Multiband gravitational wave cosmology with stellar origin black hole binaries

Niccolò Muttoni[1,2,*] Alberto Mangiagli,[1,3] Alberto Sesana,[1,4] Danny Laghi,[5,6,2] Walter Del Pozzo,[5,6]
David Izquierdo-Villalba,[1] and Mattia Rosati[1]

[1]*Department of Physics, G. Occhialini, University of Milano—Bicocca,
Piazza della Scienza 3, 20126 Milano, Italy*
[2]*Laboratoire des 2 Infinis—Toulouse (L2IT-IN2P3), Université de Toulouse,
CNRS, UPS, F-31062 Toulouse Cedex 9, France*
[3]*APC, AstroParticule et Cosmologie, Université de Paris, CNRS, F-75013 Paris, France*
[4]*INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy*
[5]*Dipartimento di Fisica "Enrico Fermi", Università di Pisa, Largo Pontecorvo 3, I-56127 Pisa, Italy*
[6]*INFN, Sezione di Pisa, Largo Pontecorvo 3, I-56127 Pisa, Italy*

Massive stellar origin black hole binaries (SBHBs), originating from stars above the pair-instability mass gap, are primary candidates for multiband gravitational wave (GW) observations. Here we study the possibility to use them as effective dark standard sirens to constrain cosmological parameters. The long lasting inspiral signal emitted by these systems is accessible by the future *Laser Interferometer Space Antenna* (LISA), while the late inspiral and merger are eventually detected by third generation ground-based telescopes such as the *Einstein Telescope* (ET). The direct measurement of the luminosity distance and the sky position to the source, together with the inhomogeneous redshift distribution of possible host galaxies, allow us to infer cosmological parameters by probabilistic means. The efficiency of this statistical method relies in high parameter estimation performances. We show that this multiband approach allows a precise determination of the Hubble constant H_0 with just $\mathcal{O}(10)$ detected sources. For selected SBHB population models, assuming 4 (10) years of LISA observations, we find that H_0 is typically determined at $\sim 2\%$ ($\sim 1.5\%$), whereas Ω_m is only mildly constrained with a typical precision of 30% (20%). We discuss the origin of some outliers in our final estimates and we comment on ways to reduce their presence.

I. INTRODUCTION

Recent analysis, relying on Type Ia supernovae (SNe) [1] and Planck measurements [2], revealed a tension at 4.4σ on the determination of the Hubble constant, H_0. While the former reported $H_0 = 74.03 \pm 1.42$ km s^{-1} Mpc^{-1}, the latter derived $H_0 = 67.4 \pm 0.5$ km s^{-1} Mpc^{-1}. Although a number of possible solutions have been proposed (see, e.g., [3]), the tension still persists.

The first gravitational wave (GW) detection [4] brightened the future of astrophysical observations. Among the possible theoretical [5] and cosmological applications [6], GWs offer a unique opportunity to provide an independent constraint on the H_0 parameter, thus shedding light on the evolution history of our Universe. Coalescing compact object systems are ideal standard sirens in the determination of cosmological distances and their signals bring direct information about the source's luminosity distance. However no information is carried about the redshift. If an electro-magnetic (EM) counterpart is detected, the identification of the host galaxy might provide the redshift information necessary to build the $d_L - z$ relation and constrain H_0 [7].

Even if no information about the source redshift is gathered, a measurement of H_0 can be performed by exploiting the statistical properties of the inhomogeneous redshift distribution of possible galaxy hosts [8]. This method has been developed in the last decade [e.g., [9–12]] and has been successfully applied to recent LIGO-Virgo observations, although yielding only mild constraints [6].

The future *Laser Interferometer Space Antenna* [LISA, [13]] will extend the frequency band currently explored by ground based detectors to the mHz region. Moreover, the third generation interferometers Einstein Telescope [ET, [14]] and Cosmic Explorer [CE, [15]] will remarkably enhance the GW sensitivity from the ground, thus enabling the exploration of a larger portion of the Universe. LISA will detect the early inspiral of stellar black hole binaries (SBHBs) [16]. A fraction of these systems will coalesce within only few years from the first LISA observation, becoming observable by ground based detectors, thus fostering a multiband approach.

*n.muttoni@campus.unimib.it

Combining information from ground and space detectors might provide unique scientific outcomes. This is particularly true in the case of multiband SBHB exploitation for cosmological measurements. While LISA will determine the sky position of the GW source to great accuracy, due to the long persistence of the signal in band, ET will pin down its luminosity distance, thanks to the high signal-to-noise ratio (S/N). By combining these two measurements, the origin of the signal can be constrained to a small 3D volume, encompassing only a small number of candidate galaxy hosts. Therefore, multiband GW sources might prove to be a particularly powerful class of "dark" standard sirens.

The idea of multiband GW sources was first proposed in the context of intermediate mass black holes (IMBHs) [17] and later revised in light of the observation of GW150914 [16]. Although the idea is appealing, recent estimates of the SBHB merger rate and mass function coupled with the current LISA sensitivity curve result in rather pessimistic prospects of detection. In fact, when considering only BHs below the pair-instability mass gap, LISA might perhaps detect only a handful of such multiband sources [18]. However, things change when the BH mass function is extended beyond the mass gap. As shown in [19], depending on the details of the SBHB population and on the duration of the LISA mission, "above-gap" SBHBs might dominate the population of multiband sources [20], with up to about a hundred of detected systems.

Here we exploit SBHBs above the pair-instability gap jointly observed by LISA and ET (which will be our 3G detector default choice) as dark standard sirens to infer the cosmological parameters via statistical methods. We simulate a population of inspiralling SBHBs in LISA, focusing on the systems that are going to coalesce during the LISA time mission. We perform parameter estimation of the multiband binaries under the Fisher matrix formalism with LISA and ground-based detectors. Combining the sky localization uncertainty provided by LISA and the estimate on the luminosity distance by ET, we construct error boxes in the sky. We populate these volumes with realistic galaxy catalogs [21], and we infer H_0, as well as other cosmological parameters, by applying a statistical nested sampling algorithm.

A very similar approach was taken in [12]. Their study focused on the inference of cosmological parameters via LISA-only observations of SBHBs below the pair-instability mass gap. Assuming different detector configurations, they found that the Hubble constant is determined between ∼5% and ∼2%. In the past few years, however, LISA sensitivity has been downgraded by 50% at the high frequency end [22], and the latest observing run by LIGO/Virgo reported an intrinsic merger rate of 24 Gpc^{-3} yr^{-1} for SBHBs [6]. This will inevitably impact the results of [12], yielding a worse estimate of H_0. Therefore, above-gap SBHBs could very well represent the only possibility to perform precision

cosmology in this mass range, unless EM counterpart are found to be associated to SBHB mergers [23].

Even if EM-based measurements are going to be improved by the time LISA and ET start observing, we stress that GW-based observations rely on completely independent assumptions and systematics, and for distance measurements they do not require complex calibrations as for the construction of a cosmic distance ladder.

The paper is structured as follows: in Sec. II we present the assumptions under which we perform our analyses. In particular, in Sec. II A we describe the astrophysical GW sources that we consider in our work and their observational properties. In Sec. II B we present the analytical framework adopted to perform the SBHB parameter estimation. Section II C is devoted to the description of the error box construction and population, and Sec. II D, describes the statistical framework and the numerical implementation of the inference of cosmological parameters from SBHB observations. Our main results are presented and extensively discussed in Sec. III, while Sec. IV summarizes the main features of our work and lays out future plans.

II. PRELIMINARIES

A. Binary population

A multiband approach requires multiband sources. In this work we explore the realm of SBHBs above the pair-instability mass gap. Current stellar evolution models predict the lack of stellar black holes in the mass range between ∼60 M$_\odot$ and ∼120 M$_\odot$ [24]. Depending on the metallicity, the hot core of very massive stars might undergo electron-positron pairs production. These processes soften the star's equation of state and bring the star to collapse. The rising temperatures from the contraction then trigger thermonuclear runaway reactions that completely disrupt the progenitor in a luminous *pair-instability supernova* (PISN) event, thus leaving no remnant [25]. On the contrary, no physical process can halt the collapse triggered in the last life stages of very massive stars, and the progenitor leaves a BH remnant with mass greater than ∼120 M$_\odot$ [24]. In this study we will consider SBHBs composed by such "above-gap" objects, i.e., BHs with masses in the range 120 M$_\odot$ − 300 M$_\odot$, and we will assume the mass gap has a sharp cutoff at [60, 120] M$_\odot$ (but look also at [26] for the effects of star rotation and compactness on the BHs mass distribution).

When bounded in coalescing binaries, above-gap BHs emit a GW signal which crosses multiple frequency bands: from the mHz, where LISA observes the long lasting inspiral phase of the system, up to $\mathcal{O}(10^2)$ Hz, where ET detects the last-inspiral cycles, the merger, and the ringdown. A multiband signal requires to be detected (i.e., it has to be revealed with S/N greater than a fixed threshold) by both interferometers, with LISA being the first of them.

These loud GW sources were extensively investigated in [19] and in this work we rely on those results to perform our analysis. Assuming the optimistic scenario developed by the authors, the LISA merger rate of above-gap GW sources is estimated to be between $R \simeq 10$ yr^{-1} and $R \simeq 14$ yr^{-1}. A LISA mission time of 4 (10) years would then reveal ~40 (~140) multiband events. However, as in [19], we first consider three different subpopulations of SBHBs: below-gap (above-gap) binaries, where both components are below (above) 60 M$_\odot$ (120 M$_\odot$), and across-gap binaries with the primary (secondary) BH above (below) 120 M$_\odot$ (60 M$_\odot$). We consider the same models described in that study, i.e., an optimistic (sSFR-sZ) and pessimistic (mSFR-mZ) model plus two intermediates ones (mSFR-sZ and sSFR-mZ). The optimistic model features an higher star formation rate and lower metallicity, hence leading to more massive SBHBs than the pessimistic one. We defer to the original paper for extensive details on how the population of SBHBs is built and we summarize here only the main properties of our binary population model.

We consider two models for the SFR and the average metallicity evolution as function of redshift. The pessimistic SFR and metallicity are taken both from [27] while the optimistic SFR and metallicity are adopted from [28,29], respectively. We assume a power-law stellar initial mass function (IMF) $\xi(M_\star, \alpha) \propto M_\star^{-\alpha}$ extending in the range [8, 350], with $\alpha = 2.7$ for the pessimistic SFR and $\alpha = 2.35$ for the optimistic one. Single stars are evolved with the code SEVN [24] and the resulting BHs are combined assuming a flat mass-ratio distribution in [0.1, 1] and a log-flat time delay distribution in [50 Myr, t_{Hubble}]. The frequency distribution in LISA is computed in the quadrupole approximation for circular orbits following [30].

We stress that our population model is simplified in many ways. Perhaps most importantly, the resulting BH mass function is obtained by evolving individual stars only, neglecting the effect of binary interaction. In fact, binary evolution models have been generally been implemented for stars up to ≈ 150 M$_\odot$ [31,32], therefore preventing the possibility to study above-gap remnants in binaries. Moreover, we consider only the standard field formation channel, neglecting alternatives such as dynamical capture [33], hierarchical mergers [34] and accretion and mergers in AGN disks [35]. In this respect, our BH binary population can be considered a "toy model" to provide a proof of principle demonstration of the multiband approach.

In our analysis there are two important differences with respect to [19]: we normalized the overall rate at 25 Gpc^{-3} yr^{-1} to best match the updates from the latest LIGO/Virgo results and we changed the cosmology choosing the same cosmological values used for the generation of the light cones (see Sec. II C for more details) for internal consistency. For each subpopulation and each model, we generated 30 Monte Carlo realizations of the expected SBHB population. At this step, each binary is characterized

by the two component masses, the merging redshift and the initial frequency. The redshift is sampled between 10^{-3} and ~19 in log scale, while the initial frequency is in the interval $[2 \times 10^{-4}, 1]$ Hz. We checked that extending the minimum frequency range to lower frequencies did not affect the number of observable systems. The frequency is then converted into a coalescence timescale t_c, via the standard quadrupole approximation [36], assuming circular binaries. Being interested in multiband events, we selected only systems with < 20 years.[1] Assuming a LISA mission lifetime of 10 years, we expect exquisite estimates on the coalescence time for those systems, of the order $\Delta t_c \simeq$ [1–10] s [16], allowing the unequivocal identification of the binary by ground-based detectors.

B. Parameter estimation formalism

To simulate observations with LISA and ET, we evaluate the error on the estimated parameters of each source by means of the Fisher information matrix. The general output of a detector is a time series $s(t)$ given by the superposition of the noise contribution $n(t)$ and a GW signal $h(t)$, if present:

$$s(t) = n(t) + h(t). \tag{1}$$

Assuming stationary, Gaussian white noise, the S/N produced by the GW signal is described by

$$\left(\frac{S}{N}\right)^2 = (h|h), \tag{2}$$

where the round brackets $(|)$ refer to the inner product between two real functions $A(t)$ and $B(t)$ defined as

$$(A|B) = 4\mathrm{Re} \int_0^{+\infty} df \frac{\tilde{A}^*(f)\tilde{B}(f)}{S_n(f)}. \tag{3}$$

Here the tilde labels Fourier transformed quantities, while the star denotes complex conjugate quantities. The S_n term represents the *power spectral density* (PSD) of the detector in Hz^{-1} units.

The GW signal $h(t, \Theta)$ produced by a spinning, precessing, and eccentric binary in a detector is characterized by a set of 17 parameters $\Theta = \{\Theta_1, \ldots, \Theta_{17}\}$. In the limit of high S/N, the probability (or likelihood) that the observed signal is described by Θ, given an output $s(t)$, is

$$p(\Theta|s) \propto \exp\left[-\frac{1}{2}(\partial_i h|\partial_j h)\Delta\Theta_i\Delta\Theta_j\right], \tag{4}$$

---

[1]Even though LISA will be able to detect systems much further from coalescence, these SBHBs will be characterized by large error (especially on luminosity distance) and therefore will add little information to the measurement of cosmological parameters.

where ∂_i denotes the derivative of the signal $h(t, \Theta)$ with respect to the parameter Θ_i. Equation (4) describes a multivariate Gaussian distribution centred in Θ and with covariance matrix $\Sigma = (\partial_i h | \partial_j h)^{-1}$. The Fisher information matrix Γ is then defined as

$$\Gamma_{ij} = (\partial_i h | \partial_j h), \tag{5}$$

and the parameter uncertainties and covariances are thus contained in its inverse Σ. Moreover, since LISA and ET are independent detectors, we can construct a more informative Fisher matrix by simply adding the individual ones:

$$\Gamma_{ET+LISA} = \Gamma_{ET} + \Gamma_{LISA}. \tag{6}$$

The new matrix contains the features of both detectors, therefore yielding better constraints on source parameters, in particular on sky location and luminosity distance.

For each binary sampled from our distributions as described in Sec. II A, we compute the S/N and Fisher matrix in LISA. We adopt the same sensitivity curve described in [37]. The signal in LISA is described by the inspiral-only precessing waveform presented in [38], assuming random spin magnitude in [0, 1] aligned with the binary angular momentum for consistency with the waveform adopted for ground-based detectors, as discussed below. Sky position and direction of the binary orbital angular momentum are randomly sampled from a uniform distribution on the sphere. We choose $S/N = 8$ as threshold for LISA detection and for each detected binary we compute the 15×15 Fisher matrix according to Eq. (5). Due to the numerical nature of the problem, we must check the outcome of the subsequent inversion process. We checked the discrepancy between the matrix product $(\Gamma \cdot \Sigma)_{ij}$ and the Kronecker symbol δ_{ij} through

$$\varepsilon_{inv} = \max_{i,j} |(\Gamma \cdot \Sigma)_{ij} - \delta_{ij}|, \tag{7}$$

and we consider the inversion successful if $\varepsilon_{inv} \leq 10^{-3}$ (see the appendix in [39] for details on the procedure).

All binaries detected by LISA are then analyzed with a pipeline for ground-based detectors. The ET Fisher matrix [again defined through Eq. (5)] is computed using the PYTHON library PYCBC [40], which can perform the inner products in Eq. (3) for a number of waveform approximants. Since ET will also detect the merger and ringdown, we adopt the *PhenomD* waveform [41] to model the full GW signal. Due to its domain of definition, the waveform only allows us to study nonprecessing spin-aligned binaries, as mentioned before. Therefore, of the original 17 source parameters we neglect the 4 corresponding to the x and y spin components of each BH. Two of the remaining source parameters are eccentricity related. However, GW emission eventually circularizes any eccentric orbit (but see also [42] for alternative scenarios): we choose to assume

only circular binaries, and we therefore ignore these parameters. The binaries and their signals are evolved through the high frequency band probed by ET, starting from 3 Hz, and we set an S/N threshold of 12 for the 3G interferometer detection.

Even considering circular binaries with aligned spins, we are still left with a 11×11 matrix that keeps track of the two BH masses and dimensionless spins, the luminosity distance, right ascension and declination of the source, the inclination and polarization angles, the merger time and phase. However, we find that in the case of ET the matrix is either ill-conditioned or singular [43], and the algorithm struggles with the matrix inversion process.[2] Thus, we are forced to select the larger subset of parameters that leads to acceptable ε_{inv} values. For what concerns the aim of this work, we choose to exclude the polarization and phase angle and the merger time. We do not expect the first two angles to have a strong impact on our parameter estimation and LISA will be anyway able to provide accurate estimates of the coalescence time early during the inspiral phase. In other words, our Fisher matrices are computed for the following subset of 8 parameters: the two BH masses M_1 and M_2 (with the condition $M_1 > M_2$), the luminosity distance d_L, the two BH spins χ_1 and χ_2, the right ascension RA and declination DEC of the binary and the inclination ι.[3] In Table I we show how we model the remaining quantities. We therefore reduce the Fisher matrices computed by LISA and ET to the 8 aforementioned parameters, add them, and invert the sum to get the correlation matrix. In this sense, we are not just looking at what each individual detector is able to perform, but we are summing the information matrices from the two detectors as in Eq. (6). The diagonal of the correlation matrix Σ contains the variance of each parameter, while the other entries represent their correlations. The error on d_L can be read directly out of the diagonal elements, whereas the sky position uncertainty area, in units of rad^2, is simply recovered through [44]

$$\Delta\Omega = 2\pi \sqrt{\Sigma_{\varphi_N \varphi_N} \Sigma_{\mu_N \mu_N} - (\Sigma_{\mu_N \varphi_N})^2}, \tag{8}$$

where $\Sigma_{\mu_N \mu_N}$ and $\Sigma_{\varphi_N \varphi_N}$ correspond to the diagonal elements for the right ascension and declination of the source and $\Sigma_{\mu_N \varphi_N}$ is their correlation value.

---

[2]To invert the matrices we adopt the LU decomposition.
[3]Although the ET 11×11 matrix turns out to be ill-conditioned, the LISA one is not. As a sanity check supporting the use of a reduced matrix, we computed the errors on the sky position and luminosity distance for LISA-only observations both for an 11×11 and 8×8 Fisher matrix. We found comparable results in the two cases, with the latter providing slightly more accurate numbers (by a factor ≈ 1.5) on average.

TABLE I. Parameters that characterize the Fisher matrix. We denote with a capital N subscript the celestial coordinates defined through $\theta_N = \frac{\pi}{2} - \text{DEC}$ and $\varphi_N = \text{RA}$. Furthermore, we take the natural logarithm of M_1, M_2, and d_L so to deal with relative errors.

Quantity	Parameter
Mass 1	$\ln M_1$
Mass 2	$\ln M_2$
Luminosity distance	$\ln d_L$
Spin 1	χ_1
Spin 2	χ_2
RA	φ_N
DEC	$\mu_N = \cos\theta_N$
Inclination	ι

C. Error box construction and population

The uncertainties on the luminosity distance and the sky location allow us to constrain the volume where the GW signal comes from, and galaxies inside it represent possible host candidates. We build the error boxes of the events following the procedure outlined in [45]. For a given cosmology, a $d_L \pm \sigma_{d_L}$ measure translates in a $z \pm \sigma_z$ interval, which is obtained by inverting the relation

$$d_L(z) = c(1+z) \int_0^z \frac{dz'}{\text{H}(z')}. \tag{9}$$

Here $\text{H}(z)$ represents the Hubble parameter as a function of redshift and its expression in a ΛCDM Universe is given by

$$\text{H}(z) = \text{H}_0 \sqrt{\Omega_m (1+z)^3 + \Omega_\Lambda}. \tag{10}$$

Since each set of cosmological parameters yields a different $d_L - z$ relation, the $z \pm \sigma_z$ redshift interval is extended in order to take into account for the prior ranges of the cosmological parameters one is willing to infer. From a measure of luminosity distance, we therefore construct a redshift range $[z^-, z^+]$, whose bounds represent the lowest and highest z obtained when the cosmology varies within the prior range of the cosmological parameters, which will be specified in Sec. II D. Furthermore, due to the peculiar velocity v_p of galaxies, the *cosmological* and *apparent* redshifts of a GW host might differ from each other. This is indeed an uncertainty source that we must account for. We characterize this error with $\sigma_{pv}(z) = (1+z)\sigma_{v_p}/c$, with $\sigma_{v_p} = 500 \text{ km s}^{-1}$, which is consistent with the standard deviation of the radial peculiar velocity distribution observed in the Millennium run [46]. In conclusion, each $d_L \pm \sigma_{d_L}$ measure translates in a $[z^- - \sigma_{pv}(z^-), z^+ + \sigma_{pv}(z^+)]$ redshift range. This corresponds to the redshift shell of the Universe encompassing all galaxies with a redshift consistent with the GW measured luminosity

distance, once the cosmological prior and peculiar velocities are taken into account.

The next step is to populate those redshift shells with realistic galaxy catalogs. To this end, we make use of a custom light cone built specifically for this purpose by the authors. In particular, the light cone is assembled based on the methodology presented in [21], which uses the semi-analytical model L-Galaxies applied on the Millennium dark matter merger trees. Specifically, the light cone generated for this work is complete up to $z = 1$ and it contains several physical properties such as mass, magnitudes, observed and geometrical redshift for all the galaxies included in it. Given that we are mainly interested in the low redshift Universe, where our results might be affected by cosmic variance caused by narrow angular apertures, we decide to set the light cone aperture to 1/8th of the full sky. We refer the reader to [21] for further details about the light cone construction. As an illustrative example, Fig. 1 displays the spatial distribution of the $M_{\text{gal}} > 10^{10} \text{ M}_\odot$ galaxies[4] inside $\sim$1 deg declination slice.

We fix the parameters of the simulation so that $\text{H}_0 = 73 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.25$, and $\Omega_\Lambda = 0.75$. Even though these values do not reflect the state-of-the-art estimates, this has no impact on the analysis. Our work aims to show the potential of this GW based measurement for a given Universe, which we are allowed to customize within our pipeline.

To locate and populate the error boxes, for each GW event we adopt the following procedure:

(1) We list all the galaxies within our light cone with a cosmological redshift within the $z \pm \sigma_z$ interval consistent with the ET + LISA $d_L \pm \sigma_{d_L}$ measurement and the true (i.e., the Millennium) cosmology.

(2) We randomly[5] select a galaxy among them and we denote it as the "true host" of the GW event. This galaxy position is described by the set $\Theta^{\text{th}} = \{d_L^{\text{th}}, \mu_N^{\text{th}}, \varphi_N^{\text{th}}\}$.

(3) We draw a (μ_N', φ_N') pair of celestial coordinates from the uncertainty distribution in Eq. (4) centered in Θ^{th}.

(4) We consider a 3σ region in $\Delta\Omega$, computed according to Eq. (8) and centered in (μ_N', φ_N'). This procedure ensures that the volume is not artificially centered onto the true host, but in a nearby point in the sky consistent with the sky location error given by the GW measurement.

(5) We select all the galaxies with $M_{\text{gal}} \geq 3 \times 10^{10} \text{ M}_\odot$[6] within $\Delta\Omega \times [z^- - \sigma_{pv}(z^-), z^+ + \sigma_{pv}(z^+)]$ and we associate to each one a hosting probability consistent

---

[4]The mass resolution of the simulation is $M_* = 10^{10} \text{ M}_\odot$.
[5]However, we reject those selected near the boundary of the light cone, so that we avoid cutting the edges of the $\Delta\Omega$ ellipses.
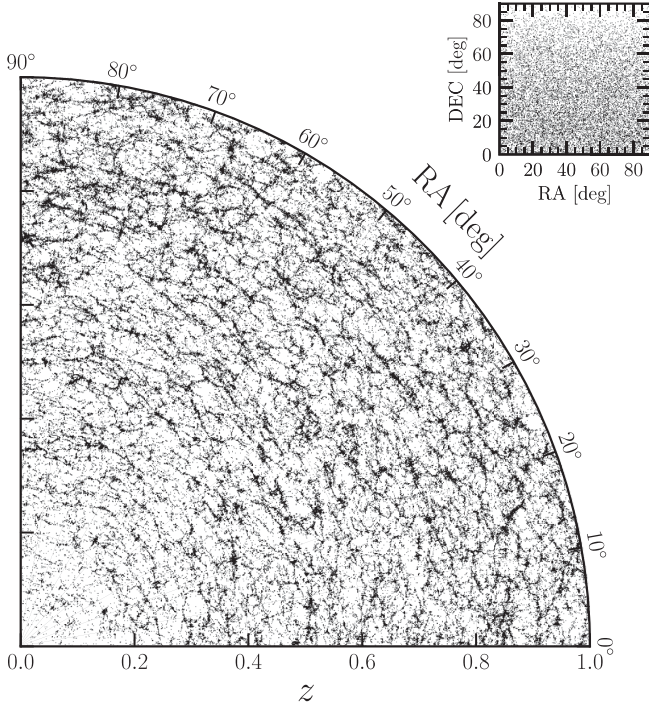[6]See appendix for a discussion on the impact of the adopted mass threshold.

FIG. 1.   Visual representation of a slice of the light cone adopted in this work. The major panel shows the galaxy distribution in the z-RA plane, while the top-right corner plot displays the galaxy distribution in the RA-DEC plane.

with the marginalized sky location error given by the GW measurement (see section II D). These represent all the possible host candidates for the GW event.

The outcome of such a method is displayed in Fig. 2, where we present different error boxes to show the main features and how they may affect the inference of cosmological parameters. In particular, the left column shows an error box with an optimal galaxy clustering in the true cosmology region, thus helping the inference of cosmological parameters. The right column, instead, depicts a noninformative event, due to the misleading information provided by the large galaxy cluster at higher redshift. The middle column is an average error box and it becomes useful to the inference when it is cross correlated with other GW events.

D. Bayesian inference

The problem of measuring quantities in nonrepeatable experiments has a Bayesian nature by definition. The entire framework relies on the Bayes theorem, which states that, given a set of data D and a model hypothesis H, the probability of H given D is

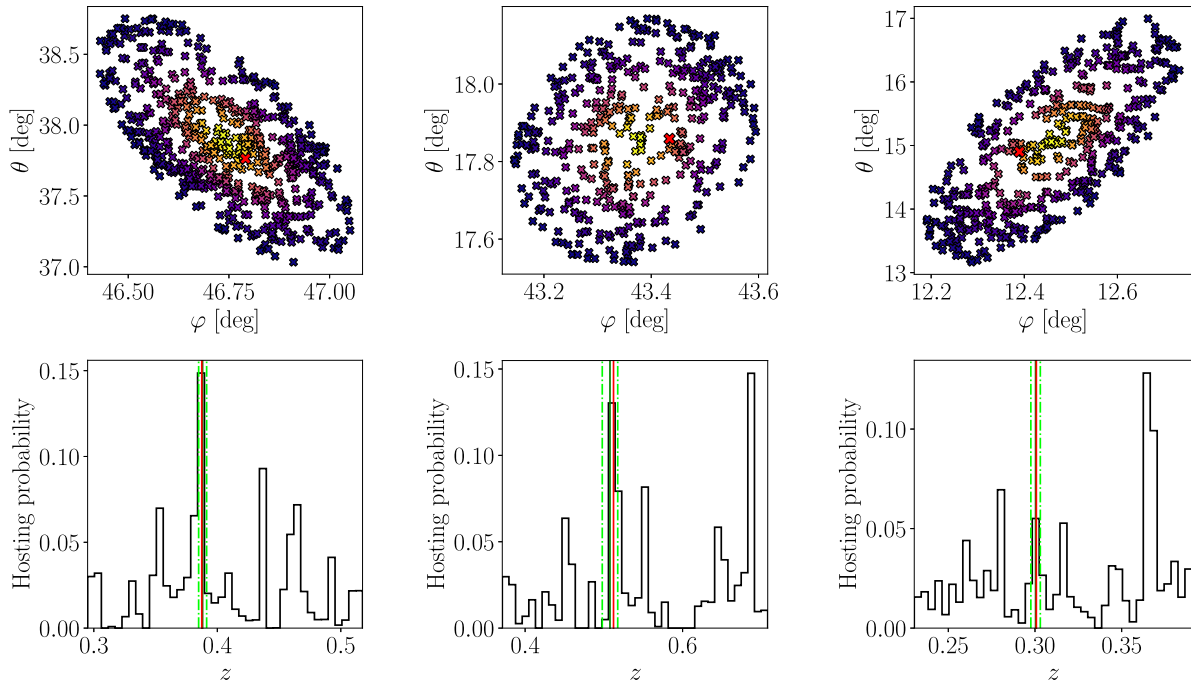$$p(H|D) = \frac{\mathcal{L}(D|H)\pi(H)}{\mathcal{Z}}, \qquad (11)$$

where

FIG. 2.   Three different error boxes are displayed column-wise. The top row depicts the distribution of galaxies in the (θ, φ) plane, and the color scale denotes the associated hosting probability, from dark blue (low) to yellow (high). The histograms on the bottom row represent the galaxy distributions over redshift, weighted on the hosting probability. According to the true cosmology, the dark green solid line denotes the best ET + LISA z measurement, while the light green dotted lines represent the 2σ redshift interval. The red color in each panel denotes the selected "true host".

(i) $p(H|D)$ is the *posterior distribution*, stating the degree of belief in H after the measure.

(ii) $\mathcal{L}(D|H)$ is the *likelihood function*, which is known once we assume a model hypothesis.

(iii) $\pi(H)$ is the *prior distribution*, which amounts to our degree of belief in H before the measure.

(iv) $\mathcal{Z}$ is the *evidence*, a central quantity in model selection studies. Since here we are interested just at the posterior distribution, we can neglect this factor and simply renormalize the posterior at the end of the computation.

In our case, D represents the detected GW events, while H denotes a particular assumed cosmological model that will define the parameter space to be explored. The aim of Bayesian inference is to determine the posterior distribution. To sample the posterior distribution $p(H|D)$ we first need to define the model (which defines the likelihood function $\mathcal{L}$) and the prior distribution.

1. Single GW event likelihood

In the next derivation, we follow the arguments detailed in [12]. Consider a set of n GW events $\mathbf{gw} = \{gw_1, \ldots, gw_n\}$ and let $\mathcal{S} = \{H_0, \Omega_m, \ldots\}$ be a set of cosmological parameters to be inferred. Each GW event is reasonably independent from the others, therefore the likelihood in Eq. (11) can be rewritten as the product of the single GW event likelihoods[7]:

$$\mathcal{L}(\mathbf{gw}|\mathcal{S}) = \prod_{i=1}^{n} \mathcal{L}(gw_i|\mathcal{S}). \qquad (12)$$

The single GW event quasilikelihood is obtained through the marginalization over the source parameters. By defining $\mathbf{x} = \{d_L, z, \theta, \varphi\}$, we can write

$$\mathcal{L}(gw_i|\mathcal{S}) = \int d\mathbf{x}\,\mathcal{L}(gw_i|\mathbf{x}, \mathcal{S})\pi(\mathbf{x}|\mathcal{S}). \qquad (13)$$

As in [12], we assume that the integral over the sky position can be performed analytically. Hence we are left with

$$\mathcal{L}(gw_i|\mathcal{S}) = \int dd_L\,dz\,\mathcal{L}(gw_i|d_L, z, \mathcal{S})\pi(d_L|z, \mathcal{S})\pi(z|\mathcal{S}). \qquad (14)$$

As shown in Eq. (9), a given cosmological model fixes a $d_L - z$ relation. Therefore, among d_L, z, and $\mathcal{S}$, only two of them are independent quantities. We choose d_L to be the dependent one, and the prior on the luminosity distance becomes, as in [12,47], the following Dirac's delta

---

[7]This quantity is more rigorously called *quasilikelihood*, due to the fact that it is obtained through the marginalization of the likelihood over *nuisance* parameters.

$$\pi(d_L|z, \mathcal{S}) = \delta(d_L - d_L(z, \mathcal{S})), \qquad (15)$$

where $d_L(z, \mathcal{S})$ is computed through Eq. (9). Under the Fisher information matrix formalism, the likelihood of a GW event is a Gaussian distribution in luminosity distance, whose mean value $\langle d_L \rangle$ and width σ_{d_L} are given by the matched filtering technique. In addition to the instrument uncertainty of the luminosity distance, we also account for the systematic error due to weak-lensing, as modeled through σ_{WL} (see Eq. (7.3) of [48]). Therefore, once we marginalize over the luminosity distance, we obtain

$$\mathcal{L}(gw_i|d_L(z, \mathcal{S}), z, \mathcal{S}) \propto \exp\left[-\frac{1}{2}\frac{(d_L(z, \mathcal{S}) - \langle d_L \rangle)^2}{\sqrt{\sigma_{d_L}^2 + \sigma_{WL}^2}}\right]. \qquad (16)$$

Finally, as in [12], the prior over the GW event redshift is built in order to take into account for the peculiar velocities of galaxies in the catalog. Each galaxy j is assigned with a hosting probability w_j proportional to the distance in the (θ, φ) plane between the host candidate and the relocated GW event. In particular, this quantity is computed by marginalization of Eq. (4) over the luminosity distance [45]

$$w_j = \int dd_L\, p(\mathbf{\Theta}|s), \qquad (17)$$

with $\mathbf{\Theta} = \{d_L, \mu_N, \varphi_N\}$, and by evaluating it at (μ_{N_j}, ϕ_{N_j}), which are the sky coordinates of the jth galaxy within the error volume of a GW event.

The prior over the GW event redshift is therefore chosen as a discrete sum of K Gaussians, to account for the galaxy redshift uncertainty, each weighted by its w_j value:

$$\pi(z|\mathcal{S}) \propto \sum_{j=1}^{K} w_j \exp\left[-\frac{1}{2}\left(\frac{z - z_j}{\sigma_{pv_j}}\right)^2\right]. \qquad (18)$$

Here j runs over the K galaxies inside the error box, while $\sigma_{pv_j} = \sigma_{pv}(z_j)$. The single GW quasilikelihood then reads

$$\mathcal{L}(gw_i|\mathcal{S}) \propto \int_{z_{min}}^{z_{max}} dz \exp\left[-\frac{1}{2}\frac{(d_L(z, \mathcal{S}) - \langle d_L \rangle)^2}{\sqrt{\sigma_{d_L}^2 + \sigma_{WL}^2}}\right] \\ \times \sum_{j=1}^{N} w_j \exp\left[-\frac{1}{2}\left(\frac{z - z_j}{\sigma_{pv_j}}\right)^2\right], \qquad (19)$$

where z_{min} and z_{max} are the lower and upper integration bounds and correspond to the minimum and maximum GW redshift obtained from the prior on $\mathcal{S}$ and inverting the $d_L - z$ relation in Eq. (9).

2. Prior choices and application

In this work, we highlight the need of an independent estimate of the local universe expansion rate due to the current tension between the most recent estimates from EM surveys [1,2]. We therefore choose to infer a set of two different cosmological parameters which count the Hubble constant[8] and the matter energy density parameter, $\mathcal{S} = \{h, \Omega_m\}$. We choose conservative flat prior distributions for each quantity, in particular $h \in [0.6, 0.86]$ and $\Omega_m \in [0.04, 0.5]$.

The numerical implementation of the method described in Sec. II D is achieved through COSMOLISA, a public software package [49] based on a nested sampling algorithm [50]. The primary output of nested sampling is the evidence $\mathcal{Z}$, producing samples of the posterior distribution as a side product (see [51] for the basics of the method). We explore the parameter space with 5000 live points and evolve, at each iteration, the lowest-likelihood one through a *Markov Chain Monte Carlo* (MCMC) until $\mathcal{Z}$ is computed at a given accuracy level.

III. RESULTS AND DISCUSSION

A. Multiband approach

We analysed the results of the parameter estimation for the three binary subpopulations (below-gap, across-gap, above-gap) and the four scenarios (mSFR-mZ, mSFR-sZ, sSFR-mZ, sSFR-sZ). As Table II shows, however, not all the cases are suitable multiband candidates for the inference of cosmological parameters. For instance, across-gap binaries are extremely rare systems regardless of the scenario, while below-gap binaries reach interesting numbers only assuming 10 years of observation and the most optimistic models. As a matter of fact, above-gap binaries are the most promising ones, in particular under the sSFR-sZ scenario. Even if the corresponding intermediate models are still encouraging, in this work we focus on the optimistic one, as we need to impose other cuts on the catalog that further shrink the sample of useful systems.

The first remarkable results are the high performances of the cooperation between space-born and third generation ground-based interferometers. In Fig. 3 we report the uncertainty distributions for d_L and $\Delta\Omega$ for the above-gap and the sSFR-sZ model, assuming 10 years of LISA mission time. We consider LISA, ET, and LIGO-Virgo network at design sensitivity. Since the latter does not yield sufficient accuracy in parameter estimation to perform meaningful cosmological measurements, we focus on ET and LISA. Individually, both detectors hardly achieve the precision reached through their joint exploitation: for the sky location, the median of the ET + LISA distribution, in fact, decreases by an order of magnitude with respect to the

---

[8]We define $h = H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$, so that it is dimensionless and smaller than unity, in our case $h = 0.73$.

TABLE II. A showcase of the number of LISA's detection for a given population of binaries and model, assuming 4 and 10 years of mission lifetimes. These numbers represent one over the 30 realizations. Due to ET sensitivity curve, all binaries detected by LISA are also detected at ground.

		mSFR-mZ	mSFR-sZ	sSFR-mZ	sSFR-sZ
4 years	Below-gap	1.6	1.5	2.1	2.1
	Across-gap	0.1	0.3	0.2	0.7
	Above-gap	1.1	7.7	8.1	40.8
10 years	Below-gap	5.1	5.5	7.3	8.9
	Across-gap	0.2	0.7	0.5	3.1
	Above-gap	4.2	28.4	27.8	134.1

single detector ones, with LISA being slightly better in the determination of this parameter. The same improvement concerns the luminosity distance, with ET being the best probe to measure this quantity, as expected.

In Table III we show the medians of the uncertainty distributions for each parameter in the single detectors and in the network ET + LISA. Crucially for this work, ET + LISA yields an improvement factor of 100 in sky localization and of 10 in luminosity distance measurement precision. It should be noticed that the combination of the two also improves the estimate of all other parameters, most noticeably spin magnitudes.
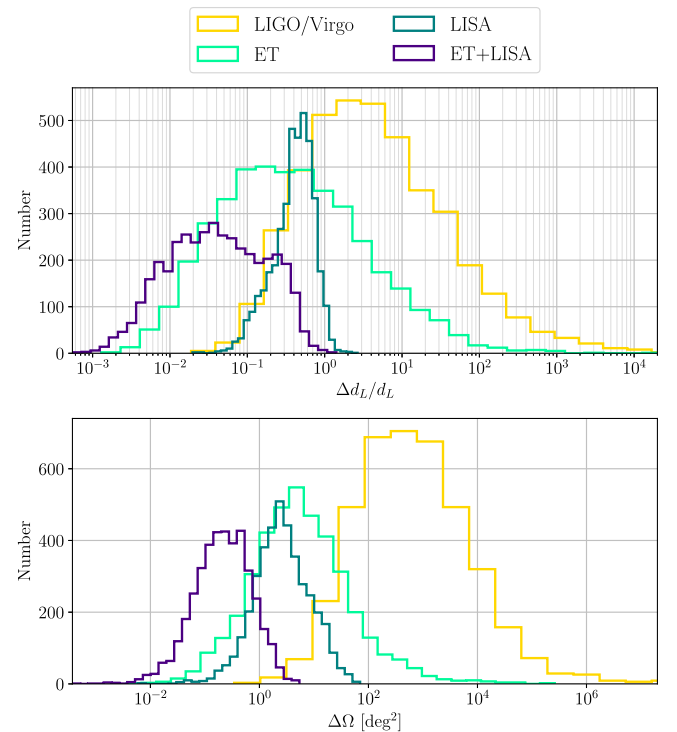

FIG. 3. Uncertainty distributions for the luminosity distance (upper panel) and the sky localization (bottom panel) in LIGO/Virgo, ET, LISA and the network ET + LISA, colors as in legend. The panels display the results assuming 10 years LISA mission time and the above-gap, sSFR-sZ scenario.

TABLE III. Medians of the uncertainty distributions of each parameter for LIGO/Virgo, ET, LISA and the network ET + LISA, assuming 10 years LISA mission time. Here $\mathcal{M}$ represents the *chirp mass* of the system, defined as $\mathcal{M} = (M_1 M_2)^{(3/5)}/(M_1 + M_2)^{(1/5)}$ and its uncertainty is obtained through error propagation formulas.

	LIGO/Virgo	ET	LISA	ET + LISA
$\Delta\Omega[\mathrm{deg}^2]$	5.6×10^2	5.1	2.5	2.2×10^{-1}
$\Delta d_L/d_L$	3.6	3.4×10^{-1}	4.3×10^{-1}	3.9×10^{-2}
$\Delta\mathcal{M}/\mathcal{M}$	3.6×10^{-1}	3.3×10^{-3}	4.1×10^{-7}	2.0×10^{-7}
$\Delta\iota/\iota$	5.9	4.7×10^{-1}	9.3×10^{-1}	6.5×10^{-2}
$\Delta\chi_1/\chi_1$	3.9	2.8×10^{-2}	3.3×10^{-1}	8.6×10^{-4}
$\Delta\chi_2/\chi_2$	3.8	3.0×10^{-2}	4.6×10^{-1}	1.4×10^{-3}

B. Inference of cosmological parameters

Among the many standard siren candidates, we impose a few cuts on the original binary catalogs. First, we request that the luminosity distance of the event is determined within 10% of precision. Then we select only the events localized with $\Delta\Omega < 1$ deg^2.[9] Moreover, due to the limited light cone extension, we request that the maximum redshift of the error box when varying the cosmology is smaller than 1. Considering 4 (10) years of LISA observations, in 30 independent realizations of the Universe, from a total of 1223 (4023) binaries we are left with 222 (510) GW events, as reported in Table IV.

1. Preliminary tests

Since the analysis over a large number of events can be computationally costly, we start by exploring a small fraction of the total 10-year catalog of events. We measure h and Ω_m by considering the closest GW events first, adding progressively farther events to the inference.

Even though the results are consistent with the true cosmology, we find estimates that are systematically biased toward large values of h and small values of Ω_m, due to the correlation between the parameters. However, as the number of events increases, this effect is mitigated and eventually disappears, as the algorithm returns well centered Gaussian posterior distributions and the accuracy of the h (Ω_m) measures evolves from 0.63σ (0.28σ) to 0.29σ (0.24σ). To understand whether this behavior is due to low redshift events only, we focus on these events in the following discussion.

The inference relies on the determination of the posterior distribution of each GW event redshift. If this step produces misleading information, the bias is propagated to the estimates of the cosmological parameters. To leading order,

---

[9]The likelihood computation becomes too expensive for GW events with more than 10^4 hosts. Those events are generally poorly localized and their large 3D volumes result in a rather uninformative host redshift distribution.

TABLE IV. Number of events expected before and after the requirements imposed on the above-gap/sSFR-sZ binary catalog, assuming both LISA mission lifetimes. The rightmost column shows the average number of events per realization once we performed the selections.

Number of events	Before cuts	After cuts	Per realization
4 years	1223	222	∼7
10 years	4023	510	∼17

if the observed redshift is underestimated (overestimated), the h and Ω_m posteriors will be biased toward low (large) and high (low) values, respectively. The accuracy of our results suggests that the closer the GW event is, the more the observed redshift is overestimated. In light of these considerations, peculiar velocities may play a role in altering significantly the apparent redshift of close objects compared to the cosmological one.

As already mentioned in Sec. II C, the galaxy catalog comes with both the geometrical and the observed redshift of each object, thus allowing us to assess the impact of the peculiar velocity on the analysis. Through the same aforementioned procedure, we therefore infer the cosmological parameters in a Universe without peculiar velocities. The results are shown in the left panel of Fig. 4, where we directly compare the estimates from a static (z_{cosmo}) and dynamic (z_{obs}) Universe as a function of the number of closest GW events. As we can see, the bias completely vanishes already with 10 events, and as the number of mergers increases, the two analyses lead to the true cosmology.

Thus, peculiar velocities play a major role in the inference of cosmological parameters with low-redshift events. In fact, the right panel of Fig. 4 shows that the v_p distribution of our light cone galaxies is not symmetrically distributed around zero at low redshifts. The quantity v_p is computed along the radial direction with positive values corresponding to a drift away from the observer. There is a clear preference for positive v_p values for galaxies at $0.05 < z < 0.1$, which is where most of the closest GW events in our catalogs occur, thus explaining the bias.

It is interesting that our light cone features a large portion of galaxies that move preferentially away from the observer. One possible cause is the limited solid angle covered by our galaxy catalog. Although a full sky would mitigate this issue, its generation would be extremely computationally expensive. Here we just notice that the bias in the parameters was found considering the closest events of the full GW source catalog, featuring 30 realizations of the experiment. In a single realization there will be perhaps only one such close source. In fact, as we will see in Sec. III B 2, this bias does not systematically appear in this case. Moreover, in a real experiment, one can in principle further mitigate any issue related to peculiar velocities by modelling the bulk motions as a function
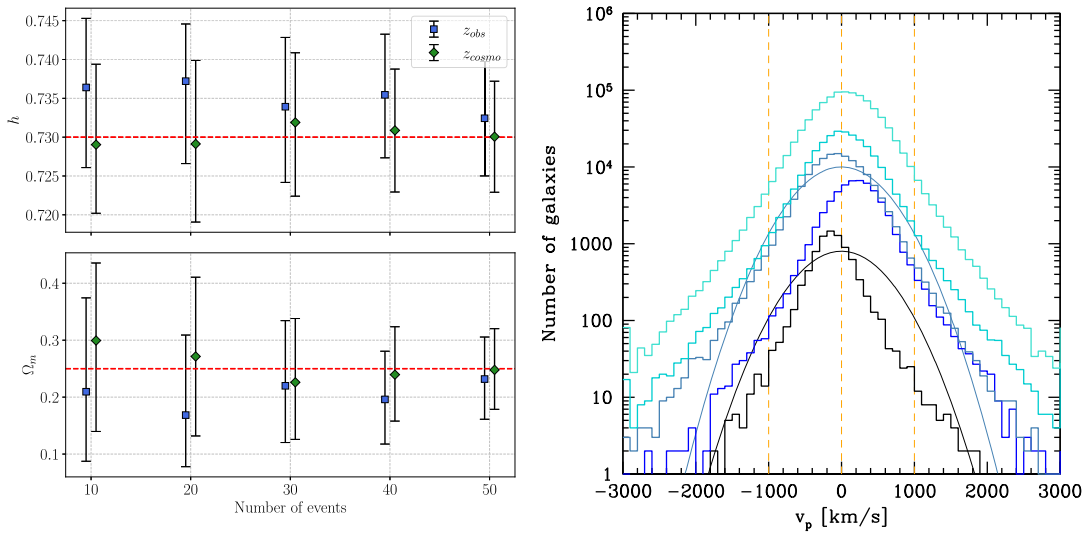
FIG. 4. Left panel: cosmological parameter measurements as a function of the number of closest GW events used for the inference assuming a Universe with and without peculiar velocities (labeled z_{obs} and z_{cosmo}, respectively). We investigate the results from a Universe with and without peculiar velocities, colors as in legend. The top plot represents the results for h, while the bottom plot displays the Ω_m estimates (median values with error bars representing the 1σ confidence level). The red dashed line represents the true value of each parameter. Right panel: distribution of peculiar velocities in our light cone for several redshift bin. Histograms from bottom to top are shown for incremental 0.05 redshift bins, starting from [0, 0.05]. To guide the eye, solid lines represent Gaussians with standard deviation $\sigma_{v_P} = 500$ km s^{-1}, as assumed in our study.

of redshift and update the inference model to keep into account for any local anisotropy.

2. Individual realizations of the full experiment

The nominal LISA mission lifetime is set to be 4 years with a possible operation extension up to 10 years [13]. To produce realistic samples of GW events for a given mission time, we divide the original catalog in independent subsets producing:

  (i) 30 realizations of either 7 or 8 events each for 4 years of observations;
  (ii) 30 realizations of 17 events each for 10 years of observations.

The sizes of both samples are easily manageable from the computational point of view. We therefore consider all individual realizations and then average the posterior distributions from the 30 subsets to characterize the accuracy and the precision of this method.

The results for the inference of h and Ω_m from all the runs are presented in Fig. 5 and Table V. We observe that the bias due to peculiar velocities does not pose a significant issue in individual realizations of the experiment because of the very small number of low redshift events. In the vast majority of the cases, the inference yields measurements of h that are both precise and accurate. By looking at Table V we can in fact appreciate that h is generally measured to better than 2% (68% credible region) and often around 1% in the 10-year case, and the true value is generally well within the 68% credible region implying

no significant biases. Ω_m is far less constrained, generally at a 20–30% level.

There are, however, a couple of realizations that yield severely inconsistent results. These outliers can be identified in realization 21 in the 4-year case and realizations 18 in the 10-year case (i.e., in 1 of our 30 realizations). In these realizations, both h and Ω_m rail against the upper end of the prior range, a behavior that is not explained by the expected bias due to peculiar velocities. The effect of this can also be seen in the averaged posteriors shown in the right panels of Fig. 5. In the 4-year case the averaged marginalized posteriors of h and Ω_m display long tails extending to the upper bound of the priors, whereas in the 10-year case this effect is largely suppressed. However, in this case a mild secondary peak appears at the boundary of the prior and leads to a slightly asymmetric distribution. Even if the averaged estimates that we obtain are largely consistent with the true cosmology, we want to focus our attention on the physical meanings (if any) of the small deviations from the expected results.

3. Origin of the problematic realizations

We extensively examined the bad 10-year realization and did not find any specific pathology in the error box construction, nor in the inference procedure. The culprit of the bias appears to lie in the relation between H_0, Ω_m and redshift that enters the single-event likelihood Eq. (19), which is proportional to the number of galaxies at a given redshift. Therefore, for a comoving-volume-uniform galaxy distribution, the weight is roughly proportional to
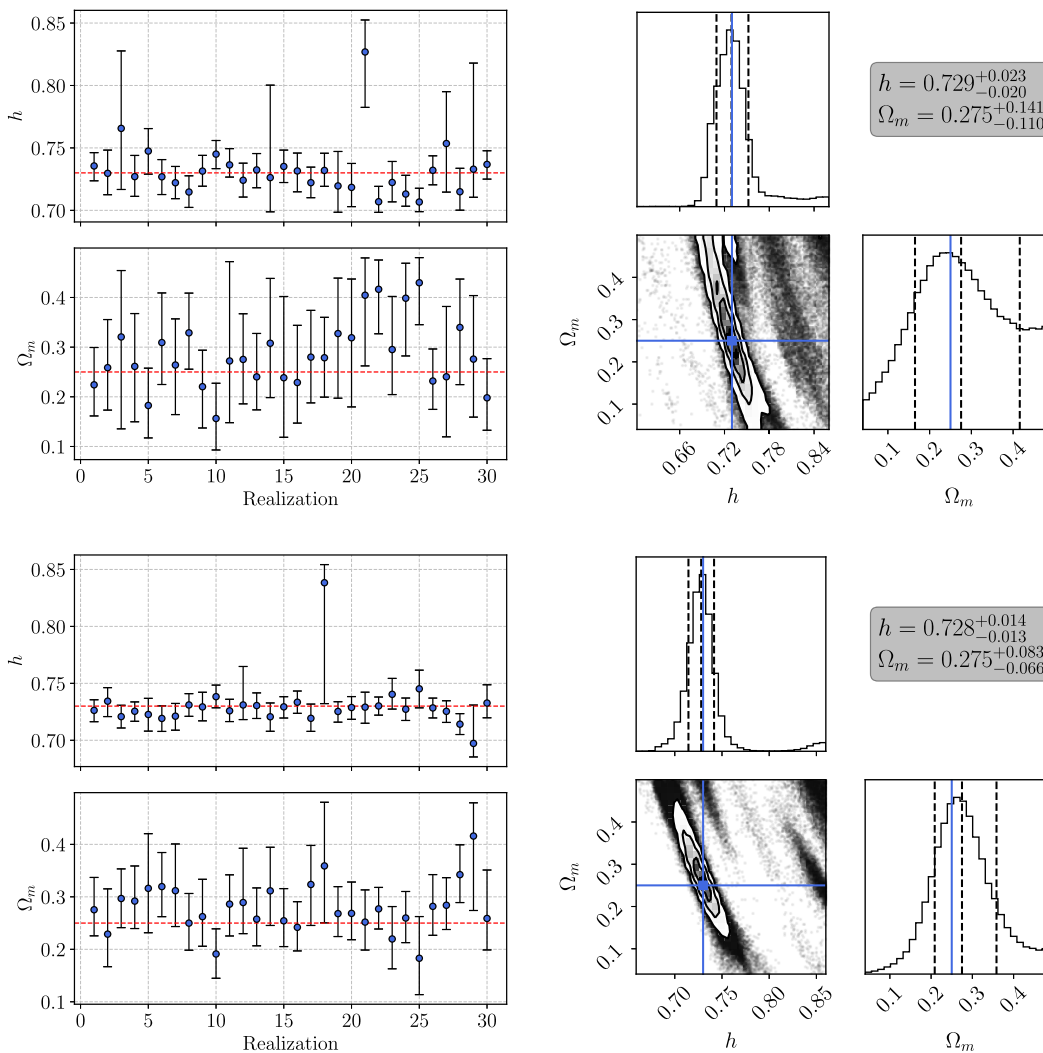
FIG. 5.   Inference results from the 30 independent realizations at the 68% confidence level. The top row refers to 4 years of LISA observations, while the bottom row refers to 10 years. Plots in the left column, the blue dots denote the median of the posterior distributions, while the red dashed lines represent the true value of each parameter. The right column displays the joint posterior distributions averaged on the 30 independent realizations. Here the gray color scale distinguishes low probability regions (light) with high probability regions (dark); the black dashed lines mark, from left to right, the 16, 50, and 84 percentile and finally the blue lines highlight the true cosmology. Plots in the right column have been made using [52].

the comoving volume shell, naturally favoring values of redshift approaching z_{max}. If the information enclosed in the clustering is not strong enough, this "high z" solution competes with the correct one and can eventually dominate. This issue is further exacerbated by the $H_0 - \Omega_m$ degeneracy. As seen from the 2D posteriors shown in Fig. 5, the two cosmological parameters are partially degenerate, showing a clear anticorrelation. Therefore, a given (d_L, z) pair can be produced by a continuum of (H_0, Ω_m) following this degeneracy. For a rectangular uniform prior in the cosmological parameters, while galaxies in the middle of the error box are consistent with a degenerate set of cosmologies, those at the boundaries (z_{min} and z_{max}) are only consistent with the corners of the prior. Since many more galaxies accumulate toward z_{max},

this creates an artificial spike at the top right corner of the parameter space. So in absence of events with zero support around z_{max}, this second mode of the posterior cannot be suppressed and would eventually dominate, essentially by design of the likelihood function.

The above interpretation is corroborated by two tests that we now discuss. The first test we performed was to artificially add to the bad 10-year realization a very informative event (VIE), i.e., an event with no (or little) support at redshifts outside those allowed by the true cosmology, in particular with zero support around z_{max}. The result of this procedure is shown in Fig. 6. It can be seen that the addition of a VIE kills the high z solution and allows the recovery of the correct cosmology, albeit with large uncertainties on Ω_m.

TABLE V. The h and Ω_m estimates for all the inference runs. The columns report: the estimates at the 68% confidence level (1σ), the precision of the measure (%), and the estimate accuracy as a fraction of σ (A).

LISA mission time	Realization	GW events	h			Ω_m		
			1σ	%	A	1σ	%	A
4 years	1	8	$0.736^{+0.011}_{-0.012}$	1.5	0.5σ	$0.224^{+0.075}_{-0.063}$	30.74	0.4σ
	2	8	$0.730^{+0.019}_{-0.017}$	2.5	$<0.1\sigma$	$0.259^{+0.097}_{-0.085}$	35.3	0.1σ
	3	8	$0.766^{+0.062}_{-0.049}$	7.3	0.6σ	$0.321^{+0.134}_{-0.185}$	49.7	0.4σ
	4	8	$0.727^{+0.017}_{-0.016}$	2.3	0.2σ	$0.261^{+0.106}_{-0.112}$	41.7	0.1σ
	5	8	$0.747^{+0.018}_{-0.018}$	2.4	1.0σ	$0.183^{+0.075}_{-0.065}$	38.5	1.0σ
	6	8	$0.727^{+0.014}_{-0.014}$	1.9	0.2σ	$0.309^{+0.100}_{-0.084}$	29.8	0.6σ
	7	8	$0.722^{+0.013}_{-0.013}$	1.8	0.6σ	$0.264^{+0.093}_{-0.100}$	36.5	0.1σ
	8	8	$0.715^{+0.013}_{-0.012}$	1.8	1.2σ	$0.329^{+0.080}_{-0.073}$	23.3	1.0σ
	9	8	$0.731^{+0.013}_{-0.012}$	1.7	0.1σ	$0.220^{+0.073}_{-0.083}$	35.5	0.4σ
	10	8	$0.745^{+0.011}_{-0.012}$	1.5	1.3σ	$0.156^{+0.071}_{-0.063}$	42.9	1.4σ
	11	8	$0.736^{+0.013}_{-0.010}$	1.6	0.6σ	$0.272^{+0.200}_{-0.124}$	59.57	0.1σ
	12	8	$0.724^{+0.014}_{-0.013}$	1.9	0.4σ	$0.275^{+0.092}_{-0.089}$	32.9	0.3σ
	13	7	$0.732^{+0.013}_{-0.014}$	1.9	0.2σ	$0.240^{+0.087}_{-0.066}$	32.0	0.1σ
	14	7	$0.726^{+0.074}_{-0.027}$	7.0	0.1σ	$0.308^{+0.131}_{-0.109}$	39.0	0.5σ
	15	7	$0.735^{+0.013}_{-0.013}$	1.8	0.4σ	$0.238^{+0.164}_{-0.120}$	59.5	0.1σ
	16	7	$0.732^{+0.014}_{-0.017}$	2.1	0.1σ	$0.229^{+0.115}_{-0.082}$	43.0	0.2σ
	17	7	$0.722^{+0.012}_{-0.012}$	1.7	0.6σ	$0.280^{+0.094}_{-0.092}$	33.3	0.3σ
	18	7	$0.732^{+0.014}_{-0.013}$	1.8	0.1σ	$0.279^{+0.081}_{-0.079}$	29.0	0.4σ
	19	7	$0.720^{+0.028}_{-0.021}$	3.4	0.4σ	$0.327^{+0.112}_{-0.130}$	36.9	0.6σ
	20	7	$0.718^{+0.019}_{-0.015}$	2.4	0.7σ	$0.319^{+0.118}_{-0.139}$	40.4	0.5σ
	21	7	$0.827^{+0.026}_{-0.044}$	4.2	2.8σ	$0.405^{+0.075}_{-0.142}$	26.9	1.4σ
	22	7	$0.707^{+0.012}_{-0.008}$	1.5	2.2σ	$0.416^{+0.059}_{-0.090}$	17.8	2.2σ
	23	7	$0.722^{+0.017}_{-0.016}$	2.2	0.5σ	$0.295^{+0.107}_{-0.091}$	33.5	0.5σ
	24	7	$0.713^{+0.015}_{-0.010}$	1.7	1.4σ	$0.398^{+0.0.70}_{-0.116}$	23.4	1.6σ
	25	7	$0.707^{+0.011}_{-0.008}$	1.3	2.5σ	$0.430^{+0.051}_{-0.084}$	15.7	2.7σ
	26	7	$0.732^{+0.012}_{-0.012}$	1.6	0.2σ	$0.232^{+0.065}_{-0.057}$	26.2	0.3σ
	27	7	$0.754^{+0.042}_{-0.039}$	5.3	0.6σ	$0.240^{+0.142}_{-0.121}$	54.6	0.1σ
	28	7	$0.715^{+0.019}_{-0.015}$	2.4.	0.9σ	$0.340^{+0.097}_{-0.115}$	31.3	0.8σ
	29	7	$0.733^{+0.085}_{-0.022}$	7.3	0.1σ	$0.276^{+0.128}_{-0.117}$	44.4	0.2σ
	30	7	$0.737^{+0.011}_{-0.012}$	1.5	0.6σ	$0.198^{+0.079}_{-0.065}$	36.4	0.7σ
10 years	1	17	$0.726^{+0.009}_{-0.010}$	1.3	0.4σ	$0.275^{+0.062}_{-0.050}$	22.2	0.5σ
	2	17	$0.734^{+0.012}_{-0.014}$	1.7	0.3σ	$0.229^{+0.086}_{-0.062}$	32.4	0.3σ
	3	17	$0.721^{+0.010}_{-0.010}$	1.4	0.9σ	$0.297^{+0.056}_{-0.055}$	18.9	0.8σ
	4	17	$0.725^{+0.008}_{-0.009}$	1.2	0.5σ	$0.292^{+0.067}_{-0.052}$	20.4	0.7σ
	5	17	$0.723^{+0.014}_{-0.015}$	2.0	0.5σ	$0.316^{+0.104}_{-0.085}$	28.9	0.7σ
	6	17	$0.719^{+0.011}_{-0.011}$	1.6	1.0σ	$0.320^{+0.065}_{-0.057}$	19.1	1.1σ
	7	17	$0.721^{+0.011}_{-0.012}$	1.6	0.8σ	$0.312^{+0.089}_{-0.068}$	25.3	0.8σ
	8	17	$0.731^{+0.010}_{-0.010}$	1.7	0.1σ	$0.250^{+0.056}_{-0.052}$	21.6	$<0.1\sigma$
	9	17	$0.720^{+0.013}_{-0.012}$	1.7	0.1σ	$0.262^{+0.071}_{-0.056}$	24.3	0.2σ
	10	17	$0.738^{+0.010}_{-0.010}$	1.4	0.8σ	$0.191^{+0.048}_{-0.046}$	24.6	1.3σ
	11	17	$0.726^{+0.010}_{-0.009}$	1.4	0.4σ	$0.286^{+0.056}_{-0.061}$	20.4	0.6σ
	12	17	$0.731^{+0.034}_{-0.013}$	3.2	$<0.1\sigma$	$0.289^{+0.103}_{-0.059}$	28.1	0.5σ
	13	17	$0.730^{+0.011}_{-0.011}$	1.5	$<0.1\sigma$	$0.257^{+0.060}_{-0.051}$	21.4	0.1σ
	14	17	$0.721^{+0.012}_{-0.013}$	1.7	$0.8\sigma.$	$0.311^{+0.083}_{-0.066}$	23.9	0.8σ
	15	17	$0.729^{+0.009}_{-0.010}$	1.3	0.1σ	$0.254^{+0.061}_{-0.049}$	21.7	0.1σ
	16	17	$0.733^{+0.010}_{-0.010}$	1.3	0.3σ	$0.242^{+0.049}_{-0.045}$	19.4	0.2σ
	17	17	$0.719^{+0.012}_{-0.011}$	1.7	0.9σ	$0.324^{+0.074}_{-0.078}$	23.6	1.0σ

(Table continued)

TABLE V. *(Continued)*

LISA mission time	Realization	GW events	h			Ω_m		
			1σ	%	A	1σ	%	A
	18	17	$0.838^{+0.016}_{-0.106}$	7.3	1.8σ	$0.359^{+0.121}_{-0.109}$	32.0	0.9σ
	19	17	$0.725^{+0.009}_{-0.009}$	1.2	0.5σ	$0.268^{+0.051}_{-0.044}$	17.7	0.4σ
	20	17	$0.729^{+0.010}_{-0.010}$	1.4	0.1σ	$0.268^{+0.060}_{-0.050}$	20.5	0.3σ
	21	17	$0.729^{+0.013}_{-0.014}$	1.9	0.1σ	$0.252^{+0.061}_{-0.053}$	22.8	$<0.1\sigma$
	22	17	$0.730^{+0.008}_{-0.008}$	1.1	$<0.1\sigma$	$0.277^{+0.041}_{-0.038}$	14.3	0.7σ
	23	17	$0.740^{+0.014}_{-0.013}$	1.8	0.8σ	$0.220^{+0.062}_{-0.057}$	27.0	0.5σ
	24	17	$0.727^{+0.010}_{-0.010}$	1.4	0.3σ	$0.260^{+0.051}_{-0.047}$	18.8	0.2σ
	25	17	$0.745^{+0.016}_{-0.017}$	2.2	0.9σ	$0.183^{+0.079}_{-0.069}$	40.7	0.9σ
	26	17	$0.728^{+0.009}_{-0.009}$	1.2	0.2σ	$0.282^{+0.061}_{-0.055}$	20.5	0.6σ
	27	17	$0.725^{+0.009}_{-0.010}$	1.3	0.5σ	$0.284^{+0.053}_{-0.046}$	17.4	0.7σ
	28	17	$0.714^{+0.009}_{-0.009}$	1.3	1.7σ	$0.342^{+0.057}_{-0.053}$	16.1	1.7σ
	29	17	$0.697^{+0.034}_{-0.012}$	3.3	1.4σ	$0.416^{+0.063}_{-0.142}$	24.7	1.6σ
	30	17	$0.733^{+0.016}_{-0.013}$	2.0	0.2σ	$0.259^{+0.092}_{-0.060}$	29.4	0.1σ

A second test consisted in changing the prior range of the analysis. If our interpretation is correct, the second mode of the solution appearing in Fig. 5 should follow the boundary of the prior. Figure 7 shows the average posterior over 30 realizations (different from those used to produce Fig. 5) of the experiment with a modified prior range $h \in [0.6, 0.95]$ and $\Omega_m \in [0.04, 0.4]$. Clearly, the secondary mode follows the boundary of the prior, while the correct solution is consistently recovered regardless of the prior range.

4. Mitigation techniques and future investigations

Although we are focusing on the "bad outcomes" of our analysis, it is worth keeping in mind that they involve only a minority of realizations of the experiment. It is nevertheless important to be able to treat these potential issues should they manifest in a future real analysis on actual data.

Even without applying any change to the inference procedure, the tests presented above provide practical ways to reject spurious solutions and correctly constrain cosmological parameters. Bad realizations generally display bimodal posteriors. Repetition of the analysis with a varying prior range should return a "steady mode" around the correct solution and a dominant spurious peak following the boundary. Moreover, events can be analyzed individually and cosmological inference can be done progressively by adding them one by one. In order to exclude peculiar velocities' side effects, we perform this test by considering a steady Universe. As we show in Fig. 8 for realization 18, we found that in our two bad realizations,

$$h = 0.731^{+0.016}_{-0.013}$$
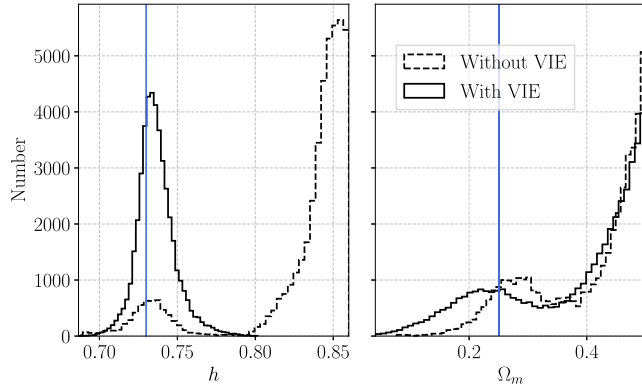$$\Omega_m = 0.265^{+0.080}_{-0.072}$$

FIG. 6. The impact of VIE on the h and Ω_m posterior distributions of realization 18 (10 years). The blue solid lines mark the true cosmology.
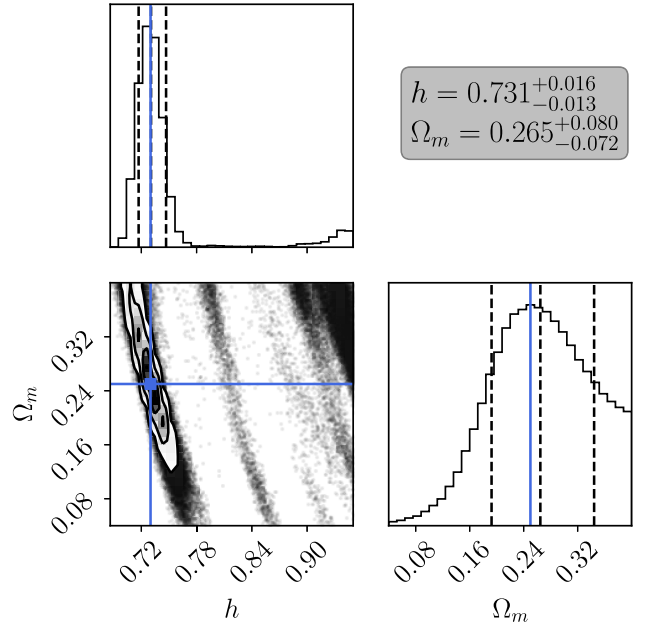
FIG. 7. Averaged posterior distribution over 30 different realizations assuming different prior ranges on h and Ω_m, as discussed in Sec. III B 3.
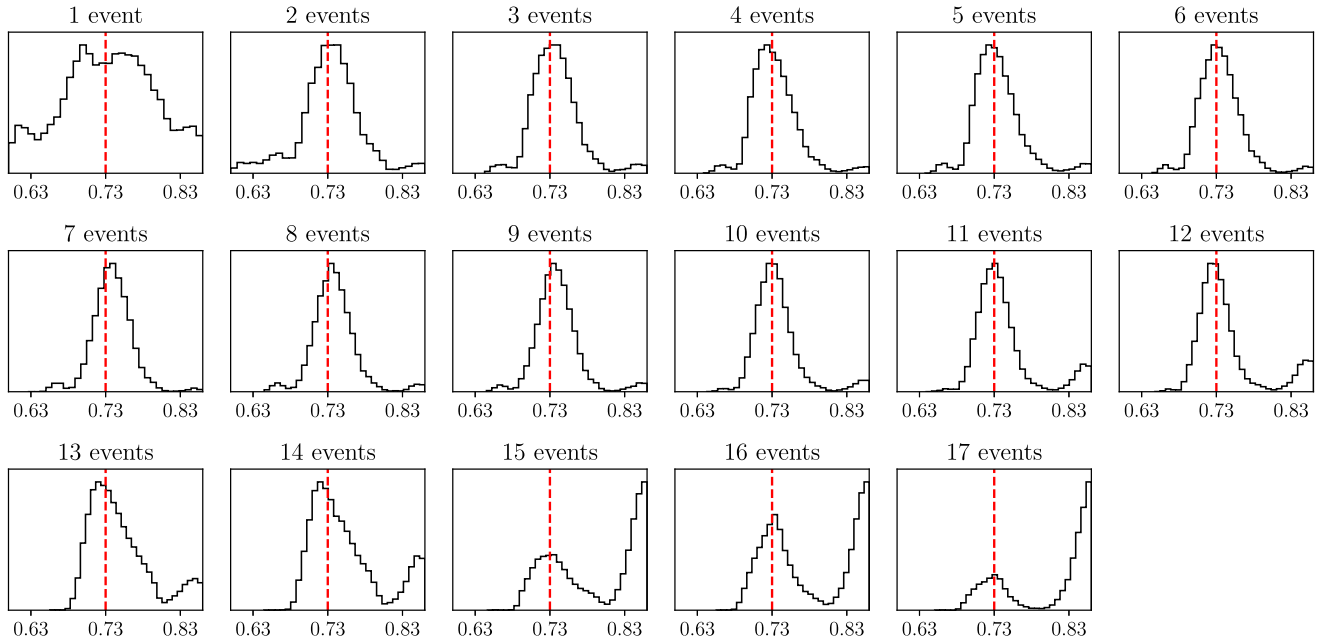
FIG. 8. Evolution of the posterior distribution of the Hubble constant from the realization 18 (10-year) as the number of events increases by one each time, as the panel titles show. Here the red dotted line represents the true value, i.e., $h = 0.73$. In this particular case, the spurious solution overcomes the right cosmology when the last few events are added to the inference.

when this procedure is employed, the joint posterior initially builds up around the correct solution due to the clustering information. This mode, however, is eventually superseded by the spurious one in the long run, if there are no events without host galaxy support around z_{max}. These two checks (varying prior and incremental analysis) could allow us to reject the spurious solution and identify the correct cosmological parameters.

It would be obviously desirable to develop an analysis that prevents side effects like the ones identified here. For example one can reweight or change the shape of the prior, so that an uninformative experiment returns a flat posterior on the parameters, regardless of the fact that the assumption of a cosmological model intrinsically brings some information on the correlation structure of the parameters. Another aspect that should be accounted for concerns the approach adopted to place each GW event in the Millennium Universe, that we discussed in Sec. II C. Within our work, the binary population and the galaxy catalog are independent entities, i.e., binaries merge at points in space where there may not be any galaxy at all: to address the problem, we randomly selected the "true host" within the redshift interval associated to the true cosmology consistent with the d_L uncertainty of the GW measurement. This ensures that an actual galaxy can be referred as the host of the GW event. The random selection should ensure that galaxies in denser regions are more likely to be drawn rather than others. However, the high performance of the multiband approach squeezes the pool of true host candidates to a narrow redshift interval, which often lacks of relevant clustering properties. Therefore, when we extend

the redshift interval to take into account for the prior ranges on the cosmological parameters, we are likely to bring into the error box much denser regions that were artificially excluded from the host selection. In practice, one should start from a given light cone and place GW events randomly within it, to ensure that the distribution of events traces the 3D clustering of galaxies. We plan to explore different experiment designs and likelihood forms in future work.

Besides these adjustments, we remark once again that the analysis performed here allows a robust determination of H_0 within 1%–2% in the vast majority of the cases. Moreover, even when the inference is biased, the spurious solution can be identified and the correct cosmology recovered.

IV. CONCLUSIONS

In this paper we explored the possibility of exploiting multiband GW astronomy to use SBHBs above the pair-instability mass gap as effective dark standard sirens.

Massive SBHBs forming from progenitor stars above the pair-instability mass gap are in fact anticipated to be loud multiband sources, detectable both by LISA and 3G detectors, for which we considered ET as an example. By combining observations in the two bands, the source 3D sky localization can be pinned down to an accuracy which is far better (by three orders of magnitude, on average) than the individual probes alone. This allows an efficient probabilistic identification of the host among all galaxies within the error volume, enabling statistical inference on the cosmological parameters.

We exploited this idea to constrain the Hubble constant H_0 and matter density fraction Ω_m under the assumption of a flat, ΛCDM Universe. We performed the parameter estimation in each detector under the Fisher information matrix formalism, and then combined the relative uncertainties in order to produce accurate estimates of the source parameters, specifically the 3D localization error volume in the sky. We then relied on the Millennium simulation to simulate the galaxy distribution across 1/8th of the sky up to $z = 1$. We placed the 3D volumes in this synthetic sky, thus creating for each of them a sample of host candidates consistent with the clustering properties of the ΛCDM Universe. Finally, we used these data to perform Bayesian inference on H_0 and Ω_m employing a nested sampling algorithm.

By analyzing a large catalog of GW events, we found that peculiar velocities might be a source of systematic errors, depending on the direction of the bulk motion of galaxies. This behavior, however, affects mostly low redshift events, and progressively vanishes when adding farther away mergers to the inference. Moreover, it can be due to limited solid angle coverage of the light cone adopted, and a full sky catalog may address the problem in the first place.

We then performed 30 realizations of the sample of observed SBHBs assuming either 4 or 10 years of LISA operations. By analyzing them, we found that multiband observations of "above-gap" SBHBs can provide a competitive measurement of the cosmological parameters $\mathcal{S} = \{h, \Omega_m\}$. In fact, assuming 4 (10) years of observation, the Hubble constant is determined down to a 1.5% (1.1%) precision, while Ω_m is measured at a 26.2% (14.3%) level. In general, the two parameters are estimated to better than $\sim 2\%$ and $\sim 30\%$ respectively.

We found, however, a couple of realizations yielding biased solutions, favoring h and Ω_m values railing against the upper end of the prior range. We traced back the insurgence of those solutions to a combination of factors, including the form of the likelihood, the $h - \Omega_m$ degeneracy and the lack of very informative events in those realizations. We notice here that this problem did not appear in the work of [12,45], who employed the same techniques. This is likely because [12] considered "below-gap" SBHBs at $z < 0.1$; at such low redshifts, the number of galaxies in the error volume is generally much smaller, and it is unlikely that all the events have support around $z_{\max}$. On the other hand, [45] investigated a limited number of realizations of their experiment, perhaps insufficient to identify a bad one. By individually analyzing those "bad realization" we provided practical ways to recognize and discard spurious solutions, allowing the recovery of the correct cosmological parameters even in their presence.

Nevertheless, the identification of this issue calls for future improvements. Among them, we underline the need to upgrade the design of the simulation so that the binaries from our population merge in crowded clusters in the very first place, as reality suggests. This can be achieved by consistently assigning a host to each GW event before the Fisher matrix pipeline, thus avoiding the need of an artificial true host selection. We plan to enhance hosting probability assignments, so to take into account also for other important parameters besides the host candidate's sky position (e.g., the mass of the galaxy). Furthermore, a reweight or change of the shape of the prior can be folded into the analysis, ensuring that an uninformative experiment returns a flat posterior on the cosmological parameters. We defer a detailed investigation of these possible improvements of the analysis to future work.

Finally, our work relies on the assumption that SBHBs above the mass gap come only from the isolated evolution channel. However BHs inside and above the mass gap can be formed in dense environment thanks to the close interplay between dynamics and stellar evolution [53,54]. These additional sources would increase the number of detected systems that could be exploited in our approach, allowing us to constrain even further the cosmological parameters.

APPENDIX: GALAXY MASS THRESHOLD

In this study, we considered only galaxies with stellar mass $M_{\mathrm{gal}} \geq 3 \times 10^{10} \, \mathrm{M_\odot}$. The motivation behind this choice is two-fold. On one hand, the vast majority of
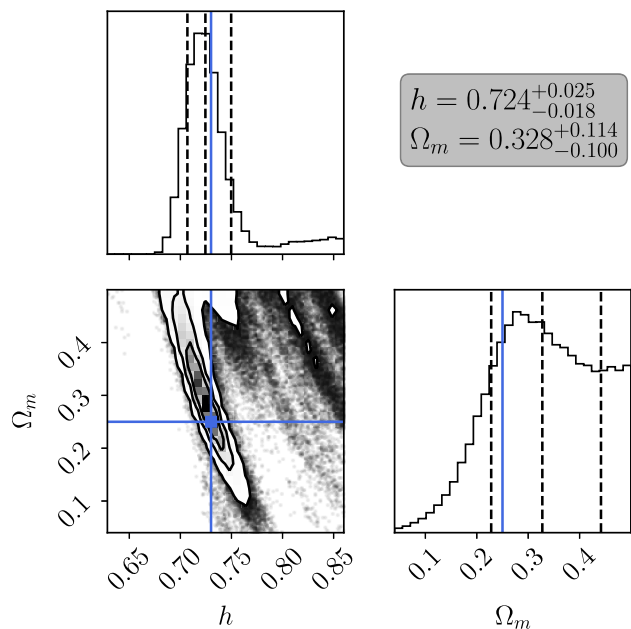

FIG. 9. Averaged joint posterior distribution of 20 different realizations assuming a lower threshold in galaxy's selection.

stars—and compact remnants—in the Universe is hosted in galaxies above this mass threshold; on the other hand, simulating large light cones down to a much lower mass threshold requires significant computing time and memory resources. It is, however, important to investigate to what extent our results are dependent on this specific choice.

To this end, we ran a test light cone considering all galaxies with $M_{gal} \geq 10^{10}$ $M_\odot$. We built the error boxes with the same procedure outlined in the main text, constructing 20 independent realizations of the above-gap SBHB population. The main result of this exercise is shown in Fig. 9. We notice here that, although the results are consistent with those obtained with a higher mass cut, the precision in the determination of the cosmological

parameter is slightly degraded. In particular, the width of the marginalized posterior on h is about 30% larger, with an average error on its determination increasing from $\approx 1.5\%$ to $\gtrsim 2\%$. Similarly, constraints on Ω_m are slightly looser, with typical precision of $\approx 30\%$. The reason of this precision loss might lurk in the peculiar clustering properties of different galaxy populations. For example, specific classes of dwarf galaxies tend to be less clustered and more evenly distributed in the sky [55], which is expected to somewhat reduce the effectiveness of our methodology. Despite this, we notice that the main results of our study are robust and we defer a more comprehensive investigation of effects such as catalog completeness and selection effects to future work.

[1] A. G. Riess, S. Casertano, W. Yuan, L. M. Macri, and D. Scolnic, Large magellanic cloud cepheid standards provide a 1% foundation for the determination of the hubble constant and stronger evidence for physics beyond λCDM, Astrophys. J. **876,** 85 (2019).

[2] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. Banday, R. Barreiro, N. Bartolo, S. Basak *et al.*, Planck 2018 results-VI. Cosmological parameters, Astron. Astrophys. **641,** A6 (2020).

[3] L. Verde, T. Treu, and A. G. Riess, Tensions between the early and late universe, Nat. Astron. **3,** 891 (2019).

[4] B. P. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari *et al.*, Observation of Gravitational Waves from a Binary Black Hole Merger, Phys. Rev. Lett. **116,** 061102 (2016).

[5] R. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Tests of general relativity with binary black holes from the second LIGO-Virgo gravitational-wave transient catalog, Phys. Rev. D **103,** 122002 (2021).

[6] B. P. Abbott *et al.*, A gravitational-wave measurement of the hubble constant following the second observing run of advanced LIGO and Virgo, Astrophys. J. **909,** 218 (2021).

[7] B. P. Abbott *et al.*, A gravitational-wave standard siren measurement of the Hubble constant, Nature (London) **551,** 85 (2017).

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

[9] D. E. Holz and S. A. Hughes, Using gravitational-wave standard sirens, Astrophys. J. **629,** 15 (2005).

[10] A. Petiteau, S. Babak, and A. Sesana, Constraining the dark energy equation of state using LISA observations of spinning massive black hole binaries, Astrophys. J. **732,** 82 (2011).

[11] 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).

[12] W. Del Pozzo, A. Sesana, and A. Klein, Stellar binary black holes in the LISA band: a new class of standard sirens, Mon. Not. R. Astron. Soc. **475,** 3485 (2018).

[13] P. Amaro-Seoane, H. Audley, S. Babak, J. Baker, E. Barausse, P. Bender, E. Berti, P. Binetruy, M. Born, D. Bortoluzzi *et al.*, Laser interferometer space antenna, arXiv:1702.00786.

[14] M. Punturo *et al.*, The Einstein telescope: A third-generation gravitational wave observatory, Classical Quantum Gravity **27,** 194002 (2010).

[15] D. Reitze *et al.*, Cosmic explorer: The U.S. contribution to gravitational-wave astronomy beyond LIGO, in *Bulletin of the American Astronomical Society* (2019), Vol. 51, p. 35, https://baas.aas.org/pub/2020n7i035.

[16] A. Sesana, Prospects for Multiband Gravitational-Wave Astronomy after GW150914, Phys. Rev. Lett. **116,** 231102 (2016).

[17] P. Amaro-Seoane and L. Santamaría, Detection of IMBHs with ground-based gravitational wave observatories: A biography of a binary of black holes, from birth to death, Astrophys. J. **722,** 1197 (2010).

[18] D. Gerosa, S. Ma, K. W. K. Wong, E. Berti, R. O'Shaughnessy, Y. Chen, and K. Belczynski, Multiband gravitational-wave event rates and stellar physics, Phys. Rev. D **99,** 103004 (2019).

[19] A. Mangiagli, M. Bonetti, A. Sesana, and M. Colpi, Merger rate of stellar black hole binaries above the pair-instability mass gap, Astrophys. J. Lett. **883,** L27 (2019).

[20] J. M. Ezquiaga and D. E. Holz, Jumping the gap: Searching for LIGO's biggest black holes, Astrophys. J. Lett. **909,** L23 (2021).

[21] D. Izquierdo-Villalba, S. Bonoli, M. Dotti, A. Sesana, Y. Rosas-Guevara, and D. Spinoso, From galactic nuclei to the halo outskirts: Tracing supermassive black holes across cosmic history and environments, Mon. Not. R. Astron. Soc. **495,** 4681 (2020).

[22] T. Robson, N. J. Cornish, and C. Liu, The construction and use of LISA sensitivity curves, Classical Quantum Gravity **36**, 105011 (2019).

[23] I. Bartos, B. Kocsis, Z. Haiman, and S. Márka, Rapid and bright stellar-mass binary black hole mergers in active galactic nuclei, Astrophys. J. **835**, 165 (2017).

[24] M. Spera and M. Mapelli, Very massive stars, pair-instability supernovae and intermediate-mass black holes with the sevn code, Mon. Not. R. Astron. Soc. **470**, 4739 (2017).

[25] S. E. Woosley, A. Heger, and T. A. Weaver, The evolution and explosion of massive stars, Rev. Mod. Phys. **74**, 1015 (2002).

[26] M. Mapelli, M. Spera, E. Montanari, M. Limongi, A. Chieffi, N. Giacobbo, A. Bressan, and Y. Bouffanais, Impact of the rotation and compactness of progenitors on the mass of black holes, Astrophys. J. **888**, 76 (2020).

[27] P. Madau and T. Fragos, Radiation backgrounds at cosmic dawn: X-rays from compact binaries, Astrophys. J. **840**, 39 (2017).

[28] L.-G. Strolger et al., The hubble higher z supernova search: Supernovae to $z \approx 1.6$ and constraints on type ia progenitor models, Astrophys. J. **613**, 200 (2004).

[29] P. Madau and M. Dickinson, Cosmic star-formation history, Annu. Rev. Astron. Astrophys. **52**, 415 (2014).

[30] P. C. Peters, Gravitational radiation and the motion of two point masses, Phys. Rev. **136**, B1224 (1964).

[31] M. Mapelli and N. Giacobbo, The cosmic merger rate of neutron stars and black holes, Mon. Not. R. Astron. Soc. **479**, 4391 (2018).

[32] F. S. Broekgaarden, E. Berger, C. J. Neijssel, A. Vigna-Gómez, D. Chattopadhyay, S. Stevenson, M. Chruslinska, S. Justham, S. E. de Mink, and I. Mandel, Impact of massive binary star and cosmic evolution on gravitational wave observations I: Black hole-neutron star mergers, Mon. Not. R. Astron. Soc. **508**, 5028 (2021).

[33] M. Zevin, J. Samsing, C. Rodriguez, C.-J. Haster, and E. Ramirez-Ruiz, Eccentric black hole mergers in dense star clusters: The role of binary-binary encounters, Astrophys. J. **871**, 91 (2019).

[34] B. Liu and D. Lai, Hierarchical black hole mergers in multiple systems: Constrain the formation of GW190412-, GW190814-, and GW190521-like events, Mon. Not. R. Astron. Soc. **502**, 2049 (2021).

[35] H. Tagawa, B. Kocsis, Z. Haiman, I. Bartos, K. Omukai, and J. Samsing, Mass-gap mergers in active galactic nuclei, Astrophys. J. **908**, 194 (2021).

[36] P. C. Peters and J. Mathews, Gravitational radiation from point masses in a Keplerian orbit, Phys. Rev. **131**, 435 (1963).

[37] A. Mangiagli, A. Klein, M. Bonetti, M. L. Katz, A. Sesana, M. Volonteri, M. Colpi, S. Marsat, and S. Babak, Observing the inspiral of coalescing massive black hole binaries with LISA in the era of multimessenger astrophysics, Phys. Rev. D **102**, 084056 (2020).

[38] A. Klein, N. Cornish, and N. Yunes, Fast frequency-domain waveforms for spin-precessing binary inspirals, Phys. Rev. D **90**, 124029 (2014).

[39] E. Berti, A. Buonanno, and C. M. Will, Estimating spinning binary parameters and testing alternative theories of gravity with LISA, Phys. Rev. D **71**, 084025 (2005).

[40] A. Nitz, I. Harry, D. Brown, C. M. Biwer, J. Willis, T. D. Canton, C. Capano, L. Pekowsky, T. Dent, A. R. Williamson et al., gwastro/pycbc: PyCBC release v1.16.9 (Zenodo, 2020), 10.5281/ZENODO.3993665.

[41] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era, Phys. Rev. D **93**, 044007 (2016).

[42] H. Tagawa, B. Kocsis, Z. Haiman, I. Bartos, K. Omukai, and J. Samsing, Eccentric black hole mergers in active galactic nuclei, Astrophys. J. Lett. **907**, L20 (2021).

[43] M. Vallisneri, Use and abuse of the Fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects, Phys. Rev. D **77**, 042001 (2008).

[44] R. N. Lang and S. A. Hughes, Measuring coalescing massive binary black holes with gravitational waves: The impact of spin-induced precession, Phys. Rev. D **74**, 122001 (2006).

[45] D. Laghi, N. Tamanini, W. Del Pozzo, A. Sesana, J. Gair, S. Babak, and D. Izquierdo-Villalba, Gravitational-wave cosmology with extreme mass-ratio inspirals, Mon. Not. R. Astron. Soc. **508**, 4512 (2021).

[46] V. Springel, S. D. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly et al., Simulations of the formation, evolution and clustering of galaxies and quasars, Nature (London) **435**, 629 (2005).

[47] W. Del Pozzo, Inference of cosmological parameters from gravitational waves: Applications to second generation interferometers, Phys. Rev. D **86**, 043011 (2012).

[48] N. Tamanini, C. Caprini, E. Barausse, A. Sesana, A. Klein, and A. Petiteau, Science with the space-based interferometer eLISA. III: Probing the expansion of the Universe using gravitational wave standard sirens, J. Cosmol. Astropart. Phys. 04 (2016) 002.

[49] W. Del Pozzo and D. Laghi, Cosmolisa (2020), https://github.com/wdpozzo/cosmolisa.

[50] J. Veitch, W. D. Pozzo, A. Lyttle, M. Williams, C. Talbot, M. Pitkin, G. Ashton, Cody, M. Hübner, A. Nitz, D. Macleod, G. Carullo, G. Davies, and Tony, johnveitch/cpnest: v0.11.3 (2021), 10.5281/zenodo.4470001.

[51] J. Skilling et al., Nested sampling for general Bayesian computation, Bayesian Anal. **1**, 833 (2006).

[52] D. Foreman-Mackey, corner.py: Scatterplot matrices in python, J. Open Source Softwaare **1**, 24 (2016).

[53] K. Kremer, M. Spera, D. Becker, S. Chatterjee, U. N. Di Carlo, G. Fragione, C. L. Rodriguez, C. S. Ye, and F. A. Rasio, Populating the upper black hole mass gap through stellar collisions in young star clusters, Astrophys. J. **903**, 45 (2020).

[54] U. N. Di Carlo, M. Mapelli, Y. Bouffanais, N. Giacobbo, F. Santoliquido, A. Bressan, M. Spera, and F. Haardt, Binary black holes in the pair instability mass gap, Mon. Not. R. Astron. Soc. **497**, 1043 (2020).

[55] D. Tanoglidis et al. (DES Collaboration), Shadows in the dark: Low-surface-brightness galaxies discovered in the dark energy survey, Astrophys. J. Suppl. Ser. **252**, 18 (2021).