


Listening across the cosmic time: standard sirens from ground- and space-based missions in the next decade

Alberto Salvarese^{1,*} , Hsin-Yu Chen¹ ,
Alberto Mangiagli²  and Nicola Tamanini³ 

¹ Department of Physics, The University of Texas at Austin, 2515 Speedway, Austin, TX 78712, United States of America

² Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlenberg 1, 14476 Potsdam, Germany

³ Laboratoire des 2 Infinis—Toulouse (L2IT-IN2P3), Université de Toulouse, CNRS, F-31062 Toulouse Cedex 9, France

E-mail: alberto.salvarese@utexas.edu

Received 30 June 2025; revised 18 August 2025

Accepted for publication 10 September 2025

Published 23 September 2025



CrossMark

Abstract

Precise measurement of the Hubble parameter will enable stringent tests of the standard model for cosmology. Standard sirens, using the luminosity distances measured by gravitational-wave observations of compact binary mergers, are expected to provide such measurements independently in the next decade. With the ground- and space-based gravitational wave observatories, the LIGO-Virgo-KAGRA (LVK) network and the Laser Interferometer Space Antenna (LISA), different types of standard sirens altogether will place constraints across a wide redshift range. In this paper, we forecast the precisions of standard siren Hubble parameter measurements and compare various scenarios, accounting for the dominant sources of systematic uncertainty. Specifically, we find a 2% constraint on H_0 , a 1.5%–3% constraint on $H(z)$ at $z = 1$, and a 3%–5% constraint on $H(z)$ at $z = 7$ when combining LVK and LISA standard sirens with precise redshift measurements from electromagnetic (EM) counterpart observations. We do not find a significant improvement when including standard sirens with no EM counterpart, but which rely on features in the black

* Author to whom any correspondence should be addressed.



Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

hole mass distribution, and the potential systematics introduced by the possible redshift evolution of such features could further degrade the measurement accuracy if not properly accounted for.

Keywords: standard sirens, gravitational waves, cosmology

1. Introduction

The cosmic expansion rate, described by the ‘Hubble parameter’ $H(z)$ [1, 2], is a fundamental quantity to measure in cosmology. High-precision measurements of $H(z)$ enable stringent tests of the standard Λ cold dark matter (Λ CDM) model [3–5] and offer avenues for investigating alternative cosmological theories. Since the first detection of gravitational wave (GW) signals in 2015, followed by the first joint observation of a GW source and its electromagnetic (EM) counterparts in 2017 [6, 7], GW observations have emerged as an independent tool to measure the cosmic expansion rate via the standard siren method [8].

GW signals from compact binary coalescences (CBCs)—such as binary neutron star (BNS), binary black hole (BBH), and neutron star–black hole mergers—carry valuable information about the physical properties of the source systems. The amplitude of the observed GW signal can be approximated as [9]

$$h_s(f) \propto \frac{M_z^{5/6}}{2D_L} f^{-7/6} g_s(\theta_{\text{JN}}), \quad (1)$$

where $s = +, \times$ denotes the GW polarization, $M_z = (1+z)M$ is the redshifted chirp mass of the binary, f is the GW frequency, $g_s(\theta_{\text{JN}})$ is a function of the binary’s inclination angle θ_{JN} , and D_L is the luminosity distance to the source. GW detectors measure a linear combination of the two polarizations: $h(f) = F_+ h_+(f) + F_\times h_\times(f)$, where F_+ and F_\times are the detector’s antenna pattern functions [9–11], which depend on the source’s sky position, polarization angle, and the detector configuration. From the amplitude of the GW signals, we can measure the luminosity distance to the GW sources. The redshifted chirp mass M_z can be inferred from the GW frequency and its evolution in time. Other parameters, if not inferred independently from other measurements (such as EM observations), contribute to the measurement uncertainty of the luminosity distance [8].

In a flat Λ CDM Universe, the luminosity distance is given by

$$D_L = c(1+z) \int_0^z \frac{dz'}{H(z')}, \quad \text{with} \quad H(z) = H_0 \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}. \quad (2)$$

Here, H_0 is the present-day ($z=0$) expansion rate of the Universe, known as the ‘Hubble constant’, while Ω_m and Ω_Λ denote the fractional densities of matter and dark energy, respectively, with the constraint $\Omega_m + \Omega_\Lambda = 1$ in a flat geometry. By measuring the luminosity distance and redshift to astrophysical sources, we can measure the Hubble parameter.

However, GW observations do not directly yield the redshift of the source. Several methods have been proposed to estimate the redshift of GW sources [8, 12–18]. The bright siren method utilizes EM counterparts produced by the binary, enabling redshift estimation by localizing the source and identifying its host galaxy. While this technique holds significant promise for constraining H_0 (e.g. to $\sim 2\%$ with $\mathcal{O}(50)$ events [19]), only one confirmed EM-bright CBC event was observed during the first three observing runs of the LIGO-Virgo-KAGRA (LVK)

observatories—though other potential EM counterpart candidates have been proposed [20–23]—in contrast to over $\mathcal{O}(100)$ CBCs lacking EM counterparts [24].

To address this limitation, alternative methods that do not require EM counterpart detection have been pursued. For example, the spectral siren method estimates redshifts by comparing the detected (redshifted) and intrinsic masses of CBC systems. The masses of CBCs can be inferred from the GW waveform; however, due to the cosmological redshift of the GW signals, the observed masses m_z are related to the source-frame masses m_s by $m_z = (1+z)m_s$. This method uses the distribution of observed masses and compares it with assumed or inferred source-frame mass distributions to extract redshift information [14, 25–27]. Especially, to mitigate biases from inaccurate assumptions about the mass distribution, recent approaches aim to jointly infer the intrinsic mass distribution and cosmological parameters in a self-consistent framework [28–31].

Other standard siren approaches that do not rely on EM counterparts have also been developed [32, 33], including the statistical dark sirens [15–17, 19, 34–41] and love sirens [42–44]. The statistical dark siren method utilizes galaxy catalogs to estimate the redshifts of all potential host galaxies within the localization volume of a GW event. By marginalizing over these candidate redshifts, one can statistically constrain cosmological parameters. The love siren approach, applicable to systems containing at least one neutron star, exploits tidal effects imprinted on the GW signal. The nuclear matter properties of neutron stars—such as the equation of state (EoS)—are unaffected by cosmological redshift. These properties govern the tidal deformability during the inspiral phase and leave characteristic imprints on the GW waveform. If the EoS is constrained, the tidal signature can be used to estimate the source-frame mass of the neutron star. Comparing the estimated source-frame mass with the observed (redshifted) mass allows for inference of the source redshift.

In this paper, we forecast measurements of the Hubble parameter by jointly analyzing bright and spectral sirens, explicitly incorporating their dominant sources of systematic uncertainty. We do not directly consider constraints from the statistical dark siren method (however, see the discussion of ‘golden dark siren’ below), as it remains limited by several key challenges. A primary concern is the incompleteness of galaxy catalogs [17, 19, 40], particularly at higher redshifts, where the catalogs become increasingly incomplete, reducing the precision of the inference of cosmological parameters. An important source of systematics for dark sirens arises from the weighting scheme used to assign host probabilities to galaxies in the catalog, which is often based on observable properties such as star formation rate or color. Recent studies [45–47] have demonstrated that incorrect weighting prescriptions can lead to biased estimates of cosmological parameters. Moreover, as the analysis extends to greater distances, the localization volumes of GW events encompass a rapidly growing number of potential host galaxies, diminishing the statistical power of this method. While these limitations may be alleviated in favorable cases—namely, golden dark sirens, which are nearby and well-localized events associated with only one or a few credible host galaxies—the expected number of such events during the current LVK era is limited [18, 48–51]. Golden dark sirens are anticipated to become significantly more informative with the advent of next-generation GW observatories, such as the Einstein telescope [52, 53] and cosmic explorer [54, 55], projected to become operational beyond 2040 [18]. Although a small number of golden dark sirens may contribute in the near term, their constraining power is expected to be modest and complementary to that of bright sirens. Accordingly, we consider their contribution to be effectively encapsulated within the LVK bright siren forecasts considered in this paper. We also exclude love sirens from our analysis due to their strong dependence on the knowledge of neutron star EoS, which

remains a major open question in nuclear physics. Given the current uncertainties in the EoS, love sirens are not expected to yield strong standard siren constraints yet [56–58].

We focus exclusively on two specific detectors that will be available before 2040: the LVK network and the Laser-Interferometer-Space-Antenna (LISA) [59], expected to provide a full dataset around 2040. LISA is expected to deliver new classes of standard sirens both with [60–66] and without EM counterpart association [67–69]. However there are no comprehensive studies on LISA spectral sirens in the literature so far, and for this reason for LISA we only consider bright sirens given by massive black hole binaries (MBHBs) [60, 61]. These massive binary coalescences can be detected by LISA up to $z \sim 10$ [60, 67, 70], and are likely to produce EM counterparts thanks to the presence of accretion disks or gas in their surroundings [71–74]. Other astrophysical sources—such as extreme mass ratio inspirals and stellar-origin compact binaries—are also expected to be detected by LISA [70] and may provide supplementary cosmological information. However, they will be mostly detected at $z \lesssim 1$ and the number of events remains modest [75, 76], limiting their contribution to standard sirens compared to that could be obtained with ground-based GW detectors. This is why we consider MBHBs the primary LISA standard sirens in this study.

2. Methods

2.1. Bayesian framework

To combine multiple GW events and infer $H(z)$, we adopt a Bayesian hierarchical framework. For a set of hyper-parameters $\vec{\Lambda}$ and observational data $\vec{\mathcal{D}}$, the Bayes theorem [77, 78] states

$$p(\vec{\Lambda}|\vec{\mathcal{D}}) = \pi(\vec{\Lambda}) \frac{\mathcal{L}(\vec{\mathcal{D}}|\vec{\Lambda})}{p(\vec{\mathcal{D}})} \quad (3)$$

where $\pi(\vec{\Lambda})$ represents the prior on $\vec{\Lambda}$, $\mathcal{L}(\vec{\mathcal{D}}|\vec{\Lambda})$ is the likelihood function, and $p(\vec{\mathcal{D}})$ is the evidence. For N_{obs} independently observed events, the likelihood of observing the set of data $\vec{\mathcal{D}} = (\vec{\mathcal{D}}^1, \dots, \vec{\mathcal{D}}^N)$ given the hyper-parameters $\vec{\Lambda}$ can be written as [79–81]

$$\mathcal{L}(\vec{\mathcal{D}}|\vec{\Lambda}) \propto \prod_{i=1}^{N_{\text{obs}}} \frac{\int d\vec{\Theta} \mathcal{L}(\vec{\mathcal{D}}^i|\vec{\Theta}) p(\vec{\Theta}|\vec{\Lambda})}{\beta(\vec{\Lambda})}, \quad (4)$$

where the function

$$\beta(\vec{\Lambda}) = \int_{\vec{\mathcal{D}} > \vec{\mathcal{D}}_{\text{th}}} d\vec{\Theta} d\vec{\mathcal{D}} \mathcal{L}(\vec{\mathcal{D}}|\vec{\Theta}) p(\vec{\Theta}|\vec{\Lambda}) \quad (5)$$

quantifies the probability that a source with parameters $\vec{\Theta}$ exceeds the detection threshold $\vec{\mathcal{D}}_{\text{th}}$ given certain hyper-parameters $\vec{\Lambda}$. This function accounts for selection effects arising from the fact that not all sources are equally detectable. In practice, $\beta(\vec{\Lambda})$ is typically computed by generating synthetic signals from the population model $p(\vec{\Theta}|\vec{\Lambda})$, injecting them into realistic detector noise, and processing them through the detection pipeline to determine whether they exceed $\vec{\mathcal{D}}_{\text{th}}$ [82]. The detection threshold could be a threshold on the signal-to-noise ratio produced by the source in the detector. We follow the methods described in [19] and [29]

to compute the selection function for LVK bright and spectral sirens, respectively. For LISA bright sirens, we adopt the assumptions of [83].

2.1.1. Bayesian framework: bright siren. For the bright sirens approach, the redshifts of the GW sources are provided by EM observation of the hosts. Therefore, for N_{obs} independent events, we have a set of combined GW and EM data $\vec{\mathcal{D}} = (\vec{\mathcal{D}}_{\text{GW}}^1, \vec{\mathcal{D}}_{\text{EM}}^1, \dots, \vec{\mathcal{D}}_{\text{GW}}^{N_{\text{obs}}}, \vec{\mathcal{D}}_{\text{EM}}^{N_{\text{obs}}})$. We can rewrite the likelihood in equation (4) as

$$\mathcal{L}(\vec{\mathcal{D}}|\vec{\Lambda}) = \prod_{i=1}^{N_{\text{obs}}} \frac{\int d\vec{\Theta} \mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i, \vec{\mathcal{D}}_{\text{EM}}^i|\vec{\Theta}) p(\vec{\Theta}|\vec{\Lambda})}{\beta(\vec{\Lambda})}. \quad (6)$$

We separate $\vec{\Theta}$ into relevant physical parameters (D_L, z) and other physical parameters $\vec{\Theta}'$ (e.g. component masses, spins, or the binary inclination angle), and write the likelihood for the i th event as

$$\begin{aligned} \int d\vec{\Theta} \mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i, \vec{\mathcal{D}}_{\text{EM}}^i|\vec{\Theta}) p(\vec{\Theta}|\vec{\Lambda}) &= \int dD_L dz d\vec{\Theta}' \mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i, \vec{\mathcal{D}}_{\text{EM}}^i|D_L, z, \vec{\Theta}') p(D_L, z, \vec{\Theta}'|\vec{\Lambda}) \\ &= \int dD_L dz d\vec{\Theta}' \mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i|D_L, z, \vec{\Theta}') \mathcal{L}(\vec{\mathcal{D}}_{\text{EM}}^i|D_L, z, \vec{\Theta}') \\ &\quad \times p(D_L, z, \vec{\Theta}'|\vec{\Lambda}). \end{aligned} \quad (7)$$

Here, we assume that the GW and EM observations are independent. We can further simplify this expression by knowing that GW data do not provide a redshift measurement in the bright sirens approach, and that EM data do not give a D_L measurement. Finally, accounting for the dependencies between the parameters in $p(D_L, z, \vec{\Theta}'|\vec{\Lambda})$ [19, 83], equation (7) becomes

$$\begin{aligned} \int d\vec{\Theta} \mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i, \vec{\mathcal{D}}_{\text{EM}}^i|\vec{\Theta}) p(\vec{\Theta}|\vec{\Lambda}) &= \int dD_L dz d\vec{\Theta}' \mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i|D_L, \vec{\Theta}') \mathcal{L}(\vec{\mathcal{D}}_{\text{EM}}^i|z, \vec{\Theta}') \\ &\quad \times \delta[D_L - \hat{D}_L(z, \vec{\Lambda})] p(z|\vec{\Lambda}) p(\vec{\Theta}'). \end{aligned} \quad (8)$$

The Dirac δ -function arises from the fact that the luminosity distance is uniquely determined once the cosmological parameters and the redshift are specified (equation (2)). $p(z|\vec{\Lambda}) = \frac{R(z)}{1+z} \frac{dV_C}{dz}(\vec{\Lambda})$ represents the probability of having a bright CBC at redshift z given the choice of hyper-parameters $\vec{\Lambda}$. Thus, by marginalizing over the parameters $\vec{\Theta}'$, and optionally fixing those that are tightly constrained by the presence of an EM counterpart (such as sky localization), we arrive at:

$$\mathcal{L}(\vec{\mathcal{D}}|\vec{\Lambda}) = \prod_{i=1}^{N_{\text{obs}}} \frac{1}{\beta(\vec{\Lambda})} \int dz \mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i|\hat{D}_L(z, \vec{\Lambda})) L(\vec{\mathcal{D}}_{\text{EM}}^i|z) \frac{R(z)}{1+z} \frac{dV_C}{dz}(\vec{\Lambda}) \quad (9)$$

where $\vec{\Lambda} = (H_0, \Omega_m)$. For bright sirens, the uncertainties of redshift are relatively small. Therefore, the merger rate $R(z)$ effectively acts almost as a constant factor in the likelihood.

2.1.2. *Bayesian framework: spectral siren.* The spectral sirens approach relies solely on information extracted from GW observations. Following the same steps outlined above, the relevant physical parameters are the detector-frame (m_z) and source-frame (m_s) component masses, D_L and z , and we can write the numerator of equation (4) for the i th event as:

$$\int d\vec{\Theta} \mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i | \vec{\Theta}) p(\vec{\Theta} | \vec{\Lambda}) = \int dm_z dm_s dD_L dz d\vec{\Theta}' \mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i | m_z, m_s, D_L, z, \vec{\Theta}') \times p(m_z, m_s, D_L, z, \vec{\Theta}' | \vec{\Lambda}). \quad (10)$$

We reformulate this equation using the product rule for joint probabilities, while explicitly accounting for parameter dependencies. Additionally, we note that GW data do not directly yield redshift or source-frame mass measurements. Therefore, we have

$$\mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i | m_z, m_s, D_L, z, \vec{\Theta}') = \mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i | m_z, D_L, \vec{\Theta}') \\ p(m_z, m_s, D_L, z, \vec{\Theta}' | \vec{\Lambda}) = p(m_z | m_s, z) p(D_L | z, \vec{\Lambda}) p(z | \vec{\Lambda}) p(m_s | \vec{\Lambda}) p(\vec{\Theta}'). \quad (11)$$

Given the definitions of luminosity distance (equation (2)) and detector-frame mass $m_z = (1+z)m_s$, once z , $\vec{\Lambda}$, and m_s are specified, both D_L and m_z are fully determined. Thus, the first two terms on the second line can be written as Dirac δ -functions, yielding:

$$\int d\vec{\Theta} \mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i | \vec{\Theta}) p(\vec{\Theta} | \vec{\Lambda}) = \int dm_z dm_s dD_L dz d\vec{\Theta}' \mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i | m_z, D_L, \vec{\Theta}') p(\vec{\Theta}') \\ \times \delta[m_z - m_z(m_s, z)] \delta[D_L - \hat{D}_L(z, \vec{\Lambda})] p(z | \vec{\Lambda}) p(m_s | \vec{\Lambda}). \quad (12)$$

The parameters that we directly measure through GW observations, and thus what actually defines the GW likelihood, are m_z and D_L . Therefore, we can use the property $\delta(f(x)) = \frac{\delta(x-x_0)}{|f'(x_0)|}$, with x_0 such that $f(x_0) = 0$, and rewrite the Dirac δ -functions in terms of m_s , and z :

$$\int d\vec{\Theta} \mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i | \vec{\Theta}) p(\vec{\Theta} | \vec{\Lambda}) = \int dm_z dm_s dD_L dz d\vec{\Theta}' \mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i | m_z, D_L, \vec{\Theta}') \\ \times p(z | \vec{\Lambda}) p(m_s | \vec{\Lambda}) p(\vec{\Theta}') \\ \times \delta\left[m_s - \frac{m_z}{1+z}\right] \left|\frac{\partial m_z}{\partial m_s}\right|^{-1} \delta\left[z - \hat{z}(D_L, \vec{\Lambda})\right] \left|\frac{\partial D_L}{\partial z}\right|^{-1}. \quad (13)$$

where $\hat{z}(D_L, \vec{\Lambda})$ represents the redshift specified by a given D_L and $\vec{\Lambda}$. For spectral sirens, $\vec{\Lambda} = (H_0, \Omega_m, \vec{\Lambda}_m, \vec{\Lambda}_z)$, with $\vec{\Lambda}_m$ and $\vec{\Lambda}_z$ describing the mass and redshift distributions, respectively.

Accounting for the different dependencies of the parameters $\vec{\Theta}$ to the hyper-parameters, from equation (4) we find

$$\begin{aligned} \mathcal{L}(\vec{\mathcal{D}}|\vec{\Lambda}) &= \prod_{i=1}^{N_{\text{obs}}} \frac{1}{\beta(\vec{\Lambda})} \int dD_L dm_z \mathcal{L}(\vec{\mathcal{D}}_{\text{GW}}^i | m_z, D_L) \frac{R(\hat{z}, \vec{\Lambda}_z)}{1 + \hat{z}} \frac{dV_C}{d\hat{z}}(H_0, \Omega_m) \\ &\times p\left(\frac{m_z}{1 + \hat{z}} \middle| \vec{\Lambda}_m\right) \left| \frac{\partial m_z}{\partial m_s} \right|^{-1} \left| \frac{\partial D_L}{\partial \hat{z}} \right|^{-1}. \end{aligned} \quad (14)$$

For the inference, we adopt the framework developed in [29], fitting for only the binary primary mass (m_1) distribution, which is parameterized as a combined Mass-peak + Power-law model [84, 85]:

$$\begin{aligned} p(m_s | \mu_m, \sigma_m, \alpha_m, m_{\text{min}}, m_{\text{max}}, f_p) &\propto \left[f_p \frac{1}{\sqrt{2\pi}\sigma_m} e^{\left(-\frac{(m_s - \mu_m)^2}{2\sigma_m^2}\right)} + (1 - f_p) m_s^{\alpha_m} \right] \\ &\times \mathcal{F}(m_{\text{min}}, m_{\text{max}}). \end{aligned} \quad (15)$$

Here, the function \mathcal{F} smoothly suppresses the mass distribution near the boundaries m_{min} and m_{max} , thus preventing sharp transitions at these mass edges. μ_m , σ_m and f_p describe the position, width and relative weight of the Mass-peak, while α_m is the Power-law index. As done in [29], we checked that including mass ratio $q = m_2/m_1$ in the inference does not alter the results.

We adopt the merger rate parameterization from [86] to model the rate evolution of BBH mergers:

$$p(z | H_0, \Omega_m, \alpha_z, \beta_z, z_p) = \mathcal{C}(\alpha_z, \beta_z, z_p) \frac{1}{1+z} \frac{(1+z)^{\alpha_z}}{1 + \left(\frac{1+z}{1+z_p}\right)^{\alpha_z + \beta_z}} \frac{dV_C}{dz}(H_0, \Omega_m), \quad (16)$$

where $\mathcal{C}(\alpha_z, \beta_z, z_p) = 1 + (1 + z_p)^{-\alpha_z - \beta_z}$ is a normalization constant that ensures the distribution integrates to unity. The parameters α_z and β_z are power-law indices that describe the merger rate at $z < z_p$ and $z > z_p$, respectively.

2.2. Systematic uncertainties of standard sirens

Constraints on cosmological parameters using combinations of different standard siren approaches have been projected (e.g. [18, 58, 87, 88]). However, such projections can be overly optimistic if systematic uncertainties are not properly taken into account. In this paper, we aim to incorporate known sources of systematics and assess their impact on the precision of the resulting measurements.

2.2.1. Systematic uncertainties: bright siren. For LVK bright sirens, known sources of systematic uncertainty have been more thoroughly investigated, and several bias mitigation methods have been developed. These systematics include GW calibration uncertainties, EM selection effects, peculiar motions of the host galaxies, and biases arising from EM constraints on the binary viewing angle [89–100].

The dominant source of systematic uncertainty for LISA bright sirens arises from weak gravitational lensing due to the inhomogeneous distribution of matter along the line of sight [60, 83, 101]. This effect causes magnification or de-magnification of the GW signal, thereby

biasing the inferred luminosity distance. The lensing-induced uncertainty increases with redshift and becomes the primary contributor to the distance uncertainty for sources at $z \gtrsim 0.6$. In this work, we model the lensing effects following the approach described in [83, 101]. More in details, we took equations (14) and (17) from [101] that model the $\sigma_{D_L, \text{lens}}$ error from lensing and summed it in quadrature with the error from LISA parameter estimation $\sigma_{D_L, \text{LISA}}$ and the peculiar velocities $\sigma_{D_L, \text{pv}}$, obtained from [74]. The total error on luminosity distance is therefore obtained as $\sigma_{D_L} = \sqrt{\sigma_{D_L, \text{LISA}}^2 + \sigma_{D_L, \text{lens}}^2 + \sigma_{D_L, \text{pv}}^2}$. As LISA noise, lensing and peculiar velocities are all expected to shift the recovered luminosity distance with respect to the true value, we also shift all the original D_L samples from LISA parameter estimation (centered around the injected value) by a random value extracted from a Normal distribution with zero mean and standard deviation σ_{D_L} .

2.2.2. Systematic uncertainties: spectral siren. For spectral sirens, the primary source of systematic uncertainty arises from the unknown source-frame mass distribution. This distribution is subject to theoretical uncertainties, including those related to stellar evolution and the history of star formation [102, 103]. In recent years, parametric and non-parametric models have been developed to jointly infer the mass distribution and cosmological parameters [25, 26, 29, 30, 104–106]. Non-parametric models offer greater flexibility by making minimal assumptions about the source-frame mass distribution. However, current methods do not consider the potential redshift evolution of this distribution. The mass distribution is expected to evolve with redshift due to the stellar evolution and the evolution of the star-formation environment across cosmic time. Several studies [26, 107, 108] have emphasized the biases that can arise when cosmological parameters are inferred without incorporating the redshift evolution in the mass distribution.

To estimate the impact of this bias, we adopt the model specified in equation (15), and allow its parameters to vary as a function of redshift. For a given hyper-parameter θ of the considered mass distribution, its evolution is modeled using a sigmoid function [109]

$$\theta(z) = \theta_0 + \frac{\theta_z - \theta_0}{1 + \exp\left\{-\frac{1}{\Delta_z}(z - \bar{z}_\theta)\right\}}. \quad (17)$$

θ_0 , and θ_z are the values of θ at $z = 0$ and $z \gg \bar{z}_\theta$, Δ_z controls the steepness of the evolution (the lower the steeper), and \bar{z}_θ represents the point at which the function transitions from lower- z to higher- z values. Specifically, we construct three distinct evolution scenarios (see more details in section 2.3.2): (i) evolution of the Mass-peak parameters (μ_m, σ_m), (ii) evolution of the Power-law parameter (α_m), (iii) combined evolution of both Mass-peak and Power-law parameters ($\mu_m, \sigma_m, \alpha_m$). In the inference, we fit equation (17) for the possible evolution of the three parameters α_m, μ_m , and σ_m for all three scenarios, adopting the uniform priors listed in appendix. We then compare them to the inference assuming no evolution to estimate the bias in cosmological parameters.

2.3. Simulations of GW and EM observations

2.3.1. Simulations: LVK bright sirens. The bright sirens expected from LVK observations are dominated by BNS mergers [110], with possibly a small contribution from NS-BH mergers [111]. BNS mergers can produce gamma-ray bursts, kilonovae, or other types of EM counterparts [112–117]. However, both the intrinsic BNS merger rate and the joint BNS-EM detection rate remain highly uncertain [118–121]. Additionally, a small number of golden dark sirens—well-localized events without EM counterpart—may contribute similarly to the bright

siren sample. In this work, we assume a total of 50 bright sirens, with a few of them may be taken as the contributions from golden dark sirens, will be observed over the next decade [120–122].

From a catalog of 1000 $1.4\text{--}1.4 M_{\odot}$ BNS mergers simulated assuming the expected sensitivities of LIGO-Hanford, LIGO-Livingston, Virgo, and KAGRA in the 5th observing run (O5, `AplusDesignh.txt`, `avirgo_05high_NEW.txt`, `kagra_05128Mpc.txt` in [118]), we randomly select 50 events. For each event, we construct the GW likelihood $\mathcal{L}(\vec{D}_{\text{GW}}^i | D_L)$ following the method developed in [123], and assign a redshift z assuming a fiducial cosmology with $H_0 = 67.74 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_m = 0.3075$, $\Omega_{\Lambda} = 0.6925$, and $\Omega_k = 0.0$ [124]. We refer to the corresponding Hubble parameter as $H_{\text{inj}}(z)$ hereafter. We assume the redshift is perfectly measured, a reasonable assumption when comparing to the D_L uncertainty in the local Universe.

We repeat this procedure five times to reduce the impact of statistical fluctuations, using equation (9) to infer cosmological parameters at each realization, and report the mean results below.

2.3.2. Simulations: LVK spectral sirens. To simulate a catalog of detected BBH events, we begin by generating a population of BBH with specified component masses and redshifts. We then use the publicly available `GWMockCat` package [125], following the methods used in [29], to apply LVK O5 selection effects and produce realistic parameter estimation posteriors.

We assume the BBH rate evolves as equation (16) with parameters $\{\alpha_z, \beta_z, z_p\} = \{1.0, 3.4, 2.4\}$, as inferred in [120]. For the primary mass distribution, we adopt a `Mass-peak+Power-law` (equation (15)) model where we fix $\{m_{\text{min}}, m_{\text{max}}, f_p\} = \{10M_{\odot}, 78M_{\odot}, 0.05\}$ [120]. We allow other primary mass distribution hyper-parameters to evolve in redshift following equation (17). We consider three evolution scenarios:

- (i) `Mass-peak only`: $\{\mu_{m,0}, \mu_{m,z}, \bar{z}_{\mu}, \Delta_{z,\mu}, \sigma_{m,0}, \sigma_{m,z}, \bar{z}_{\sigma}, \Delta_{z,\sigma}\} = \{30M_{\odot}, 15M_{\odot}, 3, 1.5, 7, 5, 3, 1.5\}$ for the evolution of the `Mass-peak` only;
- (ii) `Power-law only`: $\{\alpha_{m,0}, \alpha_{m,z}, \bar{z}_m, \Delta_{z,m}\} = \{-2.7, -3.5, 3, 1.5\}$ for the evolution of the `Power-law` component only;
- (iii) `Peak+Power-law`: the combination of (i) and (ii) for the combined evolution of `Power-law` and `Mass-peak` components.

For each scenario, the hyper-parameter values at redshift zero are based on [120]. In contrast, the hyper-parameters governing redshift evolution are not informed by specific astrophysical predictions. They are instead chosen to represent illustrative deviations from the local population, allowing us to assess the robustness of cosmological inference under a generic redshift evolution model. Since the primary goal of this study is to evaluate the potential impact of redshift evolution on spectral siren measurements, we adopt these values without asserting that they reflect the real underlying population. The secondary mass for each BBH is assigned by drawing a mass ratio $q = m_2/m_1$ from a uniform distribution over the interval $[0, 1]$, assumed to be independent of redshift. As a result, the redshift evolution of the secondary mass follows directly from the evolution of the primary mass.

The three different evolution scenarios are shown in figure 1, where the black line represents the BBHs mass distribution at redshift $z=0$, while the other colored lines show different z -evolution stages. The mass distribution either shifts toward lower characteristic masses (`Mass-peak only` and `Peak+Power-law`) or has a suppression of the high-mass tail (`Power-law only` and `Peak+Power-law`). For each mass evolution scenario, we simulate

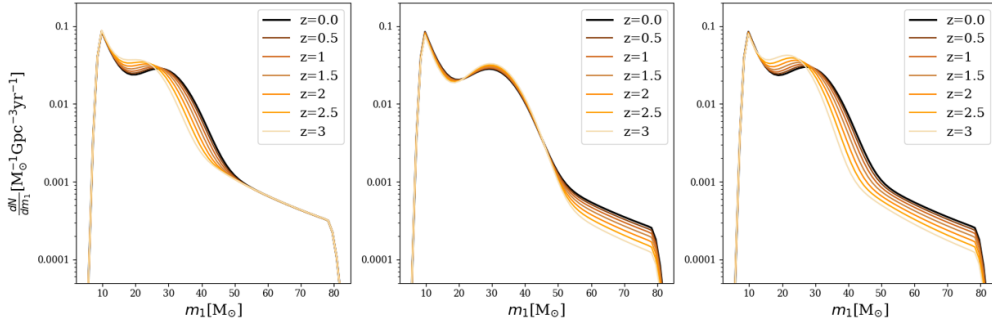


Figure 1. The three simulated primary mass distribution evolution scenarios. The left, central and right panel shows a Mass-peak only, a Power-law only, and a combined Peak+Power-law evolution in redshift, respectively.

five independent realizations of the BBH population. Each realization includes 2800 detected events—corresponding to five year observations under LVK O5 sensitivities—which we take as a representative projection of the detection yield over the next decade.

We then infer $H(z)$ with two distinct assumptions: one assumes the primary mass distribution evolves, and another assumes it does not. Both of them adopt the simulated mass distribution model equation (15). An example of the resulting marginalized posterior distributions is shown in figure 2, where the blue (orange) contours correspond to the inference assuming an evolving (non-evolving) mass distribution under the Power-law only scenario. We see that while assuming an evolving mass distribution (blue) enables a proper recovery of the fiducial cosmological parameters (black dashed lines), neglecting this evolution can lead to biased inferences (orange). Similar results are found in other evolutionary scenarios.

We use the results assuming that the primary mass distribution evolves as our estimate of the statistical uncertainty. We then compare them to the ones assuming no evolution in order to estimate the systematic uncertainties originated from the redshift evolution.

2.3.3. Simulations: LISA bright sirens. We adopt the Pop3 and Q3d MBHB formation channels described in [60] as our bright CBC events (we consider the Q3nd model unrealistic as it completely neglects the time delay between the merger of two galaxies and the merger of their corresponding MBHs). These populations are based on the results of semi-analytical models (see [126, 127] and reference therein) that can track the evolution of MBHs across cosmic time. We simulate a 5 year observing period with an 80% duty cycle.

To model the number of detectable EM counterparts, we follow the framework of [60, 83], considering three scenarios: (i) *Pessimistic*: the EM emission is collimated and the obscuration given by the surrounding gas and dust is included, giving the most conservative counterpart yield; (ii) *Optimistic*: assuming no obscuration and isotropic EM emissions; and (iii) *Moderate*: an intermediate case, given by the average between (i) and (ii). The expected number of MBHB with EM counterpart is reported in table 1. As for the previous simulations, we randomly draw the corresponding number of MBHBs for five independent realizations and infer $H(z)$ through equation (9).

For LISA bright sirens, we do not consider a perfectly measured redshift. Instead, we model the EM likelihood as a Normal distribution centered at $z = z_{\text{true}} + \mathcal{N}(0, \sigma_z)$ with standard deviation σ_z computed as described in [60, 83]. The value of σ_z depends on the method and facility used to estimate the redshift of the host galaxy. For example, if the redshift is obtained with the

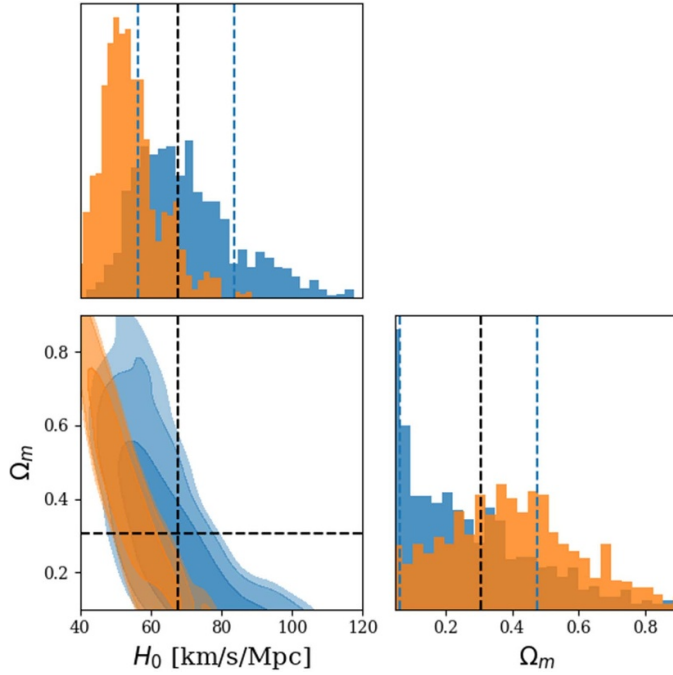


Figure 2. H_0 and Ω_m joint inference using a population of LVK spectral sirens evolved with the Power-law only evolution scenario. We show the cosmological parameter inference with (blue) and without (orange) assuming the mass distribution evolves in the inference. The different shaded regions report the 68%, 90%, and 95% credible intervals. The fiducial cosmological values are shown as black dashed lines. The dashed blue lines in the marginalized posterior distributions indicate the 68% credible intervals for the inference that assumes an evolving primary mass distribution.

Table 1. Number of LISA bright sirens over the 5 year observing period for different MBHB formation channels and EM counterpart detection scenarios [83].

	Pessimistic	Moderate	Optimistic
Pop3	2	4	6
Q3d	3	9	15

Vera Rubin Observatory, we assume $\sigma_z = 0.031 \times (1 + z)$, using photometric measurements. If the redshift is determined with the extremely large telescope and the galaxy is sufficiently bright, we assume that the redshift can be measured spectroscopically, leading to $\sigma_z = 10^{-3}$. If the galaxy is fainter, we still expect to be able to get an estimate of the redshift thanks to the Lyman- α or the Balmer break: in these cases, we assume $\sigma_z = 0.2$ and $\sigma_z = 0.5$, respectively. More details on the choice of σ_z can be found in section IV of [60].

2.4. Combination of sirens

Both LISA bright sirens and LVK spectral sirens provide constraints on $H(z)$ at higher redshifts [26, 29, 83], while LVK bright sirens are particularly well-suited for measuring the local expansion rate, H_0 . Different types of sirens provide complementary information, allowing

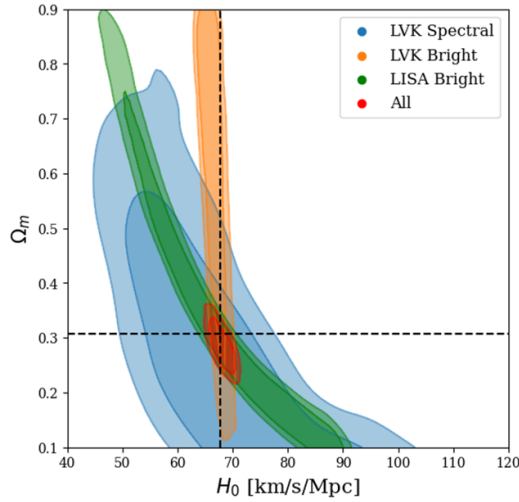


Figure 3. An example of the inferred cosmological parameter posterior from different sirens and the combined posterior. The blue, orange, and green contours show the joint (H_0, Ω_m) posterior inferred from the LVK spectral (assuming the mass distribution evolves following the Power-law only scenario), LVK bright, and Pessimistic Pop3 LISA bright sirens, respectively. The red contour shows the combined posterior of the three. The black dashed lines denote the injected H_0 and Ω_m values. The different shaded regions represent the 68% (dark) and 90% (light) credible intervals.

for constraints on the cosmic expansion rate over a broad range of redshifts. To combine the information from different siren types, we adopt a sequential inference strategy: the posterior distribution obtained from one class of sirens is used as the prior for the inference with another class. This procedure is mathematically equivalent to a single, joint inference over all the considered datasets, provided that the datasets are statistically independent and the same prior is used. We adopt this approach for computational efficiency, and because it enables direct access to the inferred posteriors from individual classes of sirens.

An example of the resulting posteriors from the individual methods and their combination is given in figure 3, which shows how combining standard siren methods can improve cosmological constraints in the (H_0, Ω_m) parameter space. In particular, LISA bright and LVK spectral sirens are both sensitive to $H(z)$ at $z > 0$ and exhibit an anti-correlation between H_0 and Ω_m . This arises from the definition of D_L (equation (2)) in a flat Λ CDM Universe: for a given luminosity distance measurement, an increase in H_0 can be compensated by a decrease in Ω_m , and vice versa. This degeneracy is more pronounced at higher redshifts, where D_L becomes sensitive to both parameters. In contrast, LVK bright sirens, which are predominantly located at low redshift, provide direct and precise constraints on H_0 , effectively breaking the degeneracy and improving the overall precision of the cosmological inference.

3. Results

In figure 4, we show the relative systematic uncertainty in $H(z)$ using the LVK spectral sirens. For each simulated realization described in section 2.3.2, we take the H_0 and Ω_m posterior samples and compute $H(z)$ following equation (2). We compute the absolute difference σ_{sys}

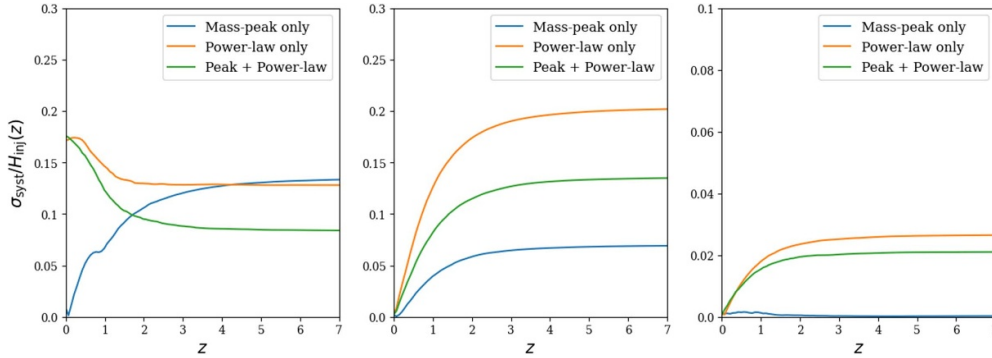


Figure 4. Relative systematic uncertainties for LVK spectral sirens. We estimate the systematics by comparing the absolute difference between the medians of the $H(z)$ posteriors inferred assuming evolving and non-evolving primary mass distribution, for three different evolution scenarios: *Mass-peak only*, *Power-law only*, and *Peak+Power-law*. Left panel: LVK spectral sirens only. Central panel: combination of LVK spectral and LVK bright sirens. Right panel: combination of LVK spectral, LVK bright, and LISA bright sirens (Pop3 *Pessimistic* scenario). Combining all classes of sources reduces the relative systematic uncertainties across redshift and mitigates the variability between different evolution scenarios.

between the medians of the $H(z)$ posterior samples inferred assuming evolving and non-evolving primary mass distribution. We then report the average of this quantity over five independent realizations.

In the case where only LVK spectral sirens are considered (left panel), we find that different mass evolution scenarios yield systematic biases ranging from $\sim 6\%$ (*Mass-peak only*) to $\sim 14\%$ (*Power-law only*) at redshift $z=1$, and from $\sim 8\%$ (*Peak+Power-law*) to $\sim 13\%$ (*Mass-peak only*) at $z=7$. By combining LVK spectral and LVK bright sirens (central panel), the systematic uncertainty at $z=0$ is negligible due to the strong constraint on H_0 provided by the bright sirens. However, the systematic increases with redshift: from $\sim 4\%$ (*Mass-peak only*) to $\sim 12\%$ (*Power-law only*) at $z=1$, and from $\sim 7\%$ (*Mass-peak only*) to $\sim 20\%$ (*Power-law only*) at $z=7$. This behavior arises from the anti-correlation between H_0 and Ω_m : fixing H_0 around its injected value constrains Ω_m to values that may be biased when the underlying mass distribution is incorrectly modeled. As a result, even if H_0 is correctly recovered, the inferred $H(z)$ deviates from the injected expansion history at $z > 0$. When all methods are combined (right panel), the joint constraint on (H_0, Ω_m) becomes significantly tighter, effectively suppressing the impact of an incorrect modeling of the mass distribution. The right panel refers to the Pop3 *Pessimistic* scenario for LISA bright sirens; however, we find that all other EM counterpart scenarios yield similar results.

We emphasize that the systematic shown in figure 4 depends on the specific redshift evolution assumed. Varying the rate at which a particular mass distribution feature evolves with redshift can significantly alter the resulting systematic bias.

In figures 5 (Pop3 formation channel) and 6 (Q3d formation channel), we compare the $H(z)$ uncertainties using different standard sirens and their combinations. The upper and lower panels correspond to the LISA bright siren *Pessimistic* and *Optimistic* EM counterpart scenarios, respectively. The left panels present the relative 1σ statistical uncertainties on $H(z)$, σ_{stat} , for LVK spectral, LVK bright, and LISA bright sirens individually, as well as for their combined results. We compute σ_{stat} as half the width of the symmetric 68% credible interval of

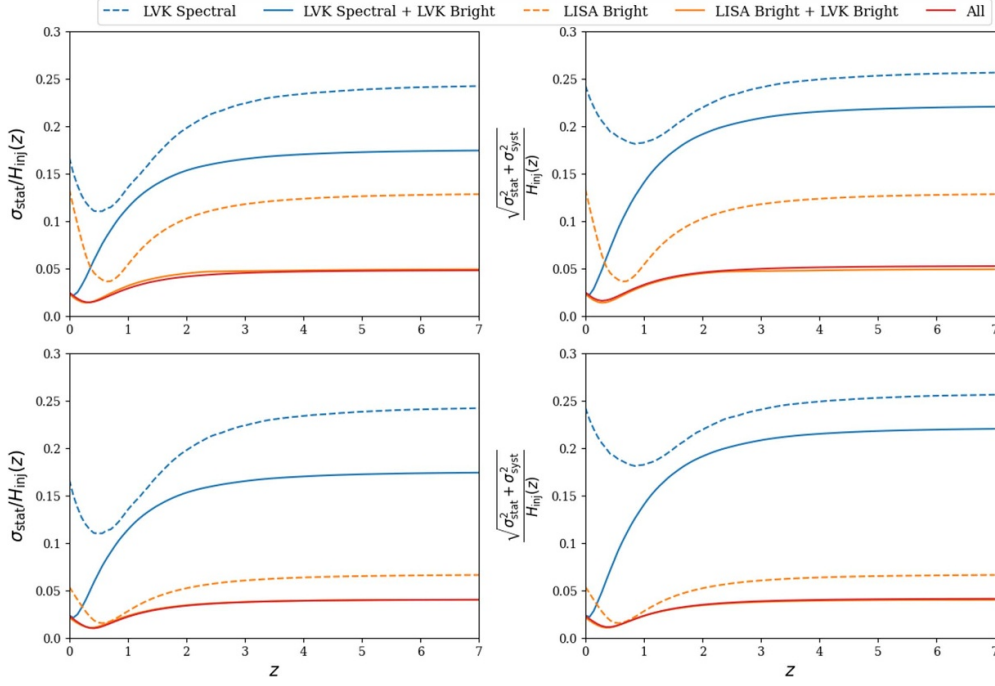


Figure 5. Left panels: Relative statistical uncertainties on $H(z)$ obtained from the different methods considered, shown for the LVK bright sirens, LVK spectral sirens with a **Peak+Power-law** evolving primary mass distribution, and the LISA bright siren Pop3 formation channel. The upper and lower panels correspond to the LISA bright siren **Pessimistic** and **Optimistic** EM counterpart scenarios, respectively. Right panels: same as the left panel but including the spectral siren systematic (σ_{syst}) uncertainties.

$H(z)$ at each redshift z . The right panels add in the systematic uncertainties estimated for the LVK spectral sirens, as given in figure 4. Here we take the **Peak+Power-law** mass evolution scenario for the LVK spectral sirens, as it represents the smallest systematic at high redshifts among the considered evolution scenarios. We note that the statistical uncertainties associated with the other two mass evolution scenarios are similar to that of the **Peak+Power-law** evolution.

We find that:

- *With only one type of sirens:* The LVK spectral sirens method alone (blue dashed line), when accounting for the redshift evolution of the source-frame primary mass distribution, achieves statistical uncertainties below $\sim 25\%$ across the entire redshift range. In particular, the inferred $H(z)$ precision reaches $\sim 14\%$ at $z=1$ and $\sim 25\%$ at $z=7$. LISA bright sirens alone (orange dashed line) constrain $H(z)$ to $\sim 5\%$ at $z=1$ and $\sim 13\%$ at $z=7$ in the **Pessimistic** scenario for both Pop3 and Q3d. In the **Optimistic** scenario, the two formation channels give slightly different results, due to the larger difference in the number of EM counterparts. At $z=1$ ($z=7$), the uncertainty reduces to $\sim 3\%$ ($\sim 7\%$) for Pop3, and to $\sim 2\%$ ($\sim 5\%$) for Q3d.

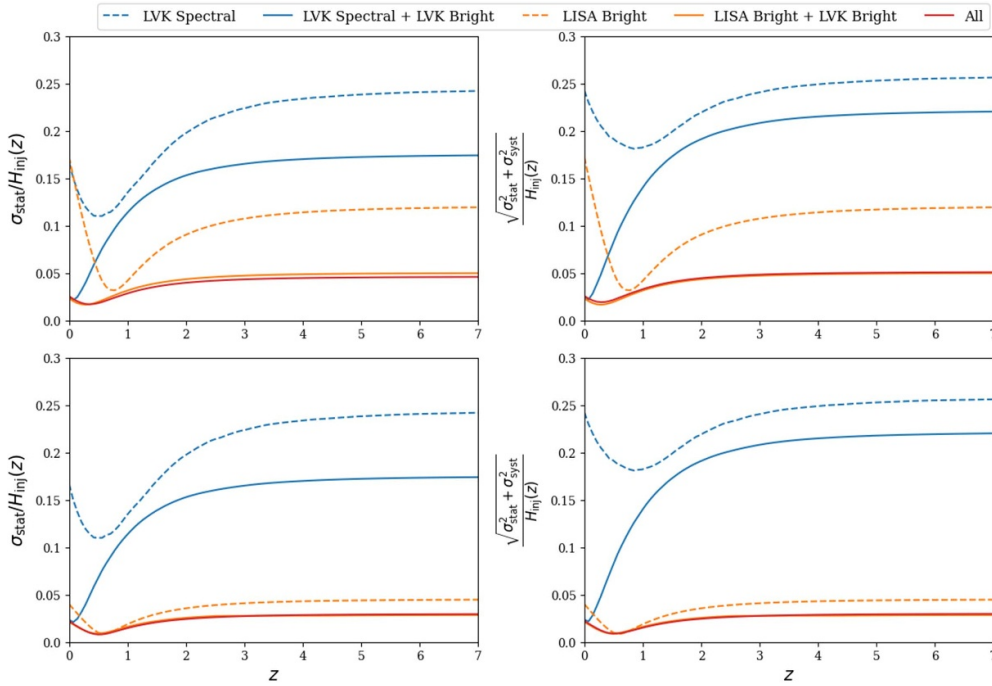


Figure 6. Same as figure 5, but for the LISA bright siren Q3d MBHB formation channel.

- *With two types of sirens:* As expected, the LVK bright sirens significantly improves the $H(z)$ constraints at low redshifts. The statistical uncertainty on H_0 is $\sim 2\%$. Due to the anti-correlation between H_0 and Ω_m (see figure 3), the constraints on H_0 improves the precision of $H(z)$ at higher redshifts. We find that the statistical uncertainty on $H(z)$ remains below 20% across all redshifts when combining LVK spectral and LVK bright sirens, and below 5% when combining LISA bright and LVK bright sirens. In particular, the tightest constraint is given by the Q3d Optimistic scenario (15 LISA bright sirens), which achieves a $H(z)$ precision of $\sim 1.5\%$ at $z = 1$ and $\sim 3\%$ at $z = 7$.
- *With all three types of sirens:* Adding the LVK spectral sirens to the LISA bright + LVK bright sirens only leads to marginal improvements *even* in the LISA bright siren Pessimistic scenario. In the LISA bright siren Optimistic scenario, the contribution from LVK spectral sirens becomes negligible, as the $H(z)$ constraints are already dominated by LISA bright sirens.
- *With the spectral siren systematics:* Even if we consider the smallest systematic at high redshifts using the Peak+Power-law mass evolution scenario (green line in the left panel of figure 4), the systematic uncertainty dominates over the statistical uncertainty in all cases involving the LVK spectral sirens. Consequently, if the systematic in the LVK spectral sirens are not accounted for, the most precise constraints on $H(z)$ come from the combination of LISA and LVK bright sirens.

4. Conclusions

In this paper, we estimate the standard siren measurements of Hubble parameter within the next decade, considering their dominant systematics.

We find that even in the *Pessimistic* EM counterpart scenario, LISA bright sirens offer tighter constraints on $H(z)$ than LVK spectral sirens. The difference is mainly due to the sensitivity of LISA to higher redshift binaries, the precise determination of the luminosity distance, and the precise determination of the redshifts for bright sirens. On the other hand, the spectral sirens are less precise due to the uncertainty of mass distribution and its evolution.

When combining LISA and LVK bright sirens, we find $\sim 2\%$ constraint on H_0 , $1.5\%–3\%$ constraint on $H(z)$ at $z = 1$, and $3\%–5\%$ constraint on $H(z)$ at $z = 7$. We do not find significant improvement when adding in the LVK spectral sirens, and the potential systematics introduced by them could further degrade the measurements.

However, spectral sirens can contribute substantially with next-generation ground-based GW interferometers, which are expected to provide sub-percent $H(z)$ measurements [18, 26] due to their higher detection rates and redshift reach. Therefore, mitigating the systematic associated with the mass evolution of the binary population will be essential to achieve such precision. This could be achieved through a combination of improved population-synthesis modeling, inference of population properties from large GW datasets with next-generation detectors, and joint constraints from independent probes [26, 51, 103, 128, 129]. In particular, future bright siren measurements could be used to calibrate the primary mass distribution for spectral sirens, including its possible redshift evolution, exploiting the fact that a consistent mass model should yield the same cosmological parameters. Such cross-calibration could reduce modeling uncertainties and improve the robustness of cosmological inference from spectral sirens.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

The authors would like to thank Gabriele Perna for the LIGO Scientific Collaboration internal review. A S and H -Y C are supported by the National Science Foundation under Grant PHY-2308752 and Department of Energy Grant DE-SC0025296. The authors are grateful for computational resources provided by the LIGO Laboratory and supported by National Science Foundation Grants PHY-0757058 and PHY-0823459. This is LIGO Document Number P2500382. N T acknowledges support from the French space agency CNES in the framework of LISA, and from the Agence Nationale de la Recherche (ANR) through the MRSEI project ANR-24-MRS1-0009-01.

Appendix. Inference priors

We report the uniform prior ranges adopted in our inferences in table A1. For \bar{z} , we choose the range $[1, 10]$, as current observational data disfavor significant evolution of the black hole mass distribution below redshift $z \sim 1$ [109].

Table A1. Prior ranges for cosmological and population hyper-parameters used in the hierarchical inference.

Cosmological parameters			
H_0	Ω_m		
(30, 120)	(0.01, 0.99)		
Evolving mass distribution parameters			
m_{\min}	m_{\max}		
(1.0, 30.0)	(30.0, 100.0)		
α_{m0}	α_{mz}	$\bar{z}\alpha_m$	Δ_z
(-5.0, -1.0)	(-5.0, -1.0)	(1.0, 10.0)	(0.0, 5.0)
μ_{m0}	μ_{mz}	$\bar{z}\mu_m$	Δ_z
(1.0, 60.0)	(1.0, 60.0)	(1.0, 10.0)	(0.0, 5.0)
σ_{m0}	σ_{mz}	$\bar{z}\sigma_m$	Δ_z
(1.0, 15.0)	(1.0, 15.0)	(1.0, 10.0)	(0.0, 5.0)
BBH redshift distribution parameters			
α_z	β_z	z_p	
(0.0, 5.0)	(0.0, 5.0)	(0.0, 8.0)	

ORCID iDs

Alberto Salvarese  0000-0001-9653-1778

Hsin-Yu Chen  0000-0001-5403-3762

Alberto Mangiagli  0000-0002-3689-1664

Nicola Tamanini  0000-0001-8760-5421

References

- [1] Hubble E 1929 *Proc. Natl Acad. Sci.* **15** 168–73
- [2] Lemaître A G 1931 *Mon. Not. R. Astron. Soc.* **91** 483–90
- [3] Dodelson S 2003 *Modern Cosmology* (Academic)
- [4] Weinberg S 2008 *Cosmology* (OUP Oxford)
- [5] Lyth D H and Liddle A R 2009 *The Primordial Density Perturbation* (Cambridge University Press)
- [6] Abbott B P et al (LIGO Scientific Collaboration and Virgo Collaboration) 2016 *Phys. Rev. Lett.* **116** 061102
- [7] Abbott B et al 2017 *Phys. Rev. Lett.* **119** 161101
- [8] Schutz B F 1986 *Nature* **323** 310–1
- [9] Maggiore M 2007 *Gravitational Waves. Vol. 1: Theory and Experiments* (Oxford University Press)
- [10] Schutz B F and Tinto M 1987 *Mon. Not. R. Astron. Soc.* **224** 131–54
- [11] Forward R L 1978 *Phys. Rev. D* **17** 379–90
- [12] Holz D E and Hughes S A 2005 *Astrophys. J.* **629** 15–22
- [13] Dalal N, Holz D E, Hughes S A and Jain B 2006 *Phys. Rev. D* **74** 063006
- [14] Taylor S R, Gair J R and Mandel I 2012 *Phys. Rev. D* **85** 023535
- [15] Del Pozzo W 2012 *Phys. Rev. D* **86** 043011

- [16] Gair J R et al 2023 *Astron. J.* **166** 22
- [17] Gray R et al 2020 *Phys. Rev. D* **101** 122001
- [18] Chen H Y, Ezquiaga J M and Gupta I 2024 *Class. Quantum Grav.* **41** 125004
- [19] Chen H Y, Fishbach M and Holz D E 2018 *Nature* **562** 545–7
- [20] Graham M J et al 2020 *Phys. Rev. Lett.* **124** 251102
- [21] Abac A G et al 2024 *Astrophys. J. Lett.* **970** L34
- [22] Magaña Hernandez I, d’Emilio V, Morisaki S, Bhardwaj M and Palmese A 2024 *Astrophys. J. Lett.* **971** L5
- [23] Kunnunkai K, Palmese A, Bulla M, Dietrich T, Farah A M and Pang P T H 2024 Kilonova emission from GW230529 and mass gap neutron star-black hole mergers (arXiv:2409.10651)
- [24] Abbott R et al 2023 *Phys. Rev. X* **13** 041039
- [25] Farr W M, Fishbach M, Ye J and Holz D E 2019 *Astrophys. J. Lett.* **883** L42
- [26] Ezquiaga J M and Holz D E 2022 *Phys. Rev. Lett.* **129** 061102
- [27] Agarwal A et al 2025 Blinded mock data challenge for gravitational-wave cosmology-I: assessing the robustness of methods using binary black holes mass spectrum (arXiv:2412.14244)
- [28] Mastrogiovanni S, Leyde K, Karathanasis C, Chassande-Mottin E, Steer D A, Gair J, Ghosh A, Gray R, Mukherjee S and Rinaldi S 2021 *Phys. Rev. D* **104** 062009
- [29] Farah A M, Callister T A, Ezquiaga J M, Zevin M and Holz D E 2025 *Astrophys. J.* **978** 153
- [30] Hernandez I M and Ray A 2024 Beyond gaps and bumps: spectral siren cosmology with non-parametric population models (arXiv:2404.02522)
- [31] Mali U and Essick R 2025 *Astrophys. J.* **980** 85
- [32] Cousins B, Schumacher K, Chung A K W, Talbot C, Callister T, Holz D E and Yunes N 2025 The stochastic siren: astrophysical gravitational-wave background measurements of the hubble constant (arXiv:2503.01997)
- [33] Ferraiuolo S, Mastrogiovanni S, Escoffier S and Kajfasz E 2025 Inferring astrophysics and cosmology with individual compact binary coalescences and their gravitational-wave stochastic background (arXiv:2503.14686)
- [34] MacLeod C L and Hogan C J 2008 *Phys. Rev. D* **77** 043512
- [35] Fishbach M et al 2019 *Astrophys. J. Lett.* **871** L13
- [36] Soares-Santos M et al 2019 *Astrophys. J. Lett.* **876** L7
- [37] Finke A, Foffa S, Iacovelli F, Maggiore M and Mancarella M 2021 *J. Cosmol. Astropart. Phys.* **JCAP08(2021)** 026
- [38] Mukherjee S, Wandelt B D, Nissanke S M and Silvestri A 2021 *Phys. Rev. D* **103** 043520
- [39] Mancarella M, Finke A, Foffa S, Genoud-Prachex E, Iacovelli F and Maggiore M 2022 Cosmology and modified gravity with dark sirens from GWTC-3 (arXiv:2203.09238)
- [40] Gray R, Messenger C and Veitch J 2022 *Mon. Not. R. Astron. Soc.* **512** 1127–40
- [41] Mukherjee S, Krolewski A, Wandelt B D and Silk J 2024 *Astrophys. J.* **975** 189
- [42] Messenger C and Read J 2012 *Phys. Rev. Lett.* **108** 091101
- [43] Chatterjee D, Hegade K R A, Holder G, Holz D E, Perkins S, Yagi K and Yunes N 2021 *Phys. Rev. D* **104** 083528
- [44] Jin S J, Li T N, Zhang J F and Zhang X 2023 *J. Cosmol. Astropart. Phys.* **JCAP08(2023)** 070
- [45] Hanselman A G, Vijaykumar A, Fishbach M and Holz D E 2025 *Astrophys. J.* **979** 9
- [46] Perna G, Mastrogiovanni S and Ricciardone A 2025 *Astron. Astrophys.* **698** A128
- [47] Naveed K, Turski C and Ghosh A 2025 Dark standard siren cosmology with bright galaxy subsets (arXiv:2505.11268)
- [48] Chen H Y and Holz D E 2016 Finding the one: identifying the host galaxies of gravitational-wave sources (arXiv:1612.01471)
- [49] Nishizawa A 2017 *Phys. Rev. D* **96** 101303
- [50] Borhanian S, Dhani A, Gupta A, Arun K G and Sathyaprakash B S 2020 *Astrophys. J. Lett.* **905** L28
- [51] Muttoni N, Laghi D, Tamanini N, Marsat S and Izquierdo-Villalba D 2023 *Phys. Rev. D* **108** 043543
- [52] Punturo M et al 2010 *Class. Quantum Grav.* **27** 194002
- [53] Branchesi M et al 2023 *J. Cosmol. Astropart. Phys.* **JCAP07(2023)** 068
- [54] Reitze D et al 2019 Cosmic explorer: the US contribution to gravitational-wave astronomy beyond LIGO (arXiv:1907.04833)
- [55] Evans M et al 2021 A horizon study for cosmic explorer: science, observatories, and community (arXiv:2109.09882)
- [56] Haster C J, Chatziioannou K, Bauswein A and Clark J A 2020 *Phys. Rev. Lett.* **125** 261101

- [57] Kashyap R, Dhani A and Sathyaprakash B 2022 *Phys. Rev. D* **106** 123001
- [58] Dhani A, Borhanian S, Gupta A and Sathyaprakash B 2022 Cosmography with bright and love sirens (arXiv:2212.13183)
- [59] Baker J et al 2019 The laser interferometer space antenna: unveiling the millihertz gravitational wave sky (arXiv:1907.06482)
- [60] Mangiagli A, Caprini C, Volonteri M, Marsat S, Vergani S, Tamanini N and Inchauspé H 2022 *Phys. Rev. D* **106** 103017
- [61] Tamanini N, Caprini C, Barausse E, Sesana A, Klein A and Petiteau A 2016 *J. Cosmol. Astropart. Phys.* **JCAP04(2016)002**
- [62] Tamanini N 2017 *J. Phys.: Conf. Ser.* **840** 012029
- [63] Toscani M, Burke O, Liu C, Zamel N B, Tamanini N and Pozzoli F 2024 *Phys. Rev. D* **109** 063505
- [64] Cormann M, Ghosh A, Escamilla-Rivera C, Hendry M A, Marsat S and Tamanini N 2022 *Phys. Rev. D* **105** 064061
- [65] Speri L, Tamanini N, Caldwell R R, Gair J R and Wang B 2021 *Phys. Rev. D* **103** 083526
- [66] Belgacem E et al (LISA Cosmology Working Group) 2019 *J. Cosmol. Astropart. Phys.* **JCAP07(2019)024**
- [67] Laghi D, Tamanini N, Del Pozzo W, Sesana A, Gair J, Babak S and Izquierdo-Villalba D 2021 *Mon. Not. R. Astron. Soc.* **508** 4512–31
- [68] Liu C, Laghi D and Tamanini N 2024 *Phys. Rev. D* **109** 063521
- [69] Muttoni N, Mangiagli A, Sesana A, Laghi D, Del Pozzo W, Izquierdo-Villalba D and Rosati M 2022 *Phys. Rev. D* **105** 043509
- [70] Auclair P et al 2023 *Living Rev. Relativ.* **26** 5
- [71] Armitage P J and Natarajan P 2002 *Astrophys. J.* **567** L9
- [72] Milosavljević M and Phinney E S 2005 *Astrophys. J.* **622** L93
- [73] Dotti M, Salvaterra R, Sesana A, Colpi M and Haardt F 2006 *Mon. Not. R. Astron. Soc.* **372** 869–75
- [74] Kocsis B, Frei Z, Haiman Z and Menou K 2006 *Astrophys. J.* **637** 27
- [75] Ruiz-Rocha K, Holley-Bockelmann K, Jani K, Mapelli M, Dunham S and Gabella W 2025 *Astrophys. J.* **981** 27
- [76] Buscicchio R, Torrado J, Caprini C, Nardini G, Karnesis N, Pieroni M and Sesana A 2025 *J. Cosmol. Astropart. Phys.* **JCAP01(2025)084**
- [77] Bayes T 1763 *Phil. Trans. R. Soc.* **53** 370–418
- [78] Jeffreys H 1998 *The Theory of Probability (Oxford Classic Texts in the Physical Sciences)* (OUP Oxford) (available at: <https://books.google.com/books?id=vh9Act9rtzQC>)
- [79] Loredo T J and Lamb D Q 2002 *Phys. Rev. D* **65** 063002
- [80] Mandel I, Farr W M and Gair J R 2019 *Mon. Not. R. Astron. Soc.* **486** 1086–93
- [81] Vitale S, Gerosa D, Farr W M and Taylor S R 2020 *Handbook of Gravitational Wave Astronomy* pp 1–60 (Springer) (available at: https://doi.org/10.1007/978-981-15-4702-7_45-1)
- [82] Farr W M 2019 *Res. Notes AAS* **3** 66
- [83] Mangiagli A, Caprini C, Marsat S, Speri L, Caldwell R R and Tamanini N 2024 Massive black hole binaries in lisa: constraining cosmological parameters at high redshifts (arXiv:2312.04632)
- [84] Abbott R et al 2021 *Astrophys. J. Lett.* **913** L7
- [85] Abbott R et al 2023 *Phys. Rev. X* **13** 011048
- [86] Callister T, Fishbach M, Holz D E and Farr W M 2020 *Astrophys. J. Lett.* **896** L32
- [87] Wang R, Ruan W H, Yang Q, Guo Z K, Cai R G and Hu B 2021 *Natl Sci. Rev.* **9** nwab054
- [88] Yang T 2021 *J. Cosmol. Astropart. Phys.* **JCAP05(2021) 044**
- [89] Huang Y, Chen H Y, Haster C J, Sun L, Vitale S and Kissel J S 2025 *Phys. Rev. D* **111** 063034
- [90] Sun L et al 2020 *Class. Quantum Grav.* **37** 225008
- [91] Sun L, Goetz E, Kissel J S, Betzwieser J, Karki S, Bhattacharjee D, Covas P B, Datrier L E H, Kandhasamy S, Lecoche Y K, Mendell G, Mistry T, Payne E, Savage R L, Viets A, Wade M, Weinstein A J, Aston S, Cahillane C, Driggers J C, Dwyer S E and Urban A 2021 Characterization of systematic error in advanced LIGO calibration in the second half of O3 (arXiv:2107.00129)
- [92] Howlett C and Davis T M 2020 *Mon. Not. R. Astron. Soc.* **492** 3803–15
- [93] Mozzon S, Ashton G, Nuttall L K and Williamson A R 2022 *Phys. Rev. D* **106** 043504
- [94] Mukherjee S, Lavaux G, Bouchet F R, Jasche J, Wandelt B D, Nissanke S, Leclercq F and Hotokezaka K 2021 *Astron. Astrophys.* **646** A65
- [95] Nicolaou C, Lahav O, Lemos P, Hartley W and Braden J 2020 *Mon. Not. R. Astron. Soc.* **495** 90–97
- [96] Kumar S, Nitz A H and Forteza X J 2025 *Astrophys. J.* **982** 67

- [97] Chen H Y 2020 *Phys. Rev. Lett.* **125** 201301
- [98] Chen H Y, Talbot C and Chase E A 2024 *Phys. Rev. Lett.* **132** 191003
- [99] Müller M, Mukherjee S and Ryan G 2024 *Astrophys. J. Lett.* **977** L45
- [100] Salvarese A and Chen H Y 2024 *Astrophys. J. Lett.* **974** L16
- [101] Cusin G and Tamanini N 2021 *Mon. Not. R. Astron. Soc.* **504** 3610–8
- [102] Mapelli M 2020 *Front. Astron. Space Sci.* **7** 2020
- [103] van Son L A C, de Mink S E, Callister T, Justham S, Renzo M, Wagg T, Broekgaarden F S, Kummer F, Pakmor R and Mandel I 2022 *Astrophys. J.* **931** 17
- [104] Mukherjee S 2022 *Mon. Not. R. Astron. Soc.* **515** 5495–505
- [105] Mastrogiovanni S, Pierra G, Perriès S, Laghi D, Santoro G C, Ghosh A, Gray R, Karathanasis C and Leyde K 2023 ICAROGW: a python package for inference of astrophysical population properties of noisy, heterogeneous and incomplete observations (arXiv:2305.17973)
- [106] Abbott R et al 2023 *Astrophys. J.* **949** 76
- [107] Pierra G, Mastrogiovanni S, Perriès S and Mapelli M 2024 *Phys. Rev. D* **109** 083504
- [108] Gennari V, Mastrogiovanni S, Tamanini N, Marsat S and Pierra G 2025 Searching for additional structure and redshift evolution in the observed binary black hole population with a parametric time-dependent mass distribution (arXiv:2502.20445)
- [109] Lalleman M, Turbang K, Callister T and van Remortel N 2025 No evidence that the binary black hole mass distribution evolves with redshift (arXiv:2501.10295)
- [110] Fragione G 2021 *Astrophys. J. Lett.* **923** L2
- [111] Vitale S and Chen H Y 2018 *Phys. Rev. Lett.* **121** 021303
- [112] Totani T 2013 *Publ. Astron. Soc. Japan* **65** L12
- [113] Wang J S, Yang Y P, Wu X F, Dai Z G and Wang F Y 2016 *Astrophys. J. Lett.* **822** L7
- [114] Dokuchaev V I and Eroshenko Y N 2017 Recurrent fast radio bursts from collisions of neutron stars in the evolved stellar clusters (arXiv:1701.02492)
- [115] Yamasaki S, Totani T and Kiuchi K 2018 *Publ. Astron. Soc. Japan* **70** 39
- [116] Pan Z, Yang H and Yagi K 2023 Repeating fast radio bursts from neutron star binaries: multi-band and multi-messenger opportunities (arXiv:2208.08808)
- [117] Most E R and Philippov A A 2023 *Phys. Rev. Lett.* **130** 245201
- [118] Abbott B P et al 2020 *Living Rev. Relativ.* **23** 3
- [119] Santoliquido F, Mapelli M, Giacobbo N, Bouffanais Y and Artale M C 2021 *Mon. Not. R. Astron. Soc.* **502** 4877–89
- [120] Abbott R et al (LIGO Scientific Collaboration, Virgo Collaboration and KAGRA Collaboration) 2023 *Phys. Rev. X* **13** 011048
- [121] Shah V G, Narayan G, Perkins H M L, Foley R J, Chatterjee D, Cousins B and Macias P 2023 *Mon. Not. R. Astron. Soc.* **528** 1109–24
- [122] Colangeli E, Leyde K and Baker T 2025 A bright future? prospects for cosmological tests of gr with multimessenger gravitational wave events (arXiv:2501.05560)
- [123] Chen H Y, Vitale S and Narayan R 2019 *Phys. Rev. X* **9** 031028
- [124] Ade P A R et al 2016 *Astron. Astrophys.* **594** A13
- [125] Farah A M, Edelman B, Zevin M, Fishbach M, María Ezquiaga J, Farr B and Holz D E 2023 *Astrophys. J.* **955** 107
- [126] Barausse E 2012 *Mon. Not. R. Astron. Soc.* **423** 2533–57
- [127] Barausse E, Dvorkin I, Tremmel M, Volonteri M and Bonetti M 2020 *Astrophys. J.* **904** 16
- [128] Mastrogiovanni S, Laghi D, Gray R, Santoro G C, Ghosh A, Karathanasis C, Leyde K, Steer D A, Perriès S and Pierra G 2023 *Phys. Rev. D* **108** 042002
- [129] Ye C S and Fishbach M 2024 *Astrophys. J.* **967** 6