p_{\text{gal}}(z)$ should be modeled as a sum of a term representing the redshift distribution inferred from the galaxies within the catalog and a term representing the redshift distribution of missing galaxies [12,23]

$$p_{\text{gal}}(z) = f_{\mathcal{R}} p_{\text{cat}}(z) + (1 - f_{\mathcal{R}}) p_{\text{miss}}(z), \quad (8)$$

where $f_{\mathcal{R}}$ is the average completeness fraction in the localization volume $\mathcal{R}$ of the GW event. This quantity is computed as the comoving volume-averaged probability that the galaxy survey detects a galaxy at z , $P_{\text{det}}(z)$:

$$f_{\mathcal{R}} = \frac{1}{V_c} \int dV_c P_{\text{det}}(z) = \frac{1}{V_c} \int dz P_{\text{det}}(z) \frac{dV_c}{dz}(z). \quad (9)$$

Note that this fraction is H_0 -independent since both V_c and dV_c/dz scale as H_0^{-3} , but depends on the other cosmological parameters. The term p_{cat} is the same as in Eq. (6), but now the sum extends only over all observed galaxies. In the homogeneous incompleteness correction, missing galaxies are assumed to be uniformly distributed in comoving volume [19,23,25,27,54]

$$p_{\text{miss}}(z) = \frac{1 - P_{\text{det}}(z)}{(1 - f_{\mathcal{R}}) V_c} \frac{dV_c}{dz}(z). \quad (10)$$

Therefore, the final expression for the redshift prior in the case of an incomplete galaxy catalog with a “homogeneous completion” is

$$p_{\text{gal}}(z) = \left[\frac{1}{V_c} \int dz P_{\text{det}}(z) \frac{dV_c}{dz}(z) \right] p_{\text{cat}}(z) + [1 - P_{\text{det}}(z)] \frac{1}{V_c} \frac{dV_c}{dz}(z). \quad (11)$$

While this approach correctly accounts for the overall level of catalog incompleteness, the assumption that unobserved galaxies trace the comoving volume uniformly discards any clustering information present in the data. As discussed in the following subsection, this limitation motivates the development of a more flexible, nonparametric approach, that will be described in Sec. II C.

Galaxy selection function. The incompleteness correction described previously, as well as the GP-based model that will be detailed in the next section, rely on the galaxy selection function, $P_{\text{det}}(z)$. We model this function in terms of the galaxy absolute magnitudes M . The likelihood of observing a galaxy with magnitude M given parameters $\lambda_{\mathcal{G}}$ is described by a Schechter function, $\phi(M|\lambda_{\mathcal{G}})$, normalized over the interval $[M_{\text{min}}, M_{\text{max}}]$ where it has support [55]

$$\mathcal{L}_{\text{EM}}(M|\lambda_{\mathcal{G}}) = \frac{\phi(M|\lambda_{\mathcal{G}})}{\int_{M_{\text{min}}}^{M_{\text{max}}} dM \phi(M|\lambda_{\mathcal{G}})}. \quad (12)$$

The Schechter function describes the number density of galaxies in the comoving volume per absolute magnitude. In the following analyses, we assume a Schechter function, nonevolving in redshift, defined by five parameters $\lambda_{\mathcal{G}}$: a normalization density ϕ_* , a knee absolute magnitude M_* , a

¹Technically, the sum is over all galaxies in the Universe and the restriction to galaxies in the GW localization volume happens when this is multiplied by the GW likelihood. In practice, it is common to restrict the sum to galaxies that are within the 3D localization volume of the source, for some value of H_0 within the prior range, in order to reduce computational costs.

faint end $M_{\max}$, a bright end $M_{\min}$, and a slope parameter α . Explicitly,

$$\phi(M|\lambda_G) = \begin{cases} 0.4 \ln(10) \phi_* x^{\alpha+1} e^{-x}, & M_{\min} \leq M \leq M_{\max} \\ 0, & \text{otherwise,} \end{cases} \quad (13)$$

with $x = 10^{0.4(M_* - M)}$.

Typically, Schechter parameters are quoted for a Hubble constant of $H_0 = 100h \text{ km s}^{-1} \text{ Mpc}^{-1}$. Rescaling to an arbitrary h follows the relations

$$\begin{aligned} M_{\min, \max, *} (h) &= M_{\min, \max, *} + 5 \log_{10} h \\ \phi_* (h) &= \phi_* h^3. \end{aligned} \quad (14)$$

The detection probability is obtained by integrating $\mathcal{L}_{\text{EM}}(M|\lambda_G)$ over the space of detectable magnitudes. A galaxy is considered detectable if its absolute magnitude is brighter than a threshold $M_{\text{thr}}(z)$, which follows from the survey's apparent magnitude limit m_{thr} . This is true for magnitude-limited galaxy surveys, while other surveys may have different and more complex selection functions. The conversion between apparent and absolute magnitude thresholds is

$$M_{\text{thr}}(z) = m_{\text{thr}} - 5 \log_{10} \left(\frac{d_L(z|\lambda_c)}{1 \text{ Mpc}} \right) - 25, \quad (15)$$

so that $M_{\text{thr}}(z)$ also scales as $M_{\text{thr}}(z) + 5 \log_{10} h$. The EM detection probability is then

$$P_{\text{det}}(z) = \frac{0.4 \ln(10) \phi_* h^3 \int_{M_{\min} + 5 \log_{10} h}^{M_{\text{thr}}(z) + 5 \log_{10} h} dM 10^{0.4(\alpha+1)[M_* - (M - 5 \log_{10} h)]} e^{-10^{0.4[M_* - (M - 5 \log_{10} h)]}}}{0.4 \ln(10) \phi_* h^3 \int_{M_{\min} + 5 \log_{10} h}^{M_{\max} + 5 \log_{10} h} dM 10^{0.4(\alpha+1)[M_* - (M - 5 \log_{10} h)]} e^{-10^{0.4[M_* - (M - 5 \log_{10} h)]}}} = \frac{\int_{M_{\min}}^{M_{\text{thr}}(z)} dt 10^{0.4(\alpha+1)(M_* - t)} e^{-10^{0.4(M_* - t)}}}{\int_{M_{\min}}^{M_{\max}} dt 10^{0.4(\alpha+1)(M_* - t)} e^{-10^{0.4(M_* - t)}}}, \quad (17)$$

which is independent of H_0 . Furthermore, we treat the Schechter parameters λ_G as perfectly known in the following, thereby neglecting the dependence of $P_{\text{det}}(z)$ on λ_G . In Fig. 2 we show the dependence of the Schechter function Eq. (13) and of the corresponding galaxy detection probability $P_{\text{det}}(z)$ on some of the Schechter parameters. Note that for more complex or realistic galaxy survey selection functions or more complicated scenarios, the expression for $P_{\text{det}}(z)$ must be adjusted accordingly. For example, for surveys with spatially varying depth or apparent magnitude threshold, P_{det} has to be estimated directly from the data, rather than assumed analytically. This could be achieved by fitting a selection probability to the survey data, or by

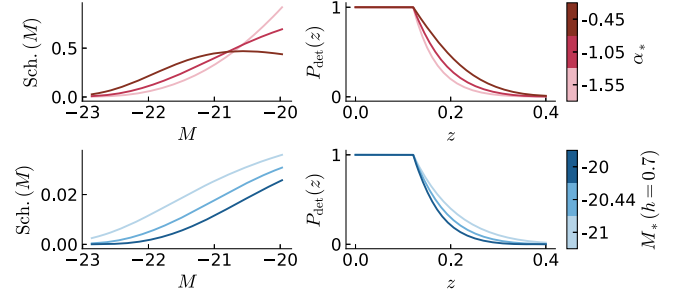


FIG. 2. Dependence of the Schechter function and of the corresponding galaxy detection probability on the Schechter parameters α_* (top panels) and $M_*(h=0.7)$ (bottom panels). The other Schechter parameters are fixed to $\phi_*(h=0.7) = 0.0149$, $M_{\min}(h=0.7) = -23.29$, and $M_{\max}(h=0.7) = -19$.

$$\begin{aligned} P_{\text{det}}(z) &= \int_{M \in \text{detectable}} dM \mathcal{L}_{\text{EM}}(M|\lambda_G) \\ &= \int_{-\infty}^{\infty} dM \Theta(M_{\text{thr}}(z) - M) \mathcal{L}_{\text{EM}}(M|\lambda_G) \\ &= \frac{\int_{M_{\min}}^{M_{\text{thr}}(z)} dM \phi(M|\lambda_G)}{\int_{M_{\min}}^{M_{\max}} dM \phi(M|\lambda_G)}. \end{aligned} \quad (16)$$

Note that $P_{\text{det}}(z)$ reduces to 1 whenever $M_{\text{thr}}(z) < M_{\max}$.

Formally, $P_{\text{det}}(z)$ depends on the cosmological parameters via the luminosity distance in $M_{\text{thr}}(z)$. In the present analysis, we fix all cosmological parameters except H_0 . Under the rescaling of the Schechter parameters, the dependence on h (and thus on H_0) cancels. Indeed, changing the integration variable to $t = M - 5 \log_{10} h$ yields

extending our hierarchical model to include additional selection function parameters.

C. Gaussian process modeling

The third approach to modeling $p_{\text{gal}}(z)$ addresses the main shortcoming of the homogeneous incompleteness correction described above: the loss of galaxy clustering information due to the uniform completion assumption. Here we present the statistical framework aimed at reconstructing the true underlying redshift distribution of galaxies from an incomplete catalog using a nonparametric approach, while retaining the over- and underdensity structure that is crucial for dark siren cosmology.

The likelihood that describes the probability of observing $N_{\text{gal}}^{\text{obs}}$ galaxies, each with redshift posterior distribution $p(z|\tilde{z}_i, \sigma_{z,i})$, given the galaxy redshift distribution p_{gal} is (see Appendix A for a detailed derivation)

$$\mathcal{L}(\{\tilde{z}_i\}|p_{\text{gal}}) = \prod_{i=1}^{N_{\text{gal}}^{\text{obs}}} \frac{\int dz P_{\text{det}}(z) \mathcal{L}(\tilde{z}_i|z, \sigma_{z,i}) p_{\text{gal}}(z)}{\int dz P_{\text{det}}(z) p_{\text{gal}}(z)}, \quad (18)$$

where the single galaxy likelihood $\mathcal{L}(\tilde{z}_i|z, \sigma_{z,i})$ corresponds to the Gaussian distribution appearing in Eq. (6). This likelihood represents an inhomogeneous Poisson process. Each observed galaxy is treated as a draw from the underlying redshift distribution $p_{\text{gal}}(z)$, marginalized over the survey's measurement uncertainty encoded in the single-galaxy likelihood $\mathcal{L}(\tilde{z}_i|z, \sigma_{z,i})$, modeled as in Eq. (7). Each factor in the numerator of the likelihood therefore represents the probability that the i th galaxy, whose true redshift is uncertain, is consistent with being drawn from p_{gal} . The term $\alpha(p_{\text{gal}}) = \int dz P_{\text{det}}(z) p_{\text{gal}}(z)$ is the fraction of the distribution p_{gal} that lies within the detectable region of the survey, and takes into account that not all galaxies are observable through the detection probability $P_{\text{det}}(z)$ defined in Eq. (16). This term thus corrects for EM selection effects. This structure is formally analogous to the likelihood used in GW population inference [52,53], in which one is interested in inferring the parameters Λ of a population model $p_{\text{pop}}(\theta|\Lambda)$ —describing, for instance, the mass and spin distributions of CBCs—given a set of GW detections characterized by source parameters θ . In this framework, the galaxy redshift distribution p_{gal} plays the role of the population model.

We do not use any particular functional form for $p_{\text{gal}}(z)$. Instead, we adopt a flexible nonparametric model, which allows the shape of the redshift distribution to be determined by the data. The main motivation is to have a model flexible enough to reconstruct the redshift distribution of the galaxies keeping over- and underdensity information, which are washed out in the “homogeneous completion method” but are crucial for dark siren cosmology, while being simple and not dependent on too many astrophysical assumptions and uncertainties. The framework we propose is based on GPs. A GP is a stochastic process, characterized by a mean function $m(x)$ and a covariance function (or kernel) $K(x, x')$, that defines a probability distribution over functions

$$g(x) \sim \mathcal{GP}(m(x), K(x, x')). \quad (19)$$

The defining property of a GP is that for any finite collection of input points $\mathbf{x} = \{x_1, \dots, x_n\}$, the marginal distribution of the corresponding function values $g(\mathbf{x}) = \{g(x_1), \dots, g(x_n)\}$ is a multivariate Gaussian $g(\mathbf{x}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where the mean vector $\boldsymbol{\mu}$ and the covariance

matrix $\boldsymbol{\Sigma}$ are defined by the GP mean and covariance functions as $\mu_i = m(x_i)$ and $\Sigma_{ij} = K(x_i, x_j)$, respectively. Typically, GPs are used in regression problems. However, here we utilize GP as a prior distribution on the space of functions describing the galaxy distribution $p_{\text{gal}}(z)$. We choose a mean function and a covariance function, described by some parameters λ_{GP} , and we build the GP. We then draw the values of $p_{\text{gal}}(z)$ on a defined redshift grid $\mathbf{z} = \{z_1, \dots, z_n\}$ from such a GP. The interpolation of the values $p_{\text{gal}}(\mathbf{z})$ gives a smooth function modeling the galaxy rate entering in the likelihood Eq. (18).

Rather than fixing λ_{GP} to a single value, we effectively marginalize over them. To do this, we construct a likelihood for the GP hyperparameters given the observed galaxy catalog and combine it, via Bayes's theorem, with a prior on λ_{GP} to obtain a posterior distribution conditioned on the data. Marginalizing over this posterior propagates the uncertainty on λ_{GP} directly into the distribution of reconstructed GP realizations $p_{\text{gal}}(z)$. This choice encodes very little prior information about the shape of the galaxy redshift distribution, besides enforcing that it must be smooth.

In the following analyses, we consider a zero-mean function. The kernel adopted, instead, is the Matern one with $\nu = 5/2$:

$$K(x_i, x_j) = \sigma^2 \left(1 + \frac{\sqrt{5}r}{\rho} + \frac{5r^2}{3\rho^2} \right) \exp \left(-\frac{\sqrt{5}r}{\rho} \right), \quad (20)$$

where $r = \sqrt{(x_i - x_j)^T (x_i - x_j)}$. This kernel was chosen because we checked that it yields results comparable to those obtained with the exponential kernel, which is the standard choice in many problems, while being significantly faster. This kernel depends on two parameters, σ and ρ . The first controls the overall amplitude scale of the covariance, while the latter describes the correlation scale, determining over what input distance the function tends to change significantly. To see the effects of these kernel parameters and function values drawn from the GP see Fig. 3.

As shown in Fig. 3, a function randomly drawn from a GP can take negative values. Since the galaxy distribution entering in the likelihood Eq. (18) must be positive definite, we define p_{gal} as a transformed realization of a GP, using the following transformation:

$$p_{\text{gal}}(z) \propto \mathcal{T}(g(z)) = l(z) + (u(z) - l(z)) \cdot \frac{1}{1 + e^{-g(z)}}. \quad (21)$$

The sigmoid function ensures the GP realization is bounded within (0, 1). This bounded output is then scaled to lie between two redshift-dependent boundaries, which scale as the comoving volume element, defined as

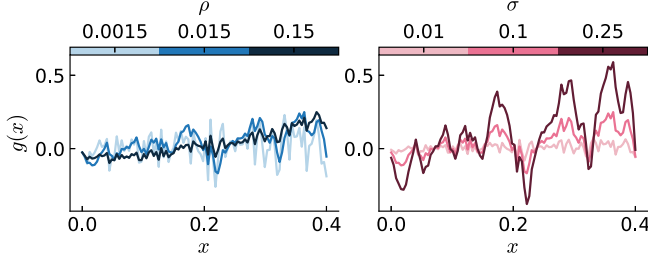


FIG. 3. Effects of the parameters ρ and σ defining the Matern 5/2 kernel on functions drawn from the corresponding GP. The different curves correspond to different random realizations drawn from the GP prior.

$$\begin{aligned} u(z) &= (1+a) \frac{dV_c/dz}{\int dz dV_c/dz}, \\ l(z) &= (1-a) \frac{dV_c/dz}{\int dz dV_c/dz}. \end{aligned} \quad (22)$$

Here, the factor a controls how much the GP realization can spread between the upper and lower boundaries. The factor a is therefore crucial to avoid GP overfitting. In general, it depends on the number of galaxies and on the redshift errors: we expect fewer fluctuations with many galaxies and large redshift errors, since the sum of many broad galaxy redshift posteriors tends to be smoother. Consequently, the value of a in a real survey must be calibrated on simulations, but it is expected to be inversely proportional to the volume of the event considered (a larger volume implies more galaxies, hence a smaller a to avoid overfitting) and to σ_z for a similar reason. We select boundaries that follow the comoving volume element because the underlying galaxy distribution in the following analyses is expected to be approximately uniform in comoving volume. This scaling is crucial to speed up the Markov chain Monte Carlo (MCMC) sampling of the likelihood Eq. (18), as described in the following sections. Note that the comoving volume elements in both boundary functions are normalized, so they no longer depend on H_0 and do not artificially introduce information about it. The effects of such transformations on random GP realizations are shown in Fig. 4.

III. RESULTS

We now apply the Bayesian framework and models described in the previous sections to simulated datasets. First, we describe the construction of the mock galaxy catalog, as well as the corresponding simulated GW catalogs. We then present the results of the GP reconstruction of the galaxy redshift distribution. Finally, we demonstrate an application to dark siren cosmology.

A. Mock catalogs

Mock galaxy catalogs. The mock galaxy catalogs considered in this work are obtained from a subsample of the

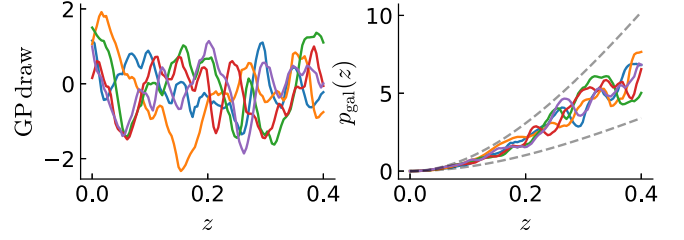


FIG. 4. Effect of the transformation Eq. (21) on random GP realizations. Left: random GP realizations. Right: random GP realizations transformed using Eq. (21), compared to the redshift-dependent boundaries Eq. (22) used in the transformation (dashed lines).

MICE grand challenge light-cone simulation (v1) [56–59]. In fact, since the intrinsic magnitude distribution of the MICE catalog does not match a Schechter function Eq. (13), we obtained the subsample via rejection sampling: starting from the full MICE catalog restricted to $z \leq 0.4$, galaxies were accepted or rejected with a probability chosen such that the resulting magnitude distribution follows the Schechter function with fiducial parameters

$$\begin{aligned} \alpha &= -1.05, & M_* &= -21.21, & \phi_* &= 0.0434 \text{ Mpc}^{-3}, \\ M_{\min} &= -24.06, & M_{\max} &= -19.77. \end{aligned} \quad (23)$$

The resulting subsample contains approximately 3.65×10^6 galaxies. Figure 5 shows the redshift and absolute magnitude distributions of the subsample, comparing them to the theoretical distributions. The redshift cut at $z < 0.4$ results in a distribution that is uniform in comoving volume for the fiducial MICEv1 cosmology, which assumes a flat Λ CDM model with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $\Omega_{m,0} = 0.25$. This redshift cut was chosen precisely to ensure a background galaxy distribution that is approximately uniform in comoving volume, simplifying the modeling of the underlying galaxy population.

For dark siren cosmology, our goal is to reconstruct the galaxy distribution along a specific LOS corresponding to a GW event’s localization area. We therefore model each LOS individually. We consider 1000 different LOS drawn

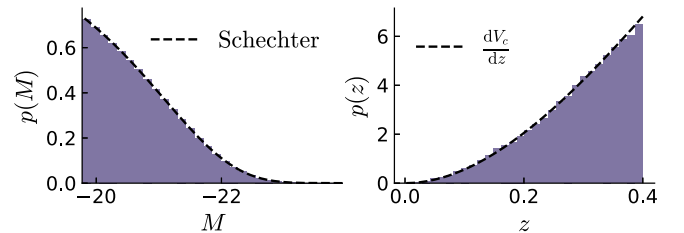


FIG. 5. Redshift and absolute magnitude distributions for the MICEv1 mock galaxy catalog subsample (histograms). The black dashed lines show the underlying theoretical redshift distribution (uniform in comoving volume) and the Schechter luminosity function from which the catalog was generated.

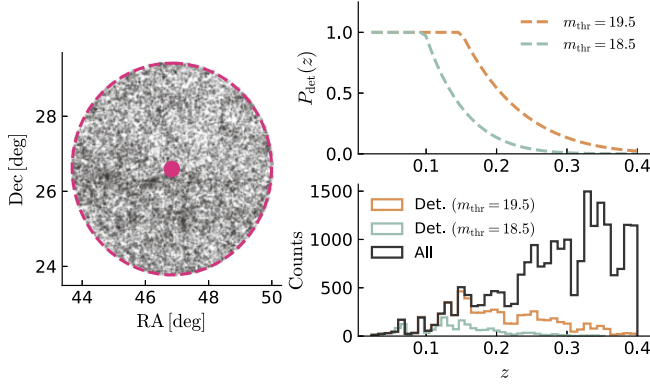


FIG. 6. Construction of a single line-of-sight galaxy sample. Left: schematic of the 25 deg^2 sky region used for each mock gravitational-wave event. Top right: the galaxy detection probability $P_{\text{det}}(z)$ as a function of redshift, resulting from the apparent magnitude cuts $m_{\text{thr}} = 19.5$ (orange) and $m_{\text{thr}} = 18.5$ (green). Bottom right: redshift histogram of the detected galaxies, shown in orange and green for $m_{\text{thr}} = 19.5$ and $m_{\text{thr}} = 18.5$, respectively, compared to the true underlying galaxy distribution in the line-of-sight volume (gray).

from the original MICE subsample, one for each mock GW event considered in the following analyses. Specifically, for each mock GW event, the LOS dataset contains all galaxies within a 25 deg^2 round sky patch centered on the LOS (see left panel of Fig. 6). This area is comparable to the 90% localization region of the best-localized BBH event from the O4a observing run [60,61] and represents the maximum area of the 4% best-localized events in O4b [29,62]. It is also representative of the pixel area used when pixelating large localization areas, $500\text{--}1000 \text{ deg}^2$, that is the area of roughly 20% of events in the GWTC-5 cosmology analysis [29].

To simulate realistic survey selection effects, we impose a detection threshold on the apparent magnitude Eq. (15), removing galaxies with $m > m_{\text{thr}}$. We study two cases, $m_{\text{thr}} = 19.5, 18.5$, corresponding to mean completeness fractions (averaged along all LOS) of 24% and 8%:

$$\begin{aligned} m_{\text{thr}} = 19.5 & \rightarrow \langle P_{\text{compl}} \rangle = 24\%, \\ m_{\text{thr}} = 18.5 & \rightarrow \langle P_{\text{compl}} \rangle = 8\%. \end{aligned} \quad (24)$$

The corresponding detection probabilities $P_{\text{det}}(z)$, defined in Eq. (16), are shown in the top right panel of Fig. 6. The bottom right panel compares the redshift distribution of the detected galaxies to the true underlying distribution in the considered LOS, illustrating the impact of the magnitude cuts.

For the galaxy redshift likelihood in Eq. (7), we investigate five scenarios for the redshift uncertainty $\sigma_{z,i}$:

- (i) spec-z (Spectroscopic errors): $\frac{\sigma_{z,i}}{1+z_i} = 1 \times 10^{-3}$.
- (ii) Intermediate case 1: $\frac{\sigma_{z,i}}{1+z_i} = 4 \times 10^{-3}$.
- (iii) Intermediate case 2: $\frac{\sigma_{z,i}}{1+z_i} = 7 \times 10^{-3}$.

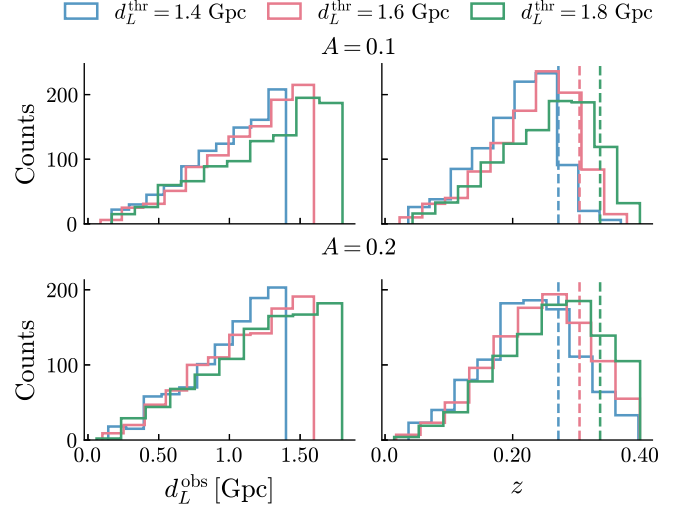


FIG. 7. Observed luminosity distance and redshift distributions of the 6 mock GW catalogs generated from 1000 LOS drawn from the MICEv1 subsample.

(iv) Intermediate case 3: $\frac{\sigma_{z,i}}{1+z_i} = 1 \times 10^{-2}$.

(v) DES (DES survey photometric errors): $\sigma_{z,i} = \min[0.013 \cdot (1+z_i)^3, 0.015]$,

where the last case represents photometric redshift uncertainties expected from the Dark Energy Survey [DES [63]], as done in [64].

Note that the observed redshift $\tilde{z}_i$ [the mean μ in Eq. (7)] is not the true MICEv1 redshift z_i , but is instead drawn from a Gaussian distribution centered at the redshift and with standard deviation $\sigma_{z,i}$, effectively perturbing the true value within its measurement uncertainty. Consequently, larger redshift errors produce a smoother observed distribution.

Mock GW catalogs. To generate GW catalogs, we assign a binary merger event to each galaxy located in the innermost region of each of the 1000 LOS considered. We associate events with all galaxies, not only those that are detected, and assign to each GW event the true redshift of its host galaxy. We then compute the true luminosity distance of each event assuming the fiducial cosmological model used to build the MICEv1 simulation. To obtain the observed GW luminosity distance, d_L^{obs} , we perturb the true luminosity distance by drawing a random value from the Gaussian distribution in Eq. (2), centered on the true distance, for a given choice of A . We consider two configurations defined by $A = 0.1$ and $A = 0.2$, corresponding to 10% and 20% errors on the observed luminosity distance, respectively. We determine which GW events are detected by applying the criterion $d_L^{\text{obs}} < d_L^{\text{thr}}$, where d_L^{thr} is the luminosity distance threshold entering the GW selection function in Eq. (4). We consider three luminosity distance thresholds: $d_L^{\text{thr}} = 1.4, 1.6$, and 1.8 Gpc . For each LOS, among the events satisfying this criterion, we then randomly select one to represent the single GW detection associated with that LOS. This

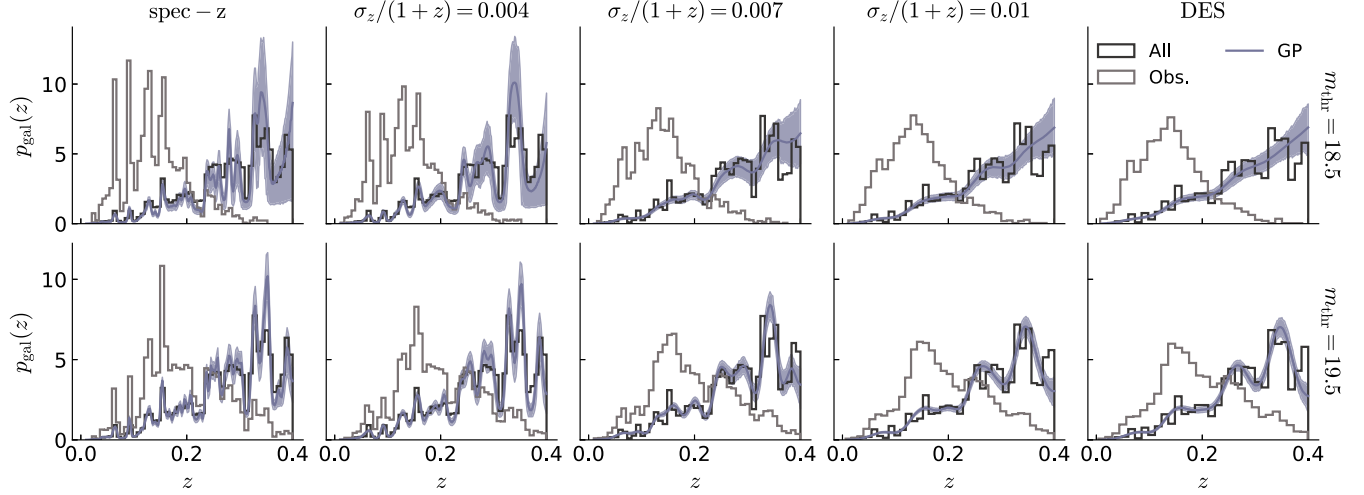


FIG. 8. Posterior distribution for $p_{\text{gal}}(z)$ obtained with GP model (purple curve) and, for comparison, the true galaxy distribution (black normalized histogram) and observed galaxy distribution (gray normalized histograms). The panels in the first row show results obtained with the catalog cut at $m > m_{\text{thr}} = 18.5$, while the second row shows results for $m_{\text{thr}} = 19.5$. The results are those relative to the single LOS plotted in Fig. 6.

procedure yields six mock GW catalogs of 1000 events each, consistent with the order of magnitude expected from the LIGO-Virgo-KAGRA O5 observing run [65]. In Fig. 7, we show the properties of the six catalogs, including their observed luminosity distance distributions and redshift distributions.

B. Galaxy redshift distribution reconstruction

We implement the likelihood Eq. (18) in JAX [66], and the construction of the GP is done using the TINYGP package [67]. This allows us to sample the likelihood Eq. (18) on GPUs using the No-U-Turn [NUTS, [68]] Hamiltonian Monte Carlo sampler implemented in the NUMPYRO framework [69]. Thanks to GPU acceleration, the fit for each LOS and redshift error assumption takes $\mathcal{O}(10)$ minutes on a single A100 GPU.

In Appendix B, we present and discuss the posterior distribution for the GP kernel parameters obtained with the LOS shown in Fig. 6. In Fig. 8, instead, we show the posterior distribution for $p_{\text{gal}}(z)$ obtained with the GP method. In particular, we plot the median and the 68% credible interval (C.I.). Top panels show results for the $m_{\text{thr}} = 19.5$ catalog, while lower ones for the $m_{\text{thr}} = 18.5$ catalog. The reconstructed $p_{\text{gal}}(z)$ follows quite precisely the true underlying galaxy distribution, especially in the $m_{\text{thr}} = 19.5$ case. In the other case, the GP reconstructs fewer features and with larger uncertainty at high redshift, where only a few galaxies are observed. Nevertheless, some high-redshift features are still reconstructed. Moreover, the larger the redshift error, the fewer features in the reconstructed redshift distribution are recovered. The way in which the GP model adds missing galaxies is by adding them where overdensities in the observed redshift distribution are present, effectively adding more galaxies where redshift clusters are

present, rather than adding them uniformly in comoving volume as done in the homogeneous completion method. This behavior is directly tied to the assumed galaxy redshift errors: larger error smooths out the observed overdensities, making them harder for the GP to distinguish from noise, and thus limiting its ability to reconstruct accurately the underlying clustering structure. In fact, as it was noted above, larger redshift errors result in fewer recovered features in the reconstructed distribution.

In Fig. 9, we compare the redshift posterior distributions obtained along the same LOS using three different methods: the GP-completed catalog (purple), the standard homogeneous completion in uniform comoving volume (UCV, green), and the complete galaxy catalog (COMP., pink). The first row shows results for the $m_{\text{thr}} = 18.5$ catalog, while the second row shows results for $m_{\text{thr}} = 19.5$. Overall, the GP-based posteriors are in better agreement with those obtained using the complete catalog than the UCV ones, as the GP method captures the clustering structure of the observed galaxy distribution rather than distributing the missing galaxies uniformly in comoving volume. In particular, the GP posteriors tend to peak at redshifts where overdensities are present in the observed catalog, closely following the features recovered by the complete catalog. The UCV method, by contrast, produces smoother and more featureless posteriors, washing out the redshift clustering information. The improvement brought by the GP completion is more pronounced for the $m_{\text{thr}} = 19.5$ catalog, where the denser observed sample allows the GP to reconstruct the underlying galaxy distribution more accurately. For the shallower $m_{\text{thr}} = 18.5$ catalog, the GP posteriors still outperform the UCV ones, although the agreement with the COMP. results is somewhat reduced at high redshift, consistently with the larger reconstruction

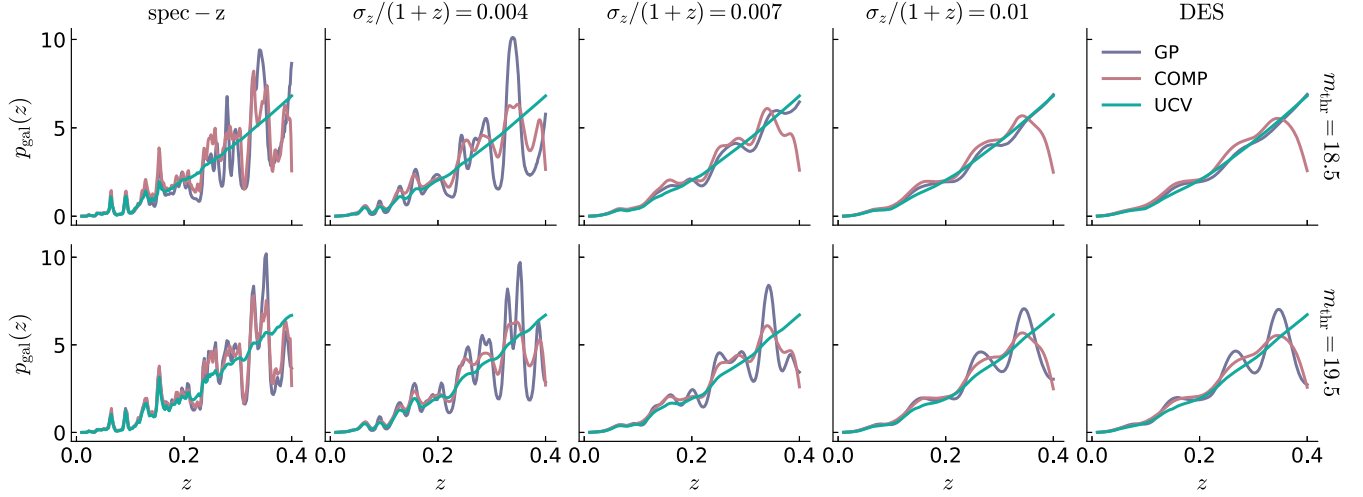


FIG. 9. Redshift posterior distributions along the LOS plotted in Fig. 6 using the catalog with $m_{\text{thr}} = 18.5$ (first row) and the one with $m_{\text{thr}} = 19.5$ (second row). In each plot, the purple curves correspond to results obtained using the GP-completed galaxy catalog (in particular we show only the median of the GP-reconstructed $p_{\text{gal}}(z)$, the green curves to results obtained using the standard homogeneous completion (UCV), and the pink curves to results obtained considering the complete galaxy catalog (COMP.).

uncertainty seen in Fig. 8. We note that the median of the reconstructed $p(z)$ can exhibit larger fluctuations than the COMP case, but from Fig. 8 we note that the COMP distribution is enclosed within the full reconstructed posterior.

C. Hubble constant constraints

Having generated the mock GW catalogs, we can now use the likelihood Eq. (1) to estimate H_0 . We model the merger rate distribution p_{rate} entering in Eq. (5) as a uniform distribution in the redshift range $[0, 0.4]$ where the MICEv1 subsample considered has support. We study three different scenarios described previously to model the p_{gal} distribution. In the first one the galaxy catalog is 100% complete and p_{gal} , in each LOS, is given by Eq. (6), where the observed galaxy redshifts $\tilde{z}_i$ are obtained by perturbing all true ones using the Gaussian likelihood for galaxy redshifts Eq. (7). In the second case, we consider the incomplete galaxy catalog obtained with the apparent magnitude cut Eq. (24) and p_{gal} is built using the homogeneous completion method Eq. (11). Finally, in the third case we use as p_{gal} the median of the predictive posterior redshift distribution reconstructed from the incomplete catalog using the Gaussian process method. This choice represents an approximation. In principle, one should marginalize over all possible realizations of p_{gal} drawn from each GP. Even more consistently, one should perform a joint inference of the redshift distribution and the H_0 posterior using both GW and galaxy data simultaneously. However, this would require accounting for all events across the ~ 1000 LOS, implying the construction and sampling of ~ 1000 distinct GP models, which makes the procedure computationally prohibitive for the present analysis.

We study how the H_0 posterior changes as a function of the parameters governing both the galaxy observations and the GW likelihood. Specifically, we vary the apparent magnitude threshold m_{thr} and the redshift uncertainty $\sigma_{z,i}$ to assess the impact of galaxy catalog quality, and the fractional distance error A and luminosity distance threshold d_L^{thr} to probe the sensitivity to GW measurement precision and detector horizon. A complete table reporting the median and 68% C.I. for each configuration studied, is provided in Appendix C.

1. Impact of galaxy survey assumptions

We first study how the galaxy survey assumptions affect the H_0 posterior. In the top two panels of Fig. 10 we show the H_0 68% C.I. for various redshift errors and for the two tested apparent magnitude thresholds, $m_{\text{thr}} = 18.5$ and 19.5. These results are obtained using the mock GW catalog simulated assuming $d_L^{\text{thr}} = 1.6$ Gpc and $A = 0.1$, containing 1000 GW events each from a different LOS.

We note that when the redshift error is sufficiently small ($\sigma_z/(1+z) \leq 0.007$), the posterior obtained with the GP method closely matches that from the complete galaxy catalog, while homogeneous completion yields a less informative posterior. This reflects the fact that the GP faithfully reconstructs the true underlying galaxy distribution and preserves features at high redshift, as in the complete case, whereas these structures are washed out by homogeneous completion. When the redshift error becomes larger ($\sigma_z/(1+z) > 0.007$), the GP posterior becomes marginally more informative than even that of the complete galaxy catalog. This modest effect, not representing a substantial improvement, occurs because the GP method is able to recover some of the true redshift

density fluctuations even in the presence of large redshift uncertainties (see Fig. 9), while in the complete case we effectively marginalize over all galaxies along the LOS. Since there are many galaxies along the LOS, each carrying a large redshift uncertainty, the resulting redshift distribution is smoother than the GP one. Physically, the GP-reconstructed distribution resembles the one obtainable by retaining only clusters of galaxies in redshift space (rather than angular clusters). This may thus introduce a small bias in regimes where the cluster distribution does not trace well the true host distribution, although we do not find evidence of this in the present analysis.

2. Impact of GW observation assumptions

We now turn to the impact of GW observation assumptions. From the last two panels of Fig. 10 one can read off the dependence on d_L^{thr} and A at fixed galaxy survey configuration ($m_{\text{thr}} = 18.5$ and spec-z redshift errors).

For the GP and complete-catalog methods, the constraints remain broadly comparable when increasing d_L^{thr} from 1.4 to 1.8 Gpc, with only a mild degradation in some configurations. By contrast, the homogeneous completion method becomes significantly less informative at large d_L^{thr} . This is because increasing the GW horizon probes higher-redshift regions where the galaxy catalog is much more incomplete and the homogeneous completion increasingly dominates the reconstructed redshift prior, washing out the clustering information relevant for dark siren cosmology. The GP method instead remains able to partially reconstruct the underlying overdensity structure even in these sparsely sampled regions, making it substantially more precise despite the large incompleteness at high redshift. As expected, the $A = 0.1$ configurations systematically yield tighter constraints than the $A = 0.2$ ones, since smaller fractional luminosity-distance uncertainties translate directly into sharper GW likelihoods and therefore more informative H_0 posteriors.

3. Summary of results

On average across all simulations tested, the GP method provides constraints that are 23% more precise than those obtained with the homogeneous completion method when using the $m_{\text{thr}} = 19.5$ galaxy catalog (see Table I). The largest improvement, 52%, is obtained in the DES case with $A = 0.1$ and $d_L^{\text{thr}} = 1.8$ Gpc. When using the 8% complete galaxy catalog ($m_{\text{thr}} = 18.5$) instead, the improvements are much larger, averaging 37%. The maximum improvement, 66%, is obtained in the $\sigma_z/(1+z) = 0.007$ case with $A = 0.1$ and $d_L^{\text{thr}} = 1.8$ Gpc. This is expected: with a shallower catalog the homogeneous completion reconstructs redshift features much less accurately, especially at high redshift where observed galaxies are sparse, whereas the GP completion is still able to recover some of the high-redshift structure and thus provide tighter constraints on H_0 .

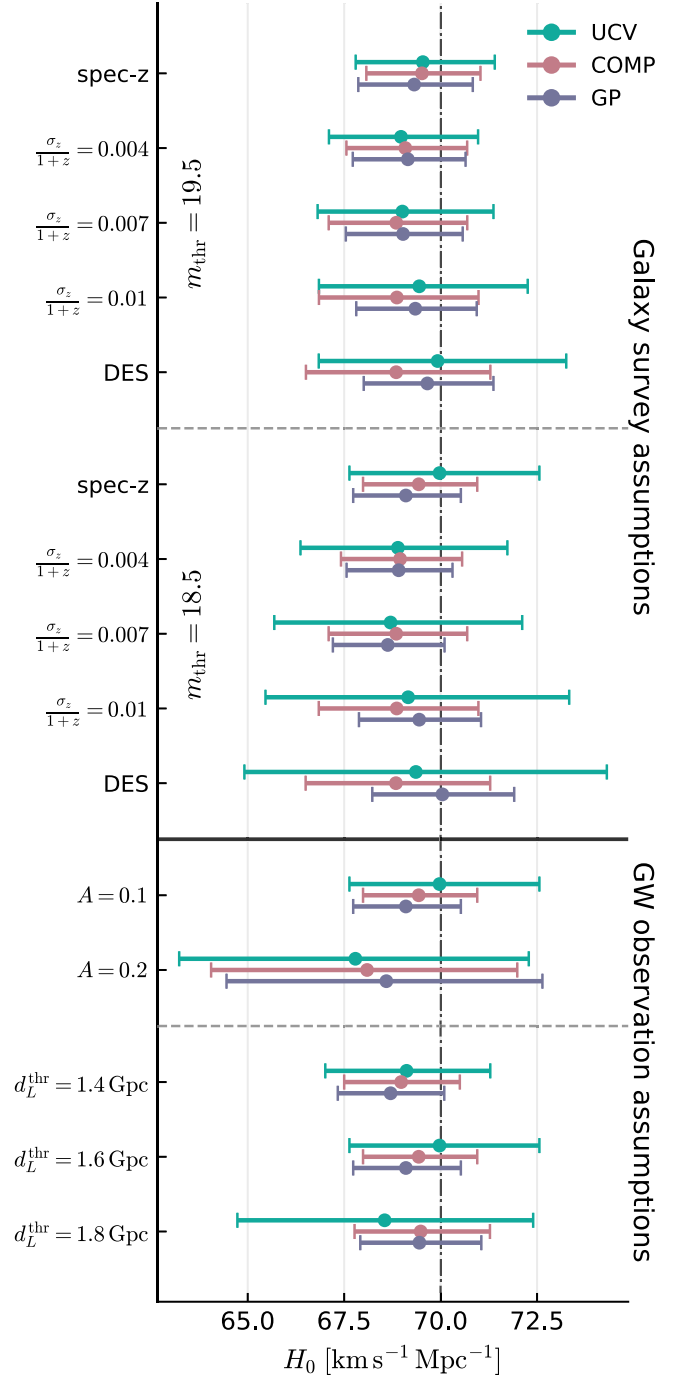


FIG. 10. 68% C.I. of the H_0 posterior for various configurations, all using the mock GW catalog with $A = 0.1$ and $d_L^{\text{thr}} = 1.6$ Gpc unless otherwise stated. The top two panels show results for different redshift errors and magnitude thresholds ($m_{\text{thr}} = 19.5$ and 18.5 , respectively). The bottom two panels fix galaxy assumptions to $m_{\text{thr}} = 18.5$ and spec-z errors, varying A (third panel) and d_L^{thr} (fourth panel). The vertical dashed line marks the injected H_0 value.

Figure 11 illustrates the two configurations that yield the largest GP improvement over UCV: the $\sigma_z/(1+z) = 0.007$, $m_{\text{thr}} = 18.5$ case (left panel) and the DES,

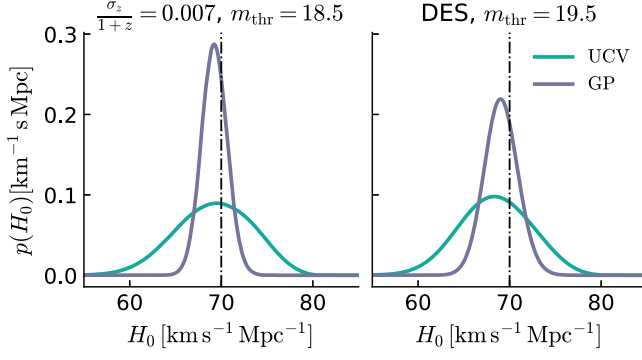


FIG. 11. H_0 posterior distributions for the two configurations exhibiting the largest GP improvement over homogeneous completion: $\sigma_z/(1+z) = 0.007$ with $m_{\text{thr}} = 18.5$ (left) and DES-like redshift errors with $m_{\text{thr}} = 19.5$ (right), both at $A = 0.1$ and $d_L^{\text{thr}} = 1.8$ Gpc.

$m_{\text{thr}} = 19.5$ case (right panel), both at $A = 0.1$ and $d_L^{\text{thr}} = 1.8$ Gpc. In both cases, the largest improvement is achieved at the maximum luminosity distance threshold considered, $d_L^{\text{thr}} = 1.8$ Gpc, where a relevant fraction of the observed GW events lie at redshifts where the galaxy catalog is most incomplete. For these events, homogeneous completion assumes a smooth redshift prior, neglecting the clustering features that are instead partially recovered by the GP method, which consequently yields markedly tighter constraints on H_0 .

IV. CONCLUSIONS

We have presented a novel method for improving H_0 inference from GW dark sirens by applying Gaussian processes directly to the reconstruction of the line-of-sight galaxy redshift distribution from incomplete catalogs. Our approach addresses a fundamental limitation of the standard homogeneous incompleteness correction, which assumes that galaxies missing from a catalog are uniformly distributed in comoving volume. While this assumption correctly accounts for the overall normalization of the galaxy number density, it washes out the over- and under-density structure of the galaxy distribution, that is the information that drives the constraining power in the dark siren method. In contrast, the GP framework reconstructs the true underlying distribution nonparametrically, preserving redshift clustering features and learning the characteristic correlation scale of the galaxy distribution directly from the data, without requiring any explicit galaxy bias or halo occupation model.

We validated the method on mock datasets derived, considering a range of galaxy survey completeness levels, redshift uncertainty assumptions, and GW detector configurations. In all cases, the GP method yields H_0 constraints consistent with the true fiducial value and systematically more precise than those from homogeneous completion, with improvements averaging 23% and 37%

for galaxy catalogs with mean completeness of 24% and 8%, respectively. The GP reconstruction is particularly valuable at large luminosity distance thresholds, where the catalog is most incomplete and the homogeneous completion is least informative.

Several directions for future work naturally follow from this analysis. First, the GP framework could be extended to incorporate galaxy weighting, assigning each galaxy a weight proportional to its expected merger rate—for example based on luminosity or stellar mass—as done in [19]. Second, the present analysis approximates the full Bayesian inference by using only the median of the GP-reconstructed redshift distribution when computing H_0 posteriors. A fully Bayesian treatment, in which the redshift distribution along each LOS and H_0 are inferred jointly, marginalizing over all GP realizations simultaneously, remains to be developed. Since this is prohibitive from the computational point of view, dedicated strategies have to be developed. Finally, the method should be applied to real GW and galaxy catalog data, to assess its performance on real observations. In this context, it will be crucial to include the mixing of spectroscopic and photometric redshifts, so that the GP can exploit the precise low-redshift measurements to better reconstruct the LOS redshift prior at high redshift, where photometric uncertainties dominate, ultimately providing improved constraints on H_0 .

ACKNOWLEDGMENTS

M. T. thanks the Max Planck Institute for Gravitational Physics—Albert Einstein Institute (Potsdam) for the hospitality while this work has been developed. M. M. acknowledges the financial contribution from the Grant No. PRIN-MUR 2022 2022NY2ZRS 001 “Optimizing the extraction of cosmological information from Large Scale Structure analysis in view of the next large spectroscopic surveys” and from the Grant No. ASI 2024-10-HH.0 “Attività scientifiche per la missione Euclid-fase E.” J. G. acknowledges support from the Simons Foundation International via Grant No. SFI-MPS- BH-00012593-06. We also acknowledge ISCRA for awarding the projects GPGW (No. HP10CF55AE, PI: M. T.) and MLGW (No. HP10CM1UJX, PI: M. T.) access to the LEONARDO supercomputer, owned by the EuroHPC Joint Undertaking, hosted by CINECA (Italy).

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

APPENDIX A: GALAXY LIKELIHOOD DERIVATION

In this section, we derive the likelihood Eq. (18) used to reconstruct the true underlying redshift distribution of

galaxies from an incomplete catalog using a nonparametric approach. The derivation follows the same hierarchical approach used for gravitational-wave population inference [15,52,53,70]. We are interested in the probability of observing $N_{\text{gal}}^{\text{obs}}$ galaxies with redshifts $\{z_i\}$, given some function $\mu(z) \equiv \frac{dN_{\text{gal}}}{dz}$ describing the true underlying galaxy population rate. Such a function can be parametrized by parameters Λ or approximated using nonparametric approaches. To be as general as possible, we write a likelihood for μ rather than for its parameters Λ .

We start by discretizing the redshift axis into fine bins of width δz_k , chosen such that each bin contains at most one detected galaxy. In each bin centered at z_k , the number of galaxies is modeled as a Poisson random variable with rate $\mu(z_k)\delta z_k$. The probability of observing ν galaxies in the k th bin is

$$p(\nu|\mu(z_k)) = e^{-\delta z_k \mu(z_k)} \frac{(\delta z_k \mu(z_k))^\nu}{\nu!}. \quad (\text{A1})$$

The full likelihood for the observed set $\{z_i\}$ is the product of two terms: the probability of detecting exactly one galaxy in each of the $N_{\text{gal}}^{\text{obs}}$ bins containing only one detected galaxy (p_1), and the probability of detecting zero galaxies in all other bins ($\bar{p}$):

$$\mathcal{L}(\{z_i\}|\mu) = \prod_{i=1}^{N_{\text{gal}}^{\text{obs}}} p_1(\delta z_i) \prod_{k \neq i} \bar{p}(\delta z_k). \quad (\text{A2})$$

a. Zero-detection bins. Let $P_{\text{det}}(z)$ denote the probability of detecting a galaxy at redshift z . For a bin with center z_k , the probability of zero detections is obtained by marginalizing over the true (unknown) number of galaxies ν in that bin—given by Eq. (A1), each weighted by the probability that not one of them is detected

$$\begin{aligned} \bar{p}(\delta z_k) &= \sum_{\nu=0}^{\infty} e^{-\delta z_k \mu(z_k)} \frac{(\delta z_k \mu(z_k))^\nu}{\nu!} (1 - P_{\text{det}}(z_k))^\nu \\ &= e^{-\delta z_k \mu(z_k)} \sum_{\nu=0}^{\infty} \frac{1}{\nu!} [\delta z_k \mu(z_k) (1 - P_{\text{det}}(z_k))]^\nu \\ &= e^{-\delta z_k \mu(z_k) P_{\text{det}}(z_k)}. \end{aligned} \quad (\text{A3})$$

The product over all empty bins—which is the second term in Eq. (A2)—therefore becomes

$$\prod_{k \neq i} \bar{p}(\delta z_k) = e^{-\sum_{k \neq i} \delta z_k \mu(z_k) P_{\text{det}}(z_k)}. \quad (\text{A4})$$

b. Single-detection bins. For a bin centered at z_i containing one detected galaxy, we must marginalize over the true number of galaxies $\nu \geq 1$, with exactly one of them

detected and the remaining $\nu - 1$ missed. The number of ways to choose which of the ν galaxies is the detected one is ν , and the probability of that specific configuration (one detected, $\nu - 1$ missed) is $P_{\text{det}}(z_i)(1 - P_{\text{det}}(z_i))^{\nu-1}$:

$$\begin{aligned} p_1(\delta z_i) &= \sum_{\nu=1}^{\infty} \nu P_{\text{det}}(z_i) e^{-\delta z_i \mu(z_i)} \frac{(\delta z_i \mu(z_i))^\nu}{\nu!} (1 - P_{\text{det}}(z_i))^{\nu-1} \\ &= P_{\text{det}}(z_i) e^{-\delta z_i \mu(z_i)} \delta z_i \mu(z_i) \\ &\quad \times \sum_{\nu=1}^{\infty} \frac{[\delta z_i \mu(z_i) (1 - P_{\text{det}}(z_i))]^{\nu-1}}{(\nu-1)!} \\ &= \delta z_i \mu(z_i) P_{\text{det}}(z_i) e^{-\delta z_i \mu(z_i) P_{\text{det}}(z_i)}, \end{aligned} \quad (\text{A5})$$

where the additional factor ν is due to the fact that there are ν ways to choose which galaxy is detected, and the factor $P_{\text{det}}(z_i)$ is the probability that the chosen galaxy is indeed the one detected. The product over all detected galaxies—the first term in Eq. (A2)—is then

$$\begin{aligned} \prod_{i=1}^{N_{\text{gal}}^{\text{obs}}} p_1(\delta z_i) &= (\delta z_i)^{N_{\text{gal}}^{\text{obs}}} e^{-\sum_{i=1}^{N_{\text{gal}}^{\text{obs}}} \delta z_i \mu(z_i) P_{\text{det}}(z_i)} \\ &\quad \times \prod_{i=1}^{N_{\text{gal}}^{\text{obs}}} \mu(z_i) P_{\text{det}}(z_i). \end{aligned} \quad (\text{A6})$$

c. Continuum limit. Substituting Eqs. (A6) and (A4) into Eq. (A2) and taking the continuum limit $\delta z \rightarrow 0$ (so that sums become integrals) yields

$$\mathcal{L}(\{z_i\}|\mu) = e^{-\int dz \mu(z) P_{\text{det}}(z)} \prod_{i=1}^{N_{\text{gal}}^{\text{obs}}} \mu(z_i) P_{\text{det}}(z_i), \quad (\text{A7})$$

where we absorbed the constant factor $(\delta z_i)^{N_{\text{gal}}^{\text{obs}}}$ into an overall normalization. Note that Eq. (A7) is the standard result for a thinned Poisson process: if the true process has intensity $\mu(z)$ and each point is independently retained with probability $P_{\text{det}}(z)$, the observed process has intensity $\mu(z)P_{\text{det}}(z)$, both for the exponential (survival) term and for the density of observed points.

d. Including galaxy redshift errors. If the galaxy catalog provides not exact redshifts but only noisy measurements $\tilde{z}_i$ for each galaxy (with measurement likelihood $\mathcal{L}(\tilde{z}_i|z, \sigma_{z,i})$), the likelihood must be marginalized over the true unknown redshift z of each detected galaxy. Since $P_{\text{det}}(z)$ is a property of the true redshift z (through the true apparent magnitude of the galaxy) and not of the noisy measurement, it must remain inside this marginalization together with $\mu(z)$:

$$\mathcal{L}(\{\tilde{z}_i\}|\mu) = e^{-\int dz \mu(z) P_{\text{det}}(z)} \times \prod_{i=1}^{N_{\text{gal}}^{\text{obs}}} \int dz \mu(z) P_{\text{det}}(z) \mathcal{L}(\tilde{z}_i|z, \sigma_{z,i}), \quad (\text{A8})$$

where $\mathcal{L}(\tilde{z}_i|z, \sigma_{z,i})$ is modeled as in Eq. (7), now interpreted as the measurement likelihood rather than a posterior on z . We stress that this assumes detection depends only on the true redshift z and not on the redshift measurement error, i.e., that whether a galaxy is detected is unaffected by the (photometric or spectroscopic) redshift uncertainty $\sigma_{z,i}$; this is a reasonable assumption since detection is set by

apparent magnitude, which is determined by the true redshift.

e. Scale-free likelihood. We are interested in inferring the galaxy distribution p_{gal} , rather than the full galaxy rate μ . To write the likelihood for p_{gal} instead of μ , we factorize the latter as

$$\mu(z) \equiv \frac{dN_{\text{gal}}}{dz} = N p_{\text{gal}}(z). \quad (\text{A9})$$

Then, we marginalize Eq. (A8) over N assuming a scale-free prior, that is $\pi(\log(N)) = \text{Unif}$:

$$\begin{aligned} \mathcal{L}(\{\tilde{z}_i\}|p_{\text{gal}}) &= \int_0^\infty \frac{dN}{N} \pi(N) e^{-N \int dz p_{\text{gal}}(z) P_{\text{det}}(z)} \prod_{i=1}^{N_{\text{gal}}^{\text{obs}}} \int dz N p_{\text{gal}}(z) P_{\text{det}}(z) \mathcal{L}(\tilde{z}_i|z, \sigma_{z,i}) \\ &= \left(\prod_{i=1}^{N_{\text{gal}}^{\text{obs}}} \int dz p_{\text{gal}}(z) P_{\text{det}}(z) \mathcal{L}(\tilde{z}_i|z, \sigma_{z,i}) \right) \int_0^\infty \frac{dN}{N} N^{N_{\text{gal}}^{\text{obs}}} e^{-N \int dz p_{\text{gal}}(z) P_{\text{det}}(z)} \\ &= \left(\frac{1}{(\alpha(p_{\text{gal}}))^{N_{\text{gal}}^{\text{obs}}}} \prod_{i=1}^{N_{\text{gal}}^{\text{obs}}} \int dz p_{\text{gal}}(z) P_{\text{det}}(z) \mathcal{L}(\tilde{z}_i|z, \sigma_{z,i}) \right) \int_0^\infty dt t^{N_{\text{gal}}^{\text{obs}}-1} e^{-t}, \end{aligned} \quad (\text{A10})$$

where we used the integration variable $t = N\alpha(p_{\text{gal}})$ and we defined

$$\alpha(p_{\text{gal}}) = \int dz p_{\text{gal}}(z) P_{\text{det}}(z), \quad (\text{A11})$$

representing the fraction of galaxies in the population that are expected to be detected. The last integral in Eq. (A10) is simply equal to a constant factor $(N_{\text{gal}}^{\text{obs}} - 1)!$ and can thus be omitted. The final expression for the scale-free likelihood is then

$$\mathcal{L}(\{\tilde{z}_i\}|p_{\text{gal}}) \propto \prod_{i=1}^{N_{\text{gal}}^{\text{obs}}} \frac{\int dz p_{\text{gal}}(z) P_{\text{det}}(z) \mathcal{L}(\tilde{z}_i|z, \sigma_{z,i})}{\alpha(p_{\text{gal}})}. \quad (\text{A12})$$

APPENDIX B: KERNEL PARAMETER CONSTRAINTS

In Fig. 12, we plot the posterior distribution for the GP kernel parameters obtained with the LOS shown in Fig. 6.

We note in particular that the mean value of ρ , defining the length scale at which the GP realizations oscillate, increases as the redshift error increases. Physically, this means that as the redshift error increases, thereby smoothing the observed redshift distribution, the GP oscillations along redshift space are smaller and less numerous. On the other hand, the amplitude scale parameter σ is comparable when varying the error on galaxy redshifts. When using the 8% complete catalog ($m_{\text{thr}} = 18.5$), the posteriors are usually wider than those obtained with the 24% complete catalog ($m_{\text{thr}} = 19.5$). This is expected since there is less data to constrain the GP kernel parameters.

APPENDIX C: ALL HUBBLE CONSTANT CONSTRAINTS

In Table I we summarize the median and the 68% C.I. for all cases studied. We note that all methods yield unbiased H_0 constraints, as the fiducial value is consistently recovered within the 68% C.I.

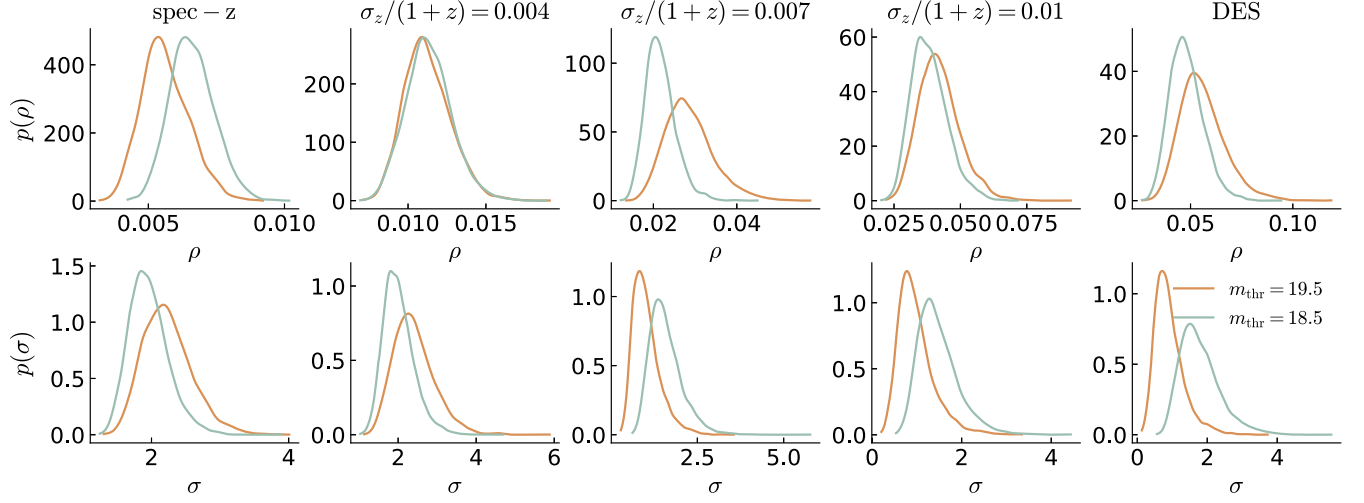


FIG. 12. Posterior distribution for the GP kernel parameters under different redshift error assumptions and for the two different magnitude cuts. These results are relative to the LOS data plotted in Fig. 6.

TABLE I. Median H_0 values and 68% C.I. (in $\text{km s}^{-1} \text{Mpc}^{-1}$) for the three cases studied: the GP-completed catalog, the standard homogeneous completion (UCV), and the complete galaxy catalog (COMP). The symmetric percentage errors are shown in parentheses.

m_{thr}	A	d_L^{thr}	Method	spec-z	0.004	0.007	0.01	DES
18.5	0.1	1.4	GP	$68.7^{+1.4}_{-1.4}$ (2.0%)	$68.8^{+1.4}_{-1.4}$ (2.0%)	$68.8^{+1.4}_{-1.5}$ (2.1%)	$68.4^{+1.6}_{-1.6}$ (2.3%)	$68.2^{+1.9}_{-1.8}$ (2.7%)
			COMP	$69.0^{+1.5}_{-1.5}$ (2.2%)	$68.5^{+1.7}_{-1.6}$ (2.4%)	$68.5^{+2.0}_{-1.8}$ (2.8%)	$68.4^{+2.4}_{-2.2}$ (3.4%)	$68.6^{+2.8}_{-2.6}$ (3.9%)
			UCV	$69.1^{+2.2}_{-2.1}$ (3.1%)	$68.4^{+2.5}_{-2.4}$ (3.6%)	$68.1^{+3.0}_{-2.9}$ (4.3%)	$68.1^{+3.6}_{-3.4}$ (5.1%)	$69.0^{+4.7}_{-4.1}$ (6.4%)
		1.6	GP	$69.1^{+1.4}_{-1.4}$ (2.0%)	$68.9^{+1.4}_{-1.3}$ (2.0%)	$68.6^{+1.5}_{-1.4}$ (2.1%)	$69.4^{+1.6}_{-1.5}$ (2.2%)	$70.0^{+1.9}_{-1.8}$ (2.6%)
			COMP	$69.4^{+1.5}_{-1.4}$ (2.1%)	$69.0^{+1.6}_{-1.6}$ (2.3%)	$68.8^{+1.9}_{-1.7}$ (2.6%)	$68.9^{+2.1}_{-2.1}$ (3.0%)	$68.8^{+2.5}_{-2.3}$ (3.5%)
			UCV	$70.0^{+2.6}_{-2.4}$ (3.6%)	$68.9^{+2.8}_{-2.5}$ (3.9%)	$68.7^{+3.4}_{-3.0}$ (4.7%)	$69.2^{+4.1}_{-3.7}$ (5.6%)	$69.4^{+4.9}_{-4.5}$ (6.8%)
	0.2	1.4	GP	$69.5^{+1.6}_{-1.6}$ (2.3%)	$69.1^{+1.5}_{-1.4}$ (2.1%)	$69.2^{+1.4}_{-1.3}$ (1.9%)	$69.4^{+1.6}_{-1.5}$ (2.2%)	$69.8^{+1.8}_{-1.8}$ (2.6%)
			COMP	$69.5^{+1.8}_{-1.7}$ (2.5%)	$69.6^{+1.9}_{-1.8}$ (2.7%)	$69.6^{+2.1}_{-2.1}$ (3.0%)	$69.7^{+2.3}_{-2.3}$ (3.3%)	$69.9^{+2.5}_{-2.6}$ (3.6%)
			UCV	$68.5^{+3.9}_{-3.8}$ (5.6%)	$68.7^{+4.0}_{-4.1}$ (5.9%)	$69.3^{+4.2}_{-4.4}$ (6.2%)	$70.2^{+4.3}_{-4.8}$ (6.5%)	$71.7^{+3.8}_{-4.9}$ (6.1%)
		1.6	GP	$68.6^{+3.1}_{-3.0}$ (4.5%)	$68.2^{+3.1}_{-2.9}$ (4.4%)	$68.3^{+3.0}_{-3.0}$ (4.4%)	$68.4^{+3.1}_{-3.1}$ (4.5%)	$68.2^{+3.4}_{-3.3}$ (4.9%)
			COMP	$68.1^{+3.0}_{-2.9}$ (4.3%)	$68.0^{+3.2}_{-3.1}$ (4.6%)	$67.8^{+3.4}_{-3.4}$ (5.0%)	$67.2^{+3.7}_{-3.7}$ (5.5%)	$66.2^{+4.0}_{-4.1}$ (6.1%)
			UCV	$66.6^{+3.5}_{-3.2}$ (5.0%)	$66.5^{+3.7}_{-3.6}$ (5.5%)	$66.3^{+4.1}_{-4.1}$ (6.2%)	$65.8^{+4.7}_{-4.7}$ (7.1%)	$65.0^{+5.3}_{-5.8}$ (8.5%)
0.2	0.1	1.4	GP	$68.6^{+4.0}_{-4.1}$ (5.9%)	$67.0^{+3.6}_{-3.7}$ (5.5%)	$66.3^{+3.5}_{-3.6}$ (5.3%)	$66.3^{+3.5}_{-3.5}$ (5.3%)	$66.5^{+3.5}_{-3.5}$ (5.3%)
			COMP	$68.1^{+3.9}_{-4.0}$ (5.8%)	$67.9^{+3.9}_{-4.0}$ (5.8%)	$68.1^{+3.7}_{-3.9}$ (5.6%)	$68.1^{+3.7}_{-4.0}$ (5.7%)	$68.0^{+3.7}_{-4.1}$ (5.7%)
			UCV	$67.8^{+4.5}_{-4.6}$ (6.7%)	$68.0^{+4.3}_{-4.5}$ (6.5%)	$68.5^{+4.2}_{-4.5}$ (6.3%)	$69.1^{+4.2}_{-4.4}$ (6.2%)	$69.9^{+4.1}_{-4.4}$ (6.1%)
		1.6	GP	$70.5^{+2.7}_{-2.9}$ (4.0%)	$69.0^{+2.6}_{-2.8}$ (3.9%)	$67.2^{+2.8}_{-2.9}$ (4.2%)	$67.1^{+2.8}_{-3.0}$ (4.3%)	$67.2^{+2.8}_{-3.2}$ (4.5%)
			COMP	$68.8^{+2.7}_{-2.9}$ (4.1%)	$68.5^{+2.7}_{-2.9}$ (4.1%)	$68.3^{+2.7}_{-3.0}$ (4.2%)	$67.9^{+2.7}_{-3.1}$ (4.3%)	$67.5^{+2.7}_{-3.1}$ (4.3%)
			UCV	$68.4^{+3.0}_{-3.3}$ (4.6%)	$68.4^{+3.0}_{-3.3}$ (4.6%)	$68.4^{+3.0}_{-3.4}$ (4.7%)	$68.5^{+3.0}_{-3.4}$ (4.7%)	$68.6^{+3.1}_{-3.4}$ (4.7%)

(Table continued)

TABLE I. (Continued)

m_{thr}	A	d_L^{thr}	Method	spec-z	0.004	0.007	0.01	DES
19.5	0.1	1.4	GP	$69.2^{+1.5}_{-1.6}$ (2.2%)	$69.0^{+1.6}_{-1.4}$ (2.2%)	$68.7^{+1.6}_{-1.5}$ (2.3%)	$68.7^{+1.6}_{-1.6}$ (2.3%)	$68.7^{+1.8}_{-1.8}$ (2.6%)
			COMP	$69.3^{+1.5}_{-1.5}$ (2.2%)	$68.9^{+1.7}_{-1.6}$ (2.4%)	$68.5^{+2.0}_{-1.8}$ (2.8%)	$68.5^{+2.3}_{-2.3}$ (3.4%)	$68.6^{+2.8}_{-2.6}$ (3.9%)
			UCV	$69.2^{+1.7}_{-1.7}$ (2.5%)	$68.6^{+1.9}_{-1.8}$ (2.7%)	$68.0^{+2.2}_{-2.1}$ (3.2%)	$67.9^{+2.6}_{-2.5}$ (3.8%)	$68.3^{+3.2}_{-3.0}$ (4.5%)
		1.6	GP	$69.3^{+1.5}_{-1.4}$ (2.1%)	$69.2^{+1.4}_{-1.5}$ (2.1%)	$69.0^{+1.6}_{-1.5}$ (2.2%)	$69.3^{+1.6}_{-1.5}$ (2.2%)	$69.7^{+1.7}_{-1.7}$ (2.4%)
			COMP	$69.5^{+1.5}_{-1.4}$ (2.1%)	$69.1^{+1.6}_{-1.5}$ (2.2%)	$68.9^{+1.8}_{-1.8}$ (2.6%)	$68.9^{+2.1}_{-2.1}$ (3.0%)	$68.8^{+2.5}_{-2.3}$ (3.5%)
			UCV	$69.5^{+1.9}_{-1.7}$ (2.6%)	$69.0^{+2.0}_{-1.9}$ (2.8%)	$69.0^{+2.4}_{-2.2}$ (3.3%)	$69.4^{+2.9}_{-2.6}$ (4.0%)	$69.9^{+3.4}_{-3.1}$ (4.7%)
		1.8	GP	$69.2^{+1.7}_{-1.7}$ (2.5%)	$69.3^{+1.7}_{-1.7}$ (2.5%)	$69.3^{+1.7}_{-1.7}$ (2.5%)	$69.2^{+1.8}_{-1.8}$ (2.6%)	$69.1^{+1.8}_{-1.8}$ (2.6%)
			COMP	$69.2^{+1.8}_{-1.8}$ (2.6%)	$69.4^{+1.9}_{-1.9}$ (2.7%)	$69.6^{+2.1}_{-2.1}$ (3.0%)	$69.7^{+2.3}_{-2.3}$ (3.3%)	$69.9^{+2.5}_{-2.6}$ (3.6%)
			UCV	$67.9^{+2.5}_{-2.4}$ (3.6%)	$67.9^{+2.8}_{-2.7}$ (4.0%)	$68.2^{+3.1}_{-3.0}$ (4.5%)	$68.0^{+3.7}_{-3.4}$ (5.2%)	$68.4^{+4.1}_{-3.9}$ (5.8%)
	0.2	1.4	GP	$68.5^{+3.1}_{-3.0}$ (4.5%)	$68.5^{+3.1}_{-3.0}$ (4.5%)	$68.4^{+3.2}_{-3.0}$ (4.5%)	$68.6^{+3.2}_{-3.2}$ (4.7%)	$68.8^{+3.3}_{-3.4}$ (4.9%)
			COMP	$68.3^{+3.0}_{-2.9}$ (4.3%)	$68.1^{+3.2}_{-3.1}$ (4.6%)	$67.8^{+3.4}_{-3.4}$ (5.0%)	$67.2^{+3.6}_{-3.7}$ (5.4%)	$66.2^{+4.0}_{-4.1}$ (6.1%)
			UCV	$67.9^{+3.2}_{-3.0}$ (4.6%)	$68.0^{+3.4}_{-3.2}$ (4.8%)	$67.7^{+3.7}_{-3.6}$ (5.4%)	$67.4^{+4.0}_{-4.0}$ (5.9%)	$66.7^{+4.4}_{-4.4}$ (6.6%)
		1.6	GP	$68.2^{+3.7}_{-3.8}$ (5.5%)	$67.9^{+3.7}_{-3.9}$ (5.6%)	$68.0^{+3.7}_{-3.7}$ (5.4%)	$68.5^{+3.5}_{-3.7}$ (5.3%)	$69.1^{+3.4}_{-3.6}$ (5.1%)
			COMP	$68.1^{+3.9}_{-3.9}$ (5.7%)	$68.0^{+3.7}_{-4.0}$ (5.7%)	$68.1^{+3.7}_{-3.9}$ (5.6%)	$68.1^{+3.7}_{-4.0}$ (5.7%)	$68.0^{+3.7}_{-4.1}$ (5.7%)
			UCV	$68.2^{+4.1}_{-4.1}$ (6.0%)	$68.0^{+4.0}_{-4.1}$ (6.0%)	$68.6^{+4.0}_{-4.0}$ (5.8%)	$69.2^{+4.0}_{-4.2}$ (5.9%)	$69.9^{+3.9}_{-4.3}$ (5.9%)
		1.8	GP	$68.6^{+2.6}_{-2.9}$ (4.0%)	$68.8^{+2.7}_{-2.9}$ (4.0%)	$68.7^{+2.6}_{-2.9}$ (4.0%)	$68.7^{+2.6}_{-2.9}$ (4.0%)	$68.6^{+2.7}_{-2.9}$ (4.1%)
			COMP	$68.6^{+2.7}_{-3.0}$ (4.2%)	$68.4^{+2.7}_{-2.9}$ (4.1%)	$68.3^{+2.7}_{-3.0}$ (4.2%)	$67.9^{+2.7}_{-3.1}$ (4.3%)	$67.5^{+2.7}_{-3.1}$ (4.3%)
			UCV	$68.5^{+2.8}_{-3.1}$ (4.3%)	$68.5^{+2.9}_{-3.1}$ (4.4%)	$68.6^{+2.9}_{-3.2}$ (4.5%)	$68.6^{+2.9}_{-3.3}$ (4.5%)	$68.5^{+3.0}_{-3.3}$ (4.6%)

- [1] N. Aghanim *et al.*, *Astron. Astrophys.* **641**, A6 (2020).
[2] A. G. Riess *et al.*, *Astrophys. J. Lett.* **934**, L7 (2022).
[3] M. Moresco *et al.*, *Living Rev. Relativity* **25**, 6 (2022).
[4] E. Di Valentino *et al.*, *Phys. Dark Universe* **49**, 101965 (2025).
[5] D. E. Holz and S. A. Hughes, *Astrophys. J.* **629**, 15 (2005).
[6] S. Nissanke, D. E. Holz, S. A. Hughes, N. Dalal, and J. L. Sievers, *Astrophys. J.* **725**, 496 (2010).
[7] B. P. Abbott *et al.*, *Phys. Rev. Lett.* **119**, 161101 (2017).
[8] B. P. Abbott *et al.*, *Astrophys. J. Lett.* **848**, L12 (2017).
[9] B. P. Abbott *et al.*, *Nature (London)* **551**, 85 (2017).
[10] B. F. Schutz, *Nature (London)* **323**, 310 (1986).
[11] W. Del Pozzo, *Phys. Rev. D* **86**, 043011 (2012).
[12] H.-Y. Chen, M. Fishbach, and D. E. Holz, *Nature (London)* **562**, 545 (2018).
[13] M. Fishbach *et al.*, *Astrophys. J. Lett.* **871**, L13 (2019).
[14] A. Palmese *et al.*, *Astrophys. J. Lett.* **900**, L33 (2020).
[15] J. R. Gair *et al.*, *Astron. J.* **166**, 22 (2023).
[16] R. Gray *et al.*, *Phys. Rev. D* **101**, 122001 (2020).
[17] G. Perna, S. Mastrogiovanni, and A. Ricciardone, *Astron. Astrophys.* **698**, A128 (2025).
[18] A. G. Hanselman, A. Vijaykumar, M. Fishbach, and D. E. Holz, *Astrophys. J.* **979**, 9 (2025).
[19] N. Borghi, M. Moresco, M. Tagliazucchi, and G. Cuomo, *Astron. Astrophys.* **706**, A199 (2026).
[20] M. VanWygarden, M. Fishbach, A. Vijaykumar, A. G. Guerrero, and D. E. Holz, *Astrophys. J.* **1008**, 52 (2026).
[21] V. Alfradique, C. R. Bom, and T. Castro, *Phys. Rev. D* **112**, 063561 (2025).
[22] M. L. Cross-Parkin, C. Howlett, T. M. Davis, and N. Khetan, *arXiv:2502.17747*.
[23] A. Finke, S. Foffa, F. Iacovelli, M. Maggiore, and M. Mancarella, *J. Cosmol. Astropart. Phys.* **08** (2021) 026.
[24] R. Gray, C. Messenger, and J. Veitch, *Mon. Not. R. Astron. Soc.* **512**, 1127 (2022).
[25] R. Gray *et al.*, *J. Cosmol. Astropart. Phys.* **12** (2023) 023.
[26] S. Mastrogiovanni, D. Laghi, R. Gray, G. C. Santoro, A. Ghosh, C. Karathanasis, K. Leyde, D. A. Steer, S. Perriès, and G. Pierra, *Phys. Rev. D* **108**, 042002 (2023).
[27] N. Borghi, M. Mancarella, M. Moresco, M. Tagliazucchi, F. Iacovelli, A. Cimatti, and M. Maggiore, *Astrophys. J.* **964**, 191 (2024).
[28] L. E. H. Datrier and M. A. Hendry, *Mon. Not. R. Astron. Soc.* **549**, stag828 (2026).
[29] The LIGO Scientific, the Virgo, and the KAGRA Collaborations, *arXiv:2605.27227*.

- [30] M. Kalomenopoulos, R. Barbieri, S. Khochfar, J. Gair, and R. J. McGibbon, [arXiv:2511.12334](#).
- [31] C. Dalang and T. Baker, *J. Cosmol. Astropart. Phys.* **02** (2024) 024.
- [32] K. Leyde, T. Baker, and W. Enzi, *J. Cosmol. Astropart. Phys.* **12** (2024) 013.
- [33] C. Dalang, B. Fiorini, and T. Baker, *J. Cosmol. Astropart. Phys.* **01** (2026) 034.
- [34] K. Leyde, T. Baker, and W. Enzi, *J. Cosmol. Astropart. Phys.* **01** (2026) 013.
- [35] R. Barbieri, M. Kalomenopoulos, M. Tagliazucchi, J. Gair, and S. Khochfar, [arXiv:2608.30599](#).
- [36] L.-G. Zhu, Y.-M. Hu, H.-T. Wang, J.-d. Zhang, X.-D. Li, M. Hendry, and J. Mei, *Phys. Rev. Res.* **4**, 013247 (2022).
- [37] K. Naveed, C. Turski, and A. Ghosh, *Mon. Not. R. Astron. Soc.* **548**, [stag605](#) (2026).
- [38] F. Beirnaert, G. Dálya, and A. Ghosh, *Mon. Not. R. Astron. Soc.* **542**, 3346 (2025).
- [39] G. Scelfo, M. Spinelli, A. Raccaelli, L. Boco, A. Lapi, and M. Viel, *J. Cosmol. Astropart. Phys.* **01** (2022) 004.
- [40] U. Dupletsa *et al.*, [arXiv:2605.12606](#).
- [41] M. Oguri, *Phys. Rev. D* **93**, 083511 (2016).
- [42] F. Calore, A. Cuoco, T. Regimbau, S. Sachdev, and P. D. Serpico, *Phys. Rev. Res.* **2**, 023314 (2020).
- [43] S. Bera, D. Rana, S. More, and S. Bose, *Astrophys. J.* **902**, 79 (2020).
- [44] S. Mukherjee, A. Krolewski, B. D. Wandelt, and J. Silk, *Astrophys. J.* **975**, 189 (2024).
- [45] T. Ghosh, S. More, S. Bera, and S. Bose, *Phys. Rev. D* **111**, 063513 (2025).
- [46] J. Ferri, I. L. Tashiro, L. R. Abramo, I. Matos, M. Quartin, and R. Sturani, *J. Cosmol. Astropart. Phys.* **04** (2025) 008.
- [47] J. Pan, Dragan Huterer, C. Avestruz, D. H. T. Cheung, E. Trott, N. Dalal, and D. Jeong, *Phys. Rev. D* **113**, 103532 (2026).
- [48] A. Pedrotti, M. Mancarella, J. Bel, M. Santoni, and D. Gerosa, *Astron. Astrophys.* **712**, A37 (2026).
- [49] I. Santiago de Matos, C. Dalang, T. Baker *et al.*, [arXiv:2512.15380](#).
- [50] A. Q. Cheng and J. Gair, [arXiv:2603.13053](#).
- [51] M. L. Cross-Parkin, C. Howlett, L. Giani, C. Blake, and T. M. Davis, [arXiv:2605.06783](#).
- [52] I. Mandel, W. M. Farr, and J. R. Gair, *Mon. Not. R. Astron. Soc.* **486**, 1086 (2019).
- [53] S. Vitale, D. Gerosa, W. M. Farr, and S. R. Taylor, *Handbook of Gravitational Wave Astronomy* (Springer, Singapore, 2022).
- [54] S. Mastrogiovanni, G. Pierra, S. Perriès, G. Pierra, S. Perriès, D. Laghi, G. Caneva Santoro, A. Ghosh, R. Gray, C. Karathanasis, and K. Leyde, *Astron. Astrophys.* **682**, A167 (2024).
- [55] P. Schechter, *Astrophys. J.* **203**, 297 (1976).
- [56] J. Carretero, F. J. Castander, E. Gaztanaga, M. Crocce, and P. Fosalba, *Mon. Not. R. Astron. Soc.* **447**, 650 (2015).
- [57] P. Fosalba, M. Crocce, E. Gaztanaga, and F. J. Castander, *Mon. Not. R. Astron. Soc.* **448**, 2987 (2015).
- [58] P. Fosalba, E. Gaztanaga, F. J. Castander, and M. Crocce, *Mon. Not. R. Astron. Soc.* **447**, 1319 (2015).
- [59] K. Hoffmann, J. Bel, E. Gaztanaga, J. Bel, E. Gaztanaga, M. Crocce, P. Fosalba, and F. J. Castander, *Mon. Not. R. Astron. Soc.* **447**, 1724 (2015).
- [60] A. G. Abac *et al.*, *Astrophys. J. Lett.* **1004**, L22 (2026).
- [61] A. G. Abac *et al.*, *Astrophys. J. Lett.* **1007**, L17 (2026).
- [62] A. G. Abac *et al.*, [arXiv:2605.27225](#).
- [63] T. Abbott *et al.*, [arXiv:astro-ph/0510346](#).
- [64] M. Soares-Santos *et al.*, *Astrophys. J. Lett.* **876**, L7 (2019).
- [65] B. P. Abbott *et al.*, *Living Rev. Relativity* **19**, 1 (2016).
- [66] J. Bradbury, R. Frostig, P. Hawkins *et al.*, <http://github.com/jax-ml/jax>.
- [67] D. Foreman-Mackey, TINYGP: The tiniest of Gaussian Process libraries, Python software, Version v0.3.0, Zenodo (2024), [10.5281/zenodo.6389737](#)
- [68] M. D. Hoffman and A. Gelman, [arXiv:1111.4246](#).
- [69] D. Phan, N. Pradhan, and M. Jankowiak, [arXiv:1912.11554](#).
- [70] T. J. Loredo, *AIP Conf. Proc.* **735**, 195 (2004).