

Lensing bias on cosmological parameters from bright standard sirens

Sofia Canevarolo,¹★ Nora Elisa Chisari¹†

¹ *Institute for Theoretical Physics, Utrecht University, Princetonplein 5, 3584 CC, Utrecht, The Netherlands.*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Next generation gravitational waves (GWs) observatories are expected to measure GW signals with unprecedented sensitivity, opening new, independent avenues to learn about our Universe. The distance-redshift relation is a fulcrum for cosmology and can be tested with GWs emitted by merging binaries of compact objects, called standard sirens, thanks to the fact that they provide the absolute distance from the source. On the other hand, fluctuations of the intervening matter density field induce modifications on the measurement of luminosity distance compared to that of a homogeneous universe. Assuming that the redshift information is obtained through the detection of an electromagnetic counterpart, we investigate the impact that lensing of GWs might have in the inference of cosmological parameters. We treat lensing as a systematic error and check for residual bias on the values of the cosmological parameters. We do so by means of mock catalogues of bright sirens events in different scenarios relevant to Einstein Telescope. For our fiducial scenario, the lensing bias can be comparable to or greater than the expected statistical uncertainty of the cosmological parameters, although non-negligible fluctuations in the bias values are observed for different realisations of the mock catalogue. We also discuss some mitigation strategies that can be adopted in the data analysis. Overall, our work highlights the need to model lensing effects when using standard sirens as probes of the distance-redshift relation.

Key words: gravitational waves – gravitational lensing: weak – cosmological parameters

1 INTRODUCTION

Since the first detection in 2015 by the LIGO-Virgo collaboration (Abbott et al. 2016), gravitational waves (GWs) have opened a new observational window on our Universe. As first observed by Schutz (1986), the gravitational wave signal from the inspiral and merger of a compact binary system carries information about our distance from the source. In fact, the amplitude of the GW signal from compact binary systems is inversely proportional to the luminosity distance of its source. This useful property turns merging binaries emitting GWs into *standard sirens* (Holz & Hughes 2005; Dalal et al. 2006), suitable for probing the expansion history of the Universe. This is of crucial importance to solve many open problems in cosmology, such as the current tension on the value of the Hubble constant H_0 in the Λ CDM model and to investigate possible extensions of the cosmological paradigm (see Di Valentino et al. 2021; Perivolaropoulos & Skara 2022, and references therein).

To probe cosmology with standard sirens, redshift information is also necessary. However, in the waveform of the standard sirens, the redshift is degenerate with a combination of the masses of the two objects in the binary system - the so-called mass-redshift degeneracy. The detection of an electromagnetic (EM) counterpart of the GW signal can provide an independent measurement of the redshift, breaking the degeneracy. A strategy for the detection of the EM counterpart can consist of an optical/IR follow-up in the localization area delimited by the GW detectors, aimed at the identification of the host galaxy, whose redshift can be measured spectroscopically or

photometrically. If the localization area of the GW detectors is very broad, or even in the case that no localization region is provided, it is possible to exploit the coincident detection of a gamma ray burst (GRB Dalal et al. 2006), and the corresponding X-ray/optical afterglow emission. In particular, the afterglow is important for the host galaxy identification.

In August 2017, the first coincident detection of the GW event GW170817 from the coalescence of two neutron stars and its EM counterpart was achieved (Abbott et al. 2017b), opening the era of multi-messenger cosmology. The EM counterpart, the GRB event, GRB170817A, followed by an afterglow in other frequencies of the EM spectrum, allowed the localization of the host galaxy of the source and led to the first constraint on the Hubble constant by GWs (Abbott et al. 2017a). At present, the strongest bound is $H_0 = 69^{+12}_{-7}$ km s⁻¹Mpc⁻¹ (Gray et al. 2023), obtained with the Third Gravitational-Wave Transient Catalogue (The LIGO Scientific Collaboration et al. 2021). These measurements, even if not yet comparable with the cosmological constraints already achieved by other EM probes, give proof that the distance-redshift relation, and thus cosmological parameters, can be independently tested with standard siren events.

In a decade from now, a third generation (3G) of ground- and space-based GW detectors, such as the Einstein Telescope (ET) (Punturo et al. 2010; Branchesi et al. 2023), Cosmic Explorer (CE) (Reitze et al. 2019; Evans et al. 2021) and LISA (Amaro-Seoane et al. 2017), are expected to measure GW signals with unprecedented sensitivity, reaching very high redshift events and accumulating an enormous number of detections. In this paper, we focus on the signals coming from Neutron Star binary (BNS) coalescence, which are the best candidate events to be accompanied by an EM counterpart, the so

★ E-mail: s.canevarolo@uu.nl

† E-mail: n.e.chisari@uu.nl

called bright sirens. The Einstein Telescope is expected to detect $\sim 10^5$ BNS per year up to redshifts of $z \simeq 2 - 3$ (Maggiore et al. 2020). A fraction of these events will be bright sirens and will allow for precise GW cosmology (Sathyaprakash et al. 2010). By the time ET enters its full operational regime, new EM facilities, such as *Euclid* (Laureijs et al. 2011) and the Vera C. Rubin Observatory (Ivezić et al. 2019), will have provided spectroscopic and photometric galaxy catalogues useful for the host galaxy determination.

Similarly to the case of light, GW signals are expected to be affected by gravitational lensing. If the alignment between the source, the lens and the observer is favorable, the gravitational waves can be strongly lensed, which can result in multiple images of the same signal. Searches for strongly lensed events have been carried out in LIGO-Virgo data but it is still a matter of discussion whether these are affected by lensing (Abbott et al. 2021b; Diego et al. 2021). The observation of such events would allow a plethora of fundamental and cosmological applications, widely explored in the literature (some examples are Sereno et al. 2011; Baker & Trodden 2017; Ezquiaga & Zumalacárregui 2020; Goyal et al. 2021). Strongly lensed events are expected to constitute a small fraction of the huge detection rates of the 3G detectors (Oguri 2018; Li et al. 2018). On the contrary, weak lensing caused by the gravitational potential of the large scale structures in the Universe is ubiquitous, and accumulates along the path of GWs, thus becoming more relevant for high redshift sources (Holz & Hughes 2005).

In this paper, we investigate the impact of lensing on the estimation of cosmological parameters from bright standard sirens measurements. We treat both weak and strong scenarios simultaneously based on the statistical distribution of values for the lensing magnification, even when it is very high. The prospect of the lensing of GWs is two-fold. It can be regarded as an extra source of information enabling the use of GWs to study the matter distribution along the line-of-sight, thus testing the standard model of cosmology and shedding light on the nature of dark energy (Congedo & Taylor 2019; Mukherjee et al. 2020; Mpetha et al. 2023; Balaudo et al. 2023). Or, it can be regarded as a nuisance in the determination of the distance-redshift relation with the standard sirens method. In such a case, it is normally accounted for as a source of noise in the distance measurement (Sathyaprakash et al. 2010; Hirata et al. 2010). Here we consider the latter view, but we treat lensing as a *systematic error* and we check whether the statistical uncertainty on the luminosity distance fully absorbs any bias of the cosmological parameters.

Two distinct aspects can source this bias: lensing selection effect and the biased estimation of the luminosity distance from the GW signal. The former effect, already studied in Cusin et al. (2019); Cusin & Tamanini (2021); Shan et al. (2021), is due to the fact that we are more likely to detect magnified events, whose Signal to Noise Ratio (SNR) is increased by the lensing magnification factor (Dai et al. 2017). In fact, lensing can act as a magnifying glass, enabling the detection of faint and far events, not accessible in a homogeneous universe. On the other hand, the non-Gaussian shape of the lensing magnification probability distribution function (PDF) might also yield a bias in the estimation of the luminosity distance. As we will see, the lensing PDF is skewed, with a tail towards high magnifications, produced by the rare overdensities of the matter field in the Universe. This highlights the importance of investigating lensing bias in the context of GW cosmology.

We adopt a flat FLRW cosmology with the following fiducial values of the parameters: $h = H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1} = 0.6736$, $\Omega_b h^2 = 0.02237$, $\Omega_c h^2 = 0.12$, where Ω_b and Ω_c are respectively the baryon and cold dark matter density parameters. These values correspond to those adopted in the AbacusSummit simulations of

Hadzhiyska et al. (2023), which we use to model gravitational lensing.

The paper is organised as follow. In Section 2, we review the theory of bright standard sirens and the impact of lensing. In Section 3, we build a realistic mock catalogue of lensed bright sirens events, detectable by a third generation ground-based detector, i.e. ET. We also compare the lensing magnification PDF obtained from lensing convergence maps with a simple log-normal analytical model, often used in the literature (Kayo et al. 2001; Das & Ostriker 2006; Martinelli et al. 2022). In Section 4, we present the methods adopted to estimate the bias on the cosmological parameters, which exploit the Fisher matrix formalism and the full Monte Carlo Markov Chain (MCMC) analysis. The results of our study are then presented in Section 5, where different scenarios are considered. In Section 6, we present a procedure to slightly mitigate the effect of lensing for the cosmological parameter estimation analysis. Finally, in Section 7, we present our discussion of the results and in Section 8 the conclusions.

2 LENSED STANDARD SIRENS

At leading order, in frequency domain and using the stationary phase approximation, the waveform of the inspiral of two compact objects can be written as (Cutler & Flanagan 1994),

$$\tilde{h}(f) = \sqrt{\frac{5}{96}} \frac{Q}{d_L} \frac{(GM(1+z))^{5/6}}{\pi^{2/3} c^{3/2}} f^{-7/6} e^{-i\Phi(f, \mathcal{M}(1+z))}, \quad (1)$$

with

$$Q = \sqrt{F_+^2(\theta, \phi, \psi)(1 + \cos^2 \iota)^2 + F_\times^2(\theta, \phi, \psi)4 \cos^2 \iota}. \quad (2)$$

$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}$ is the chirp mass of the binary composed by two individual objects of masses m_1 and m_2 , Φ the phase of the strain, G the Newton constant and c the speed of light. The function Q depends on the position of the binary in the sky, determined by the angles (θ, ϕ) , the inclination angle ι and the polarization angle ψ . $F_{+, \times}$ are antenna pattern functions that describe the response of the detector.

The luminosity distance of the source, d_L , is inversely proportional to the amplitude of the strain and is degenerate with the angles in the amplitude of the strain, in particular with the inclination angle ι (Nissanke et al. 2010). In a flat FLRW spacetime, d_L is given by:

$$d_L(z) = \frac{c(1+z)}{H_0} \int_0^z \frac{dz}{\sqrt{\Omega_m(1+z)^3 + \Omega_{DE}(1+z)^{3(1+w_{DE})}}}, \quad (3)$$

where Ω_m is the energy density in matter (baryons plus dark matter), Ω_{DE} is the energy density in dark energy, and w_{DE} is the dark energy equation of state parameter. In a flat universe, $\Omega_{DE} = 1 - \Omega_m$. For a cosmological constant, $w_{DE} = -1$.

In equation (1), the redshift of the source z is degenerate with the chirp mass, such that only the redshifted chirp mass, $\mathcal{M}_z = (1+z)\mathcal{M}$, can be measured from a GW signal. Additional information is needed to break this degeneracy. In this paper we assume that the redshift can be determined from the detection of an EM counterpart that allows the localization of the host galaxy of the compact binary coalescence. If the counterpart is a strongly beamed GRB emission, with $\iota \lesssim 20^\circ$, Nissanke et al. (2010) pointed out that it can provide a prior on the inclination angle which helps break the $d_L - \iota$ degeneracy. Alternatively, the measurement of higher order modes of the GW signals can also alleviate this degeneracy (Calderón Bustillo et al. 2021).

Possible GW sources that can be accompanied by an EM counterpart are BNS, a neutron star and a black hole (BHNS) and super-massive black holes (SMBH). SMBHs are expected to be detectable in the low frequency band covered by space-based detectors, such as LISA. In this paper, we choose the framework of ground-based detectors, in particular ET and we concentrate on the population of BNSs. We do not consider BHNS mergers, as their multi-messenger rate is quite uncertain (Fragione 2021). Nevertheless, since the lensing universally affects all GW signals, we expect our methodology could be extended to other GW sources, in other frequency bands, with or without an EM counterpart.

2.1 Lensing magnification

As studied in Laguna et al. (2010); Bertacca et al. (2018); Fonseca et al. (2023), different modifications of the GW waveform are expected due to the fluctuations of the intervening matter density field, which induce effects such as lensing, Doppler, Integrated Sachs-Wolfe, Sachs-Wolfe, Shapiro time-delays and volume distortions. In this paper, we only model the effect of lensing on the luminosity distance, which is expected to be dominant for $z \gtrsim 0.5$. The geometric optics approximation, valid when the wavelength of the GW is much smaller than the size of the lens, predicts modifications that are frequency independent, i.e. the frequency of the lensed waves remain unchanged. Within this approximation, we expect the strain of the gravitational waves to be magnified or demagnified by the structures or voids the GWs encounter along their propagating path:

$$h(f) = \sqrt{\mu} h_{\text{unlensed}}(f), \quad (4)$$

where $h_{\text{unlensed}}(f)$ is the waveform computed in the homogeneous universe, e.g. it satisfies equation (1), and we dropped the tilde that signals the frequency domain for simplicity. This modification translates into an observed luminosity distance d_L^{bias} which differs with respect to the true one by (Kocsis et al. 2006):

$$d_L^{\text{bias}}(z, \mu) = \frac{d_L(z)}{\sqrt{\mu}}, \quad (5)$$

where μ is the lensing magnification. At lowest order, the magnification is related to the lensing convergence field κ by $\mu \simeq 1 + 2\kappa$. But more generally, the relation between convergence and magnification can be more accurately quantified by (Takahashi et al. 2011, Figure 8),

$$\mu = \frac{1}{(1 - \kappa)^2}, \quad (6)$$

which is derived from the definition of magnification, neglecting the effect of the shear. In this work, we use equation (6) to allow for arbitrarily high values of μ .

It is interesting to note that there exists another regime of extreme magnification caused by the presence of microlenses along the line-of-sight. We do not attempt to model this regime explicitly in this work since we assume they could be separately identified from wave optics effects (Takahashi & Nakamura 2003; Diego 2019; Mishra et al. 2021; Kim et al. 2022).

3 MOCK CATALOG

Assuming the redshift is known from the observation of an EM counterpart, we build a mock catalogue of BNS events following the steps:

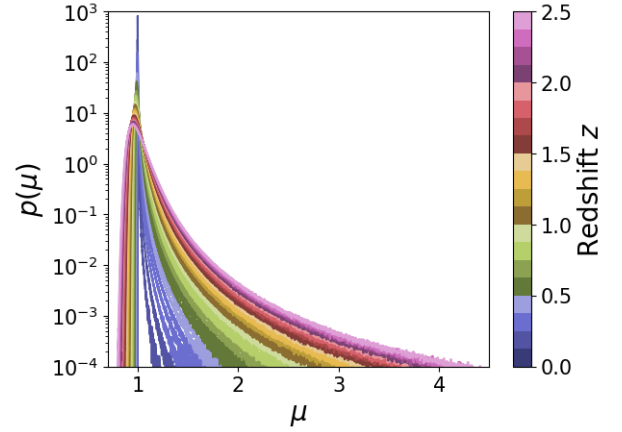


Figure 1. Lensing magnification PDF for different source redshift, as obtained from the convergence maps of Hadzhiyska et al. (2023).

(i) We draw the redshift of the events from a distribution built assuming that the BNS merger rate tracks the Madau-Fragos star formation rate (Madau & Fragos 2017), and accounting for a time delay t_d between the redshift of formation z_f and of merger z_m (Regimbau et al. 2012; Muttoni et al. 2023). The star formation rate is given by:

$$\Psi(z) = 0.01 \frac{(1+z)^{2.6}}{1 + [(1+z)/3.2]^{6.2}} [\text{M}_\odot \text{Gpc}^{-3} \text{yr}^{-1}], \quad (7)$$

and we assume the following probability distribution for the time delay,

$$p(t_d) \propto \frac{1}{t_d}. \quad (8)$$

The theoretical merger rate is then:

$$p^{(\text{th})}(z) = \frac{1}{1+z} \frac{4\pi c \chi^2(z)}{H(z)} \int_{t_d^{\text{MIN}}}^{t_d^{\text{MAX}}} dt_d \Psi(z_f(z, t_d)) \times \left(\log \left(\frac{t_d^{\text{MAX}}}{t_d^{\text{MIN}}} \right) t_d \right)^{-1}, \quad (9)$$

where $\chi(z) = c \int_0^z \frac{dz}{H(z)}$ is the comoving distance and $H(z) = H_0 \sqrt{\Omega_m (1+z)^3 + \Omega_{\text{DE}}}$ is the Hubble parameter. We assume a minimum time delay between formation and merger $t_d^{\text{MIN}} = 20$ Myr and a maximum $t_d^{\text{MAX}} = 10$ Gyr (Borhanian & Sathyaprakash 2022).

(ii) Following Mpettha et al. (2023), we account for selection effects due to the coverage of galaxy surveys. We obtain the observed probability distribution for the source redshift as:

$$p(z) = p^{(\text{th})}(z) f(z) \quad (10)$$

where $f(z)$ is the selection function of the galaxy survey and takes the form,

$$f(z) = \frac{0.9}{2} \left(1 - \tanh \left(\frac{z - z_{\text{pivot}}}{w z_{\text{pivot}}} \right) \right), \quad w = \frac{z_{\text{max}} - z_{\text{pivot}}}{2z_{\text{max}}}, \quad (11)$$

with $z_{\text{pivot}} = 1.8$ and $z_{\text{max}} = 2$. This is based on the maximum redshift coverage that future spectroscopic galaxy surveys are expected to reach (Laureijs et al. 2011). Beyond this limit, we assume that the redshift of the source cannot be estimated via spectroscopy.

We draw a set of $\{\hat{z}\}$ from the distribution $p(z)$. In Section 5, we will also study the dependence of our results on the galaxy survey selection function.

(iii) We assign to each event a value of the lensing magnification. Specifically, for each redshift \hat{z}_i in the collection $\{\hat{z}\}$, we assign a value $\hat{\mu}_i$, by drawing from the magnification PDF at redshift \hat{z}_i .

More details about the magnification PDF are given in subsection 3.1. For our base analysis, we consider the AbacusSummit suite of cosmological N-body simulations, presented in Hadzhiyska et al. (2023), and the publicly available attached lensing distributions¹. We use the “base” box with a pixel resolution of 0.21 arcmin and a coverage of an octant of the sky, up to $z \sim 0.8$, then reduced to 1800 deg². We obtain the magnification PDF from the maps of the convergence fields at all source redshifts, 47 in total, ranging from $z = 0.15$ to $z = 2.45$, and applying eq. (6). The result is shown in Fig. 1. As already said, we do not attempt to distinguish between weak and strong lensing regime, and we allow for a tail of pretty high magnifications, for example due to unidentified strongly lensed images. This can be the consequence of a single image event or of multiple images that arrive at the detector with long time delays and are recorded as separate events.

(iv) We draw values for the masses of the neutron stars from a redshift independent distribution with a Gaussian shape centered in $\bar{m} = 1.35 M_\odot$ and with standard deviation $\sigma = 0.15 M_\odot$, normalised in the interval $[1; 2]M_\odot$ (Borhanian & Sathyaprakash 2022). We also consider the case of a uniform distribution between $[1; 2.5]M_\odot$ (Belgacem et al. 2019; Ronchini et al. 2022).

(v) We draw the angles θ and ι from isotropic distributions in the range $[0, \pi]$. The polarisation angle ψ is drawn from a uniform distribution in the range $[0, \pi]$. If we assume that the EM counterpart is a beamed GRB, it might be possible to constrain the value of the inclination angle ι . Therefore, we also consider the case in which we restrict $\iota < 20^\circ$.

(vi) Using the mock collection of $\{\hat{z}, \hat{\mu}\}$ and our fiducial values of the cosmological parameters we compute the luminosity distance of each event. We assume to infer the observed luminosity distance from the full parameters estimation analysis over all the extrinsic and intrinsic parameters that characterise the GW signal. The luminosity distance obtained from this analysis is biased by the effect of lensing, as given in equation (5).

(vii) For each event in our theoretical mock catalog, we calculate the SNR ρ , defined by,

$$\rho^2 = 4 \int df \frac{|h(f)|^2}{S_n(f)}, \quad (12)$$

where $S_n(f)$ is the one-sided noise power spectral density (PSD) of a detector. In the case of ET, we have to sum over the three detectors that compose the triangular configuration. We assume the same PSD function, the ET-D sensitivity curve², for all the three detectors. In absence of correlated noise, the total SNR is:

$$\rho_{\text{tot}}^2 = \sum_{a=1}^3 \rho_a^2, \quad (13)$$

where ρ_a^2 is the SNR of the individual detector. A realistic GW interferometer is sensitive to SNRs above ρ_{lim} . We assume $\rho_{\text{lim}} = 8$, summed over the three independent detectors of ET (Zhao et al. 2011; Regimbau et al. 2015), and we create a mock catalogue of detectable events above such threshold.

For BNS events, which involve low mass objects, we can use the following approximated formula, valid during the inspiral part of

the coalescence, to compute the SNR of the event (Finn & Chernoff 1993):

$$\rho_{\text{tot}}^2 = \mu \frac{5}{6} \frac{(GM(1+z))^{5/3}}{c^3 \pi^{4/3} d_L^2(z)} \mathcal{F}^2(\theta, \iota, \psi) \int_{f_{\text{min}}}^{f_{\text{max}}} df \frac{f^{-7/3}}{S_n(f)}, \quad (14)$$

with (Regimbau et al. 2012):

$$\begin{aligned} \mathcal{F}^2(\theta, \iota, \psi) = & \frac{9}{128} (1 + \cos^2 \iota)^2 (1 + \cos^2 \theta)^2 \cos^2 2\psi \\ & + \frac{9}{32} (1 + \cos^2 \iota)^2 \cos^2 \theta \sin^2 2\psi \\ & + \frac{9}{32} \cos^2 \iota (1 + \cos^2 \theta)^2 \sin^2 2\psi \\ & + \frac{9}{8} \cos^2 \iota \cos^2 \theta \cos^2 2\psi. \end{aligned} \quad (15)$$

We take the upper integration bound f_{max} to be $f_{\text{LSO}} = \frac{1}{6\sqrt{6}(2\pi)} \frac{c^3}{G(1+z)(m_1+m_2)}$, assumed as the frequency at the last stable orbit. The lower integration bound is $f_{\text{min}} = 1\text{Hz}$.

Notice that we included the magnification factor μ in the calculation of the SNR. This is how we account for selection effects induced by lensing in our analysis. Fig. 2 shows the redshift and magnification distributions in a realisation of our mock catalog. In the lower panel, we also highlight the difference in the mock catalogue distributions that we obtain by removing μ from equation (14), and then cutting the events with SNR below the detector threshold.

In our fiducial scenario, the mock catalogue is comprised of 3,000 multi-messenger events, which we assume can be collected over a ~ 10 years operational time of ET, together with EM facilities for the redshift determination (Sathyaprakash et al. 2010; Yu et al. 2021). We also consider a more optimistic scenario in which the mock catalogue contains 30,000 bright sirens.

3.0.1 d_L uncertainty

We assume that the error on the luminosity distance is given by (Mpetha et al. 2023),

$$\sigma_{d_L}^2 = \sigma_{\text{GW}}^2 + \sigma_{\text{WL}}^2 + \left(\frac{\partial d_L}{\partial z} \right)^2 (\sigma_z^2 + \sigma_{v_{\text{pec}}}^2), \quad (16)$$

where σ_{GW} is the instrumental uncertainty of the gravitational wave detector, $\sigma_z = 0.001(1+z)$ is the spectroscopic redshift uncertainty and $\sigma_{v_{\text{pec}}} = \sqrt{\langle v_{\text{pec}}^2 \rangle} (1+z)/c$ accounts for the source peculiar velocity with $\sqrt{\langle v_{\text{pec}}^2 \rangle} = 500 \text{ km s}^{-1}$. The uncertainty due to weak lensing σ_{WL} is estimated in Hirata et al. (2010) by the following fitting formula:

$$\sigma_{\text{WL}} = 0.066 \left(\frac{1 - (1+z)^{-0.25}}{0.25} \right)^{1.8} d_L. \quad (17)$$

This uncertainty accounts for the (de)magnification of the luminosity distance inferred from a lensed gravitational wave. Our goal is to check whether this uncertainty fully absorbs any bias of the cosmological parameters.

In the analysis of actual data, the instrumental uncertainty σ_{GW} , together with the uncertainties on all the parameters that enter the GW waveform, is obtained by means of a Bayesian analysis of the GW likelihood. A quicker and efficient, although approximate way to assess the measurement accuracy of ET is to use a Fisher matrix analysis of all the GW parameters, e.g. see Pieroni et al. (2022). Since for our analysis only the uncertainty on the luminosity distance is needed, we make some simplifying assumptions on the value of

¹ https://app.globus.org/file-manager?origin_id=3dd6566c-eed2-11ed-ba43-09d6a6f08166&path=%2F

² <https://www.et-gw.eu>

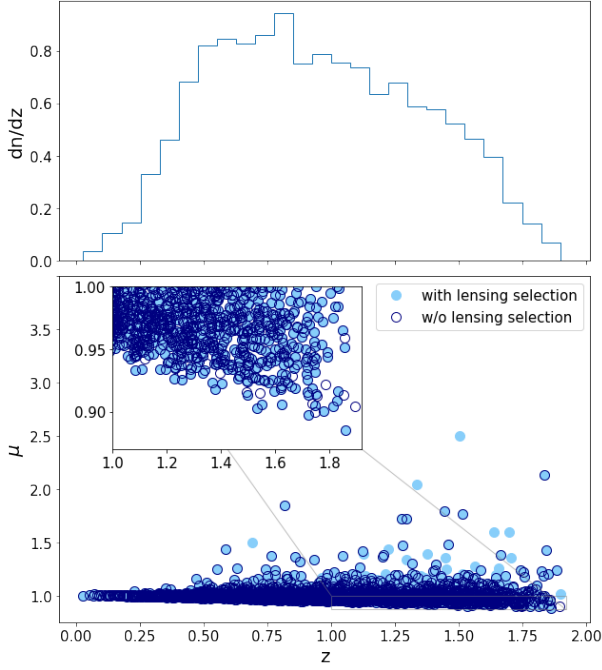


Figure 2. Upper plot: Redshift distribution of the mock catalogue of events. Lower plot: pairs of $\{\hat{z}, \hat{\mu}\}$ values. In both plots, we are showing $N_{\text{events}} = 3,000$, with SNR greater than ρ_{lim} . The SNR of the light blue filled circles is calculated including the magnification factor μ , i.e. with equation (14). Conversely, the SNR of the dark blue unfilled circles is independent of the magnification μ . The redshift distributions of the two cases are very similar, so we choose to show the histogram obtained accounting for the lensing selection effect in the upper plot.

σ_{GW} : for our base model we assume that $\sigma_{\text{GW}} = 0.02d_L$ (Mpeha et al. 2023). We consider also a less optimistic value of $\sigma_{\text{GW}} = 0.1d_L$. Moreover, if we assume that the uncertainties on d_L are uncorrelated with the errors of all the other GW parameters, we can use the formula $\sigma_{\text{GW}} = d_L/\rho$, which can be found inverting the Fisher matrix and marginalising over all the other GW parameters. This scenario is realistic in the cases when the degeneracy between luminosity distance and inclination angle can be broken, for example by the coincident detection of a GRB event. On the contrary, if the degeneracy is present, a more reliable estimate of the uncertainties is given by $\sigma_{\text{GW}} = 2d_L/\rho$ (Dalal et al. 2006; Belgacem et al. 2019).

3.1 Lensing PDF

Many different simulation suites, fitting formulae and analytical models of the PDF of κ and μ are available in the literature (see Das & Ostriker 2006; Boyle et al. 2021; Fosalba et al. 2014; Xavier et al. 2016; Takahashi et al. 2017; Gouin et al. 2019; Osato et al. 2021, as some examples). We are interested in quantifying how modelling choices in the lensing PDF can influence our results. The importance of utilising an accurate magnification PDF in the framework of standard sirens was already highlighted in Shang & Haiman (2011). The tail of highly magnified events in particular is envisaged to be difficult to capture. The resolution of the convergence maps, for example, can influence the tail and the skewness of the distribution. In fact, lowering the resolution of convergence maps has the effect of losing the rare overdensities that produce the high magnified events. Similarly, log-normal models have also the tendency of underestimating the probability of high magnifications. We also compare the

results obtained with the N-body simulations with a simple analytical log-normal model.

3.1.1 Log-Normal analytical model

We consider a probability distribution for the magnification μ , as given in Martinelli et al. (2022, Appendix C). The shape of the distribution at fixed redshift is:

$$p(\mu) = \frac{1}{\sqrt{2\pi}\sigma(\mu - \mu_{\min})} \exp\left[-\frac{(\ln(\mu - \mu_{\min}) - m)^2}{2\sigma^2}\right]. \quad (18)$$

The parameters σ and m of the distribution are redshift dependent. For a magnification distribution with mean $\langle\mu\rangle$ and variance σ_μ^2 , the parameters of the log-normal distribution are:

$$\sigma(z) = \sqrt{\log\left(\frac{\sigma_\mu^2(z)}{(\langle\mu\rangle - \mu_{\min}(z))^2} + 1\right)}, \quad (19)$$

$$m(z) = \log(\langle\mu\rangle - \mu_{\min}(z)) - \frac{\sigma^2(z)}{2}. \quad (20)$$

We assume for the base model $\langle\mu\rangle = 1$ and $\sigma_\mu^2(z) = 4\sigma_\kappa^2(z)$, where the variance of the convergence is computed from the convergence power spectrum at fixed redshift, obtained from PYCCL 2.6.1³ (Chisari et al. 2019). μ_{\min} is the minimum magnification corresponding to the case when light propagates in the empty space. It is calculated by:

$$\mu_{\min}(z) = \left(\frac{d_L(z)}{c(1+z)\lambda(z)}\right)^2, \quad \lambda(z) = \int_0^z dz \frac{1}{(1+z)^2 H(z)}. \quad (21)$$

4 ANALYSIS METHOD

We perform a parameter estimation analysis, using our mock catalogue of bright sirens, and study the results obtained neglecting the impact of lensing. We then assess whether the fiducial values of the cosmological parameters are correctly recovered, or whether lensing introduces a bias.

4.1 Fisher matrix

Applying a Fisher matrix framework (see Euclid Collaboration et al. 2020, section 3), we combine the information from multiple GW standard sirens to compute the expected uncertainties on the cosmological parameters $\vec{\theta} = (H_0, \Omega_m, \dots)$. In principle, the cosmological parameters should be inferred jointly with the astrophysical parameters which describe the BNS population. For simplicity, we keep the latter fixed, and only investigate the effect of lensing on the cosmological parameters. We adopt the following Fisher matrix (Zhao et al. 2011):

$$\Gamma_{ij} = \sum_{\text{events}} \frac{1}{\sigma_{d_L}^2(z)} \frac{\partial d_L(z)}{\partial \theta_i} \frac{\partial d_L(z)}{\partial \theta_j}, \quad (22)$$

which neglects the dependence of the error $\sigma_{d_L}(z)$ on the cosmological parameters.

Inverting the Fisher matrix Γ_{ij} , we obtain the errors on the cosmological parameters, θ_i , by,

$$\sigma_{\theta_i} = \sqrt{[\Gamma^{-1}]_{ii}} \quad (23)$$

³ <https://github.com/LSSTDESC/CCL>

This error includes the contribution due to the degeneracies with all the other parameters of the Fisher matrix. If we assume that one parameter is perfectly known, we can remove its corresponding row and column from the Fisher matrix, before inverting it.

4.1.1 Systematic error

We estimate the bias introduced by neglecting the lensing by employing the formalism of [Amara & Réfrégier \(2008\)](#). In our case, the systematic error comes from the biased luminosity distance inferred from the GW event:

$$\Delta d_L^{\text{sys}} = d_L^{\text{bias}} - d_L = \left(\frac{1}{\sqrt{\mu}} - 1 \right) d_L. \quad (24)$$

We compute the “bias vector” as:

$$B_j = \sum_{\text{events}} \frac{1}{\sigma_{d_L}^2} \Delta d_L^{\text{sys}} \frac{\partial d_L}{\partial \theta_j}. \quad (25)$$

Finally, the bias on the cosmological parameters is obtained from:

$$b_{\theta_i} = \hat{\theta}_i - \theta_i^{\text{true}} = [\Gamma^{-1}]_{ij} B_j, \quad (26)$$

where we are summing over repeated indices. The derivation of these expressions relies on a first order Taylor expansion of the likelihood, valid for small values of the bias. For larger biases, a study of the full likelihood is more appropriate.

4.2 MCMC analysis

For some interesting cases, we check our results with a full Monte Carlo Markov Chain analysis ([Hogg & Foreman-Mackey 2018](#)). Assuming a Gaussian likelihood for the cosmological parameters $\mathcal{L}(H_0, \Omega_m)$, we fit our mock catalogue of lensed bright sirens to the theoretical value of the luminosity distance,

$$\log \mathcal{L}(H_0, \Omega_m) = -\frac{1}{2} \sum_{\text{events}} \frac{1}{\sigma_{d_L}^2} \left(d_L^{\text{bias}} - d_L(H_0, \Omega_m) \right)^2, \quad (27)$$

where d_L^{bias} is constructed with the $\{\hat{z}, \hat{\mu}\}$ values in our mock catalog, while the theoretical d_L using the redshifts $\{\hat{z}\}$ only. For this, we use the EMCEE software ⁴ ([Foreman-Mackey et al. 2013](#)).

Moreover, to check whether the fiducial values of the cosmological parameters are correctly recovered, we also fit the mock catalogue to the theoretical luminosity distance rescaled by the lensing magnification of each event $\frac{d_L(H_0, \Omega_m)}{\sqrt{\mu}}$.

5 RESULTS

We perform the analysis described in Section 4 using the fiducial realisation of our mock catalog, built as in Section 3. We then consider some variations of the characteristics of the catalog.

5.1 Fiducial scenario

The redshift distribution of the bright sirens events in our fiducial scenario is obtained imposing the EM selection function in equation (11), which introduces a maximum redshift of $z = 2$, and we assume an instrumental uncertainties on the luminosity distance of $\sigma_{\text{GW}} =$

$0.02 d_L$. In this scenario, we draw the binary inclination angle from an isotropic distribution in the range $[0, \pi]$, thus allowing the possibility to detect off-axis events. We then impose the detection of beamed GRB events by restricting $\iota \lesssim 20^\circ$.

As a first step, we constrain the parameter space of the Hubble constant and the matter density parameter: (H_0, Ω_m) , which are the only parameters relevant for standard sirens within a flat Λ CDM model. We report the results for 50 different realisations of the mock catalog. For each realisation, we draw again the distributions of redshifts, magnifications, angles and masses. The result is obtained using the Fisher matrix analysis presented in Subsection 4.1, with the derivatives in equation (22) and (25) calculated analytically for stability. While we find that the statistical errors, obtained from the inverse of the Fisher matrix, are stable for different realisations of the mock catalog, the bias fluctuates over more than one order of magnitude. (As more events are considered, the value of the bias becomes more stable, but here we are restricted to 3,000 in the fiducial catalogue.)

At 1σ level, the statistical accuracies on H_0 and Ω_m are respectively 0.16% and 1.12%. This is slightly smaller than the results obtained in [Cai & Yang \(2017\)](#); [Belgacem et al. \(2019\)](#); [Jin et al. \(2020\)](#), but we emphasise that they are making different assumptions when constructing the mock catalog, in particular with a higher instrumental uncertainty on the luminosity distance and a smaller number of events, which justify the difference.

In the left panel of Fig. 3, we show the ratio between the absolute value of the bias and the statistical error on H_0 (upper panel) and Ω_m (lower panel), for each realisation of the mock catalog. It is worth noting that the error bars reported in this and the following plots represent the marginalised uncertainties on the corresponding cosmological parameters σ_{θ_i} , obtained from the inverse of Fisher matrix, as in equation (23). In the right panels of Fig. 3, we display the histograms obtained with the ratios of the bias and the statistical uncertainty. We find that lensing induces a negative (positive) systematic error for H_0 (Ω_m); in fact, for both the cosmological parameters, the values of the bias do not oscillate around zero. While for H_0 , we find that the absolute mean value of bias has the same order of magnitude than the statistical uncertainty, for Ω_m the peak of the distribution of the ratio $b_{\Omega_m}/\sigma_{\Omega_m}$ is around ~ 2 , signaling that, in this scenario, the systematic error induced by lensing can dominate the statistical one.

The amplification of the SNR for magnified events makes them more likely to exceed the SNR threshold with respect to the demagnified ones. This can also be noted in the lower panel of Fig. 2. The lensing selection effect tends to be more important at high redshift, when the probability for highly magnified events increases. Magnified events look closer than they are, and their biased luminosity distance is better fit by higher values of Ω_m . We check the amount of contribution to the bias given by the lensing selection effect. To do so, we set $\mu = 1$ in the calculation of the SNR, in equation (14), which corresponds to neglecting the selection effect of lensing, and we find that the bias on both H_0 and Ω_m is reduced to a value equal or less than 1σ for all realisations of the mock catalog. Furthermore, for the different realisations we find that the bias fluctuates around zero, showing both negative and positive values; as a consequence the average ratios becomes $b_{H_0}/\sigma_{H_0} \sim 0.2$ and $b_{\Omega_m}/\sigma_{\Omega_m} \sim -0.1$. We conclude that the bias induced by the selection effect alone contributes the most to our results.

5.1.1 Parameter space

In Fig. 4, we show the results obtained with the same set of realisations of the mock catalog, but assuming Ω_m is perfectly known. This

⁴ <https://emcee.readthedocs.io/en/stable/>

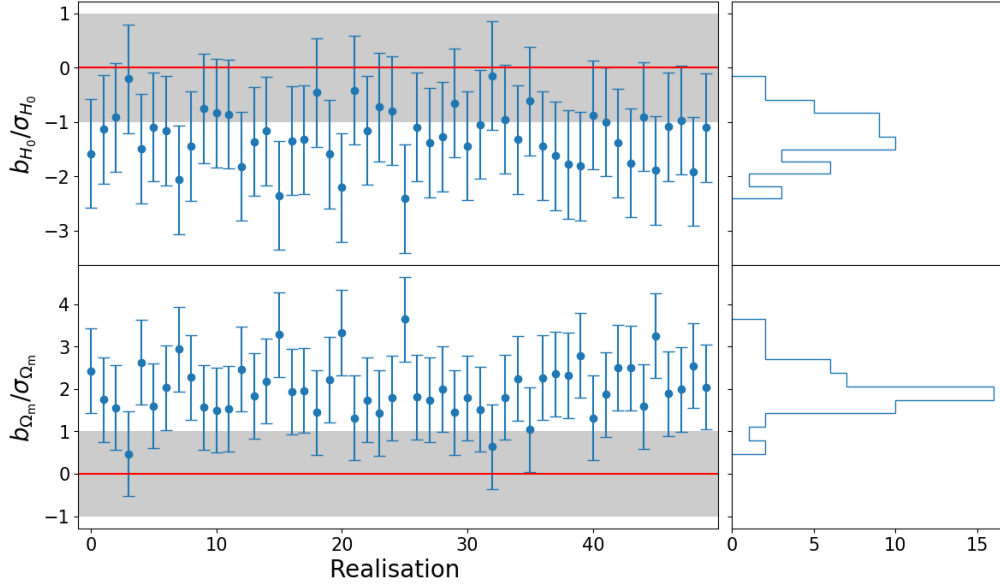


Figure 3. In the left panels: ratio of the bias and statistical error on H_0 (upper panel) and Ω_m (lower panel) for 50 different realisations of the mock catalog, each with $N_{\text{events}} = 3,000$. The gray bands highlight the region of 1σ significance from zero. The right panels show the histograms of the ratios.

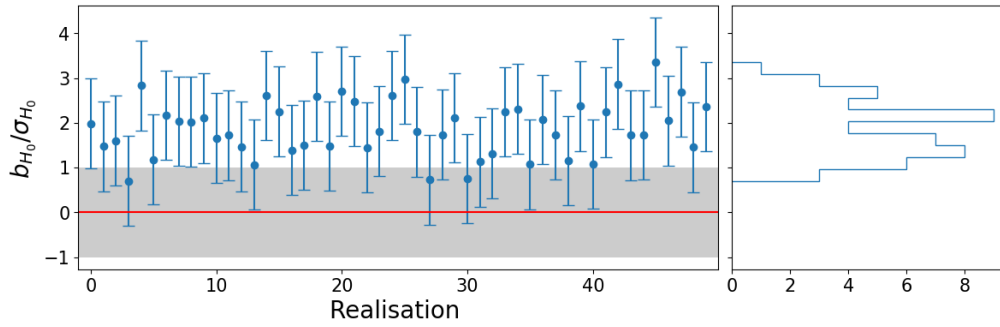


Figure 4. In the left panel: ratio of the bias and statistical error on H_0 , assuming that Ω_m is perfectly known, for 50 different realisations of the mock catalog, each with $N_{\text{events}} = 3,000$. The gray bands highlight the region of 1σ significance from zero. The right panel shows the histogram of the ratios.

is a common assumption in recent studies of standard sirens using the real data of LIGO/Virgo (Abbott et al. 2017a; Palmese et al. 2020). In our 3G scenario, we can imagine that the matter density parameter will be already very well constrained by other EM probes, while the Hubble constant might still suffer of some residual tension. It is then interesting to check at what level lensing could impact the Hubble tension. As expected, we find that the absolute magnitude of the bias is larger than in the previous case, with a peak of the distribution in the right panel of Fig. 4 around ~ 2 . Unlike the previous case, the bias on H_0 is positive. This is consistent with the fact that selection effect induced by lensing gives, on average, a smaller inferred luminosity distance. Given that Ω_m is fixed, the freedom to better fit the data is only left to H_0 , which is inversely proportional to d_L , as in equation (3).

Finally, we consider the possibility to use GWs to test extensions of the Λ CDM model, in particular the nature of dark energy. Our fiducial value of the dark energy equation of state parameter is $w_{\text{DE}} = -1$, which corresponds to the case of a cosmological constant. We are interested in knowing whether lensing can prevent us from recovering the correct value of w_{DE} . We underline that our aim is not studying the impact of lensing in alternative theory of gravity,

since the lensing simulations suite Hadzhiyska et al. (2023), we are using, was carried in a Λ CDM cosmology. In this work, we simply check whether lensing can bias the fiducial values of w_{DE} within a Λ CDM cosmology, thus risking drawing incorrect conclusions from the data. We show the result of this analysis in Fig. 5, for the same 50 realisations of the mock catalog. We find that the average ratios are $b_{H_0}/\sigma_{H_0} \sim 0.6$, $b_{\Omega_m}/\sigma_{\Omega_m} \sim 1.9$ and $b_{w_{\text{DE}}}/\sigma_{w_{\text{DE}}} \sim -1.1$. As can be seen from the lower panel in Fig. 5, for some realisations of the catalog, the lensing induced bias on w_{DE} can have a significance of more than 2σ , highlighting the necessity that additional care is needed in order to use GWs to shed light on the nature of dark energy. Once again, we also notice that by changing the parameter space, we find that the bias moves to different directions, being now positive for both H_0 and Ω_m and negative for w_{DE} .

5.2 Variations from the fiducial scenario

To assess the impact of our assumptions on the previous results, we vary one assumption at a time and leave all the other characteristics of the mock catalogue equal to what was described for the fiducial scenario.

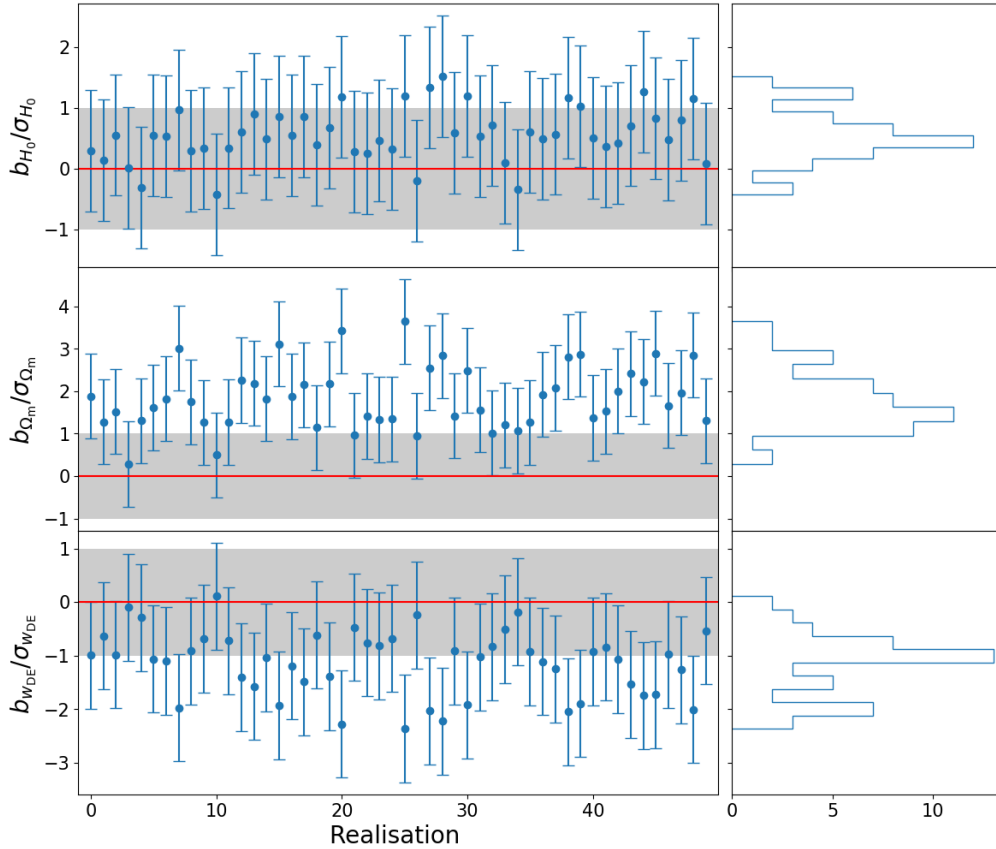


Figure 5. In the left panels: ratio of the bias and statistical error on H_0 (upper panel) and Ω_m (central panel) and w_{DE} (lower panel) for 50 different realisations of the mock catalog, each with $N_{\text{events}} = 3,000$. The gray bands highlight the region of 1σ significance from zero. The right panels show the histograms of the ratios.

We start from the instrumental uncertainties on d_L , and we repeat the analysis assuming that $\sigma_{GW} = 2d_L/\rho$, as done in Dalal et al. (2006). In this analysis, we pick one realisation of the mock catalog, chosen between the 50 different realisations already studied in the fiducial scenario, as a representative example; we consider in particular the realisation corresponding to the first point in the left panels of Fig. 3, which for the remainder of the discussion we call r_0 . Assuming $\sigma_{GW} = 0.02d_L$, we obtained $b_{H_0}/\sigma_{H_0} \sim -1.5$ and $b_{\Omega_m}/\sigma_{\Omega_m} \sim 2.7$. Instead, with the choice $\sigma_{GW} = 2d_L/\rho$ we find that the statistical error increases to 0.5% and 3.8% respectively for H_0 and Ω_m , while the absolute values of the systematic bias remain almost unchanged. As a consequence, for this specific realisation of the mock catalogue with $\sigma_{GW} = 2d_L/\rho$ the ratios decrease to $b_{H_0}/\sigma_{H_0} \sim -0.4$ and $b_{\Omega_m}/\sigma_{\Omega_m} \sim 0.8$. Similar values can be found assuming $\sigma_{GW} = 0.1d_L$, more details about this case are reported in Appendix A. In Fig. 6, we display the comparison between the Fisher confidence ellipses obtained utilising the two different assumptions on σ_{GW} discussed above. Notice that the fiducial ellipses are centered at the point corresponding to our fiducial cosmology, while the biased ellipses are shifted by the lensing bias, calculated as in Subsection 4.1.1. Alternatively, one might consider $\sigma_{GW} = d_L/\rho$, which corresponds to the Fisher matrix error obtained marginalising over all the other GW parameters. In this case, we find $b_{H_0}/\sigma_{H_0} \sim -0.8$ and $b_{\Omega_m}/\sigma_{\Omega_m} \sim 1.4$.

We also explore how some assumptions in the distributions of parameters that characterise the GW event impact our results. We draw 50 new realisations of the mock catalogue and quote the average bias

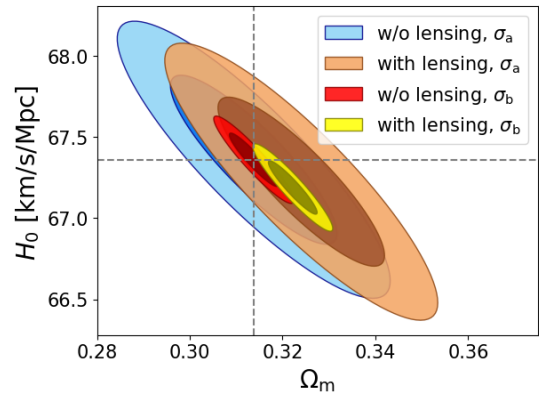


Figure 6. 1σ and 2σ ellipses for H_0 and Ω_m , calculated using the realisation r_0 of the mock catalogue and changing the assumptions on the instrumental uncertainty on d_L : $\sigma_a = 2d_L/\rho$ and $\sigma_b = 0.02d_L$. The fiducial ellipses, labelled with "w/o lensing", are centered at the correct cosmology, while the biased ones ("with lensing") show the systematic displacement. The intersection of the dotted lines represents the fiducial values of the parameters in standard Cosmology.

in Table 1. Future high-energy detectors, such as THESEUS (Amati et al. 2018; Stratta et al. 2018) and TAP (Camp 2017), are expected to detect GRBs at redshifts higher than ~ 2 . Therefore we might assume that all the events in our simulated catalogue will have, besides the GRB emission, a detectable afterglow, useful for the redshift de-

Table 1. Values of the average ratios of b_{H_0}/σ_{H_0} , $b_{\Omega_m}/\sigma_{\Omega_m}$ and 1σ uncertainties, for three variations of our fiducial scenario: removing the EM selection function given in eq. (11) (first row), assuming a uniform mass distribution for the two NSs the the binary system (second row) and restricting the inclination angle to be smaller than 20° (third row). The results are obtained averaging over 50 realisations of the mock catalog.

Model	$\frac{b_{H_0}}{\sigma_{H_0}}$	σ_{H_0} [Km/s/Mpc]	$\frac{b_{\Omega_m}}{\sigma_{\Omega_m}}$	σ_{Ω_m}
w/o $f(z)$	-2.2	0.1	3.3	0.003
NS $p(m_{1/2})$ uniform	-1.0	0.1	1.6	0.004
$\iota < 20^\circ$	-0.6	0.1	1.0	0.004

termination. This would justify removing the EM selection function in equation (11). The resulting redshift distribution shows a tail of events up to redshift ~ 3 . As given in Table 1, on average, the absolute values of the bias increases for both cosmological parameters, H_0 and Ω_m . This is expected since the few events in the tail of high redshift do not properly sample the lensing PDF and, in addition, they are more likely to be highly magnified. Given that the Abacus-Summit simulations cover up to $z = 2.45$, we have to extrapolate the PDF to higher redshifts. Since we have only a few events above this maximum redshift of $z = 2.45$, they mostly fall near the peak of the distribution and should not be affected by how we choose to perform the extrapolation.

We check the dependence of our results on the mass distribution of the NS in the binary system. The fiducial scenario assumed a Gaussian distribution for both the components of the binary, as explained in Sec. 3. We draw the value of the two masses from a uniform distribution between $[1; 2.5]M_\odot$ and display in Table 1 the average bias. The result is similar to the case of a Gaussian distribution for the NS masses, with a slight average decrease in the bias on Ω_m of a factor 0.8.

Furthermore, we restrict the value of the inclination angle $\iota \lesssim 20^\circ$, assuming that only beamed GRBs can be detected. This case is reported in the last row of Table 1. The average lensing bias on both the cosmological parameters tends to decrease, in particular of a factor 0.6 for H_0 and 0.5 for Ω_m . This is expected, since on-axis events are louder, and thus less affected by lensing selection effects. We note that the choice $\iota < 20^\circ$ might be conservative. For example, the detection of the radio afterglow may allow to relax the previous constraint (Dobie et al. 2021). We consider the two scenarios, ι drawn from an isotropic distribution and $\iota < 20^\circ$, as two extreme cases.

5.2.1 Log-Normal analytical model

We repeat the analysis using the analytical Log-Normal model, presented in Sec. 3.1. The shape of this magnification PDF, with respect to the one derived from simulations, differs the most around the peak of the distribution, where most of the events fall. In fact, the shape of PDF from simulations is more sharply peaked around $\mu \sim 1$, such that demagnified events have higher values of the magnification. The tail of the two distributions also shows differences, in particular, being underestimated in the Log-Normal model, where the probability of highly magnified events ($\mu \gtrsim 1.5$) is smaller. We compare the results obtained from the same realisation, r_0 , of the mock catalogue but with different magnification values. We use the same distributions for redshifts, masses and angles, while we draw the magnifications from the two different PDF, sampled by the same random numbers. We then calculate the SNR (14) of the events in the two cases and

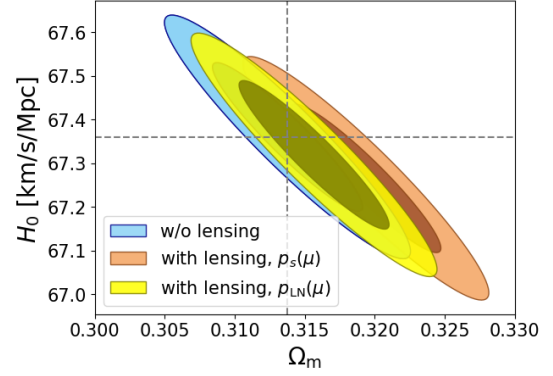


Figure 7. 1σ and 2σ ellipses for H_0 and Ω_m , calculated drawing the lensing magnifications from the Log-Normal model $p_{LN}(\mu)$ (yellow ellipse) and the simulations $p_s(\mu)$ (orange ellipse). The fiducial ellipse, labelled with "w/o lensing", are centered at the correct cosmology, while the biased ones ("with lensing") show the systematic displacement. The intersection of the dotted lines represents the fiducial values of the parameters in standard Cosmology.

keep only the ones that exceed the threshold. In Fig. 7, we show the confidence ellipses obtained with this analysis.

We find that the values of systematic bias of H_0 and Ω_m calculated using the Log-Normal model are less than half the ones obtained with the simulations, in particular with a decrease of a factor ~ 0.4 for H_0 and ~ 0.3 for Ω_m . A similar decrease in the cosmological parameters bias can be found by utilizing the PDF from simulations and removing the events with $\mu \gtrsim 1.4$. This result highlights the importance of an accurate modelling of the magnification PDF, and in particular of its tail, which is underestimated in the Log-Normal model. Note that other works have found Log-Normal descriptions of the lensing PDF that are in better agreement with simulations (Marra et al. 2013).

Even though we consider the PDF obtained from the N-body simulations to be more reliable, recent simulations do not represent the final truth about the magnification PDF and its tail. Resolution can impact the details of the tail. Moreover, baryonic effects, which are not included in the simulations we use, may also have an impact. According to Osato et al. (2021, Figure 5), baryons affect the convergence field distribution at $\sim 10\%$ level, with a smaller impact at higher redshifts.

5.2.2 Number of events

Finally, we discuss an optimistic scenario in which we consider a factor of 10 more events, i.e. we build a mock catalogue with $N_{\text{events}} = 30,000$. The statistical error, at 1σ level, reduces to 0.06% for H_0 and to 0.3% to Ω_m , one order of magnitude tighter than the current EM constraints, for example by Planck Collaboration et al. (2020). This high level of accuracy is achieved keeping the same assumption as in the previous case on the instrumental errors on d_L , i.e. $\sigma_{d_L} = 0.02d_L$. Even though these numbers are probably too optimistic even for 3G detectors, we find this idealised scenario suitable to show the effect of the systematic bias. In fact, for different realisations of the mock catalogue the dispersion of the bias around the mean turns out to be smaller, as it is clear from the two upper panels of Fig. 8, although a smaller number of realisations is considered. Moreover, under these assumptions, the systematic bias is not compatible with zero, with a significance of more than 3σ for H_0 and 6σ for Ω_m . In the lower panel of Fig. 8, the confidence 1σ and 2σ ellipses for one realisation of the mock catalogue are shown. Similarly as before, we choose the first point of the upper panel of Fig. 8 as a representative example and

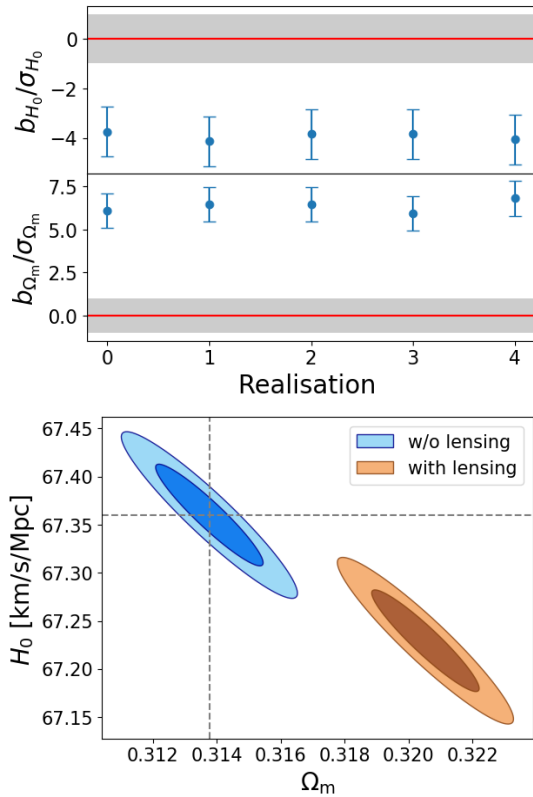


Figure 8. Upper plot: ratio of the bias and statistical error on H_0 and Ω_m for 5 different realisations of the mock catalog, $N_{\text{events}} = 30,000$. The gray bands highlight the region of 1σ significance from zero. Lower plot: ellipses showing the 1σ and 2σ contours obtained considering the realisation r'_0 of the mock catalog.

we called it r'_0 . Also in this scenario, we find that the lensed events are better fitted by smaller (higher) values of H_0 (Ω_m).

5.3 MCMC validation

All the results showcased in previous sections were performed with the Fisher matrix formalism. However, as explained in Section 4, we validate the results with the full likelihood with a MCMC analysis. We compare the outcome of the two methods for a few representative cases. In Fig. 9, we display the comparison for the fiducial scenario, corresponding to the realisation r_0 of the mock catalog. As can be seen from the plot, the MCMC contour lines match the Fisher ellipses with very good agreement. Given that we expect the Fisher approach to fail for high value of the bias, we test the result for a realisation of the mock catalogue showing one of the biggest systematic bias among those observed in our analysis. We consider a realisation of the mock catalog, which we called r_1 , obtained removing the EM selection function while keeping all the other assumptions unchanged with respect to our fiducial scenario. With the Fisher analysis, we obtain $b_{H_0} = -0.37$ Km/s/Mpc with $\sigma_{H_0} = 0.12$ Km/s/Mpc, and $b_{\Omega_m} = 0.016$ with $\sigma_{\Omega_m} = 0.003$. As for the previous case, we find perfect agreement with the MCMC result, obtaining the same significant figures.

Finally, we check our results adding statistical noise in value of the luminosity distances in our mock catalog. In the previous analysis we fitted the biased luminosity distance to its theoretical value, as given in equation (27). However, the measured luminosity distances from

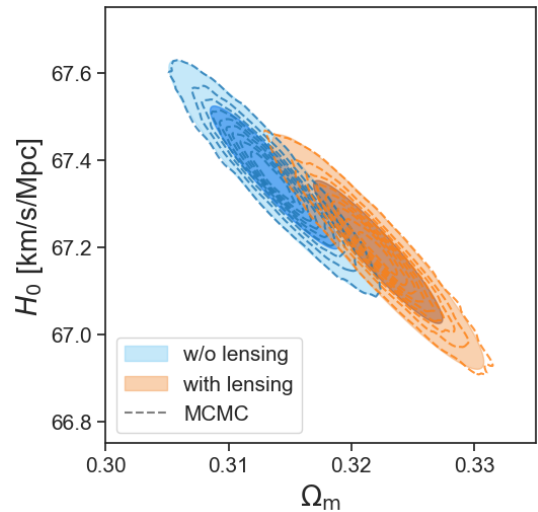


Figure 9. Confidence contours obtained with the Fisher formalism (1σ and 2σ filled ellipses) and the MCMC analysis (dotted lines). The contours are obtained using the same base realisation of the mock catalogue with $N_{\text{events}} = 3,000$.

actual observations are noisy, in addition to biased. To account for this, we sum to d_L^{bias} in equation (27) a statistical error Δd_L , which we draw from a Gaussian distribution with width equal to the expected uncertainty on d_L . We repeat the previous analysis, considering again the same realisation r_0 of the mock catalog, and drawing different realisations of the noise. We notice that even in the fiducial case, i.e. fitting the mock data to the theoretical value of the luminosity distance re-scaled by the magnification factor, the confidence contours are not perfectly centered at the point corresponding to the fiducial values of the cosmological parameters. In particular, we see that, for different realisations of the noise, the position of the contours are randomly scattered around the point corresponding to the fiducial cosmology, even though always remaining consistent at 1σ level. On the other hand, for different realisations of the noise, we find that the value of the bias, computed as the displacement between the mean values of the confidence contours, remains stable and equal to the one obtained without including the statistical noise in the data.

6 MITIGATION

We would like to find a procedure to mitigate the impact of the lensing bias in the estimation of the cosmological parameters. Previous “delensing” attempts in the literature focus on the reduction of statistical errors induced by lensing. The correction is achieved thanks to an independent estimation, which relies on EM observations, of the amount of magnification expected along the line of sight, with an event-by-event procedure (Shapiro et al. 2010; Hilbert et al. 2011; Wu et al. 2023). Even though a similar strategy could be successfully used to address the issue of systematic errors, we consider an alternative, less costly way to statistically mitigate the lensing impact of a catalogue of bright sirens.

For each redshift of our mock catalogue $\{\hat{z}\}$ we draw from the magnification PDF a second independent value of the magnification, which we call $\{\hat{\mu}_M\}$. We use $\{\hat{\mu}_M\}$ to correct the biased luminosity distance, which is obtained from the parameter estimation of the GW. This correction is applied after the selection of events with SNR greater than the SNR threshold. For each event, the mitigated

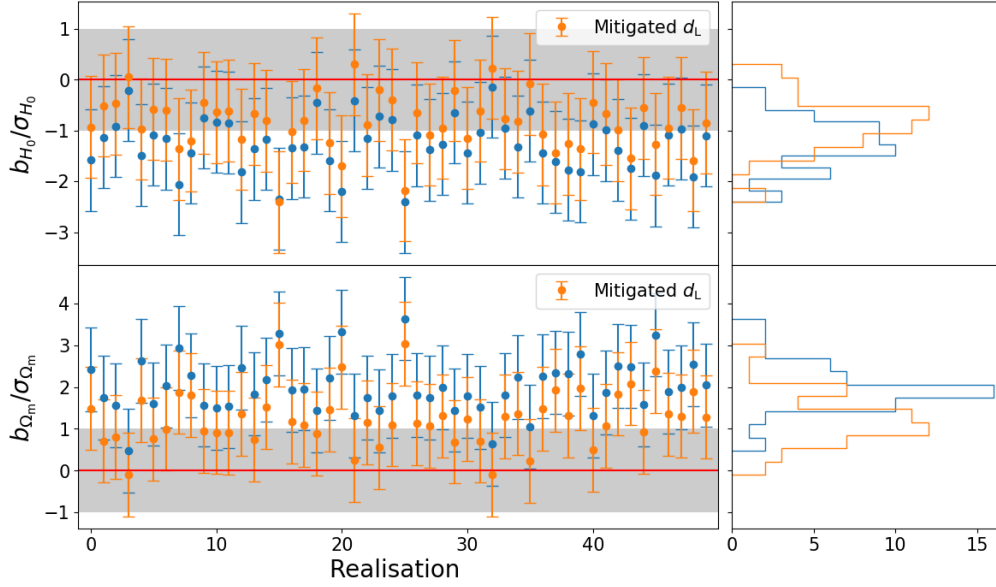


Figure 10. Ratio of the bias and the statistical uncertainty on H_0 (upper panel) and Ω_m (lower panel) for 50 different realisations of the mock catalog, $N_{\text{events}} = 3,000$, with mitigation. The blue points show the absolute value of the bias, without any correction and correspond to the points in Fig. 3; the orange points show the absolute value of the bias obtained after applying the mitigation on the luminosity distance, as in equation (28). The gray bands highlight the region of 1σ significance from zero. In the right panels: histograms of the mitigated ratios in orange and the ones without correction in blue.

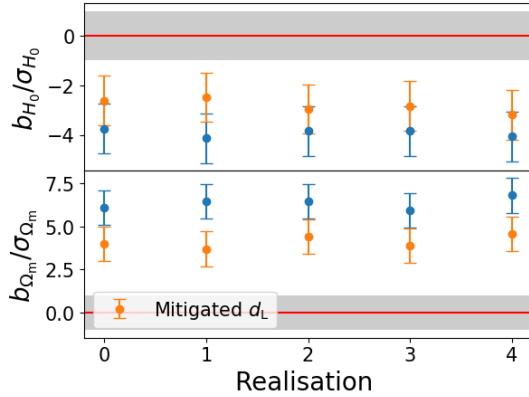


Figure 11. Ratio of the bias and statistical uncertainty on H_0 (upper panel) and Ω_m (lower panel) for 5 different realisations of the mock catalog. In blue the systematic bias and in orange the one obtained from the mitigated d_L , $N_{\text{events}} = 30,000$.

luminosity distance d_L^{MIT} is given by:

$$d_L^{\text{MIT}} = d_L^{\text{bias}}(\mu) \sqrt{\mu_M} = d_L^{\text{true}} \sqrt{\frac{\mu_M}{\mu}}, \quad (28)$$

where d_L^{true} is the true luminosity distance of the event. We calculate the bias with the Fisher formalism, using d_L^{MIT} instead of the biased luminosity distance. To help reducing fluctuations due to events in the tail of the magnification distribution, we repeat the procedure 10 times and take the average of the mitigated values of the bias for each realisation. The result is shown in in Fig. 10 for our fiducial mock catalog. As can be seen, for almost all the realisations we are able to reduce the systematic error, even though the amount of correction is different. In fact, the mitigation procedure reduces the systematic bias on both the cosmological parameters or in the worst case, it keeps it unchanged. The peaks of the histograms of the ratios in the right

panels of Fig. 10 also shifts to lower values with $b_{H_0}/\sigma_{H_0} \sim -0.9$ and $b_{\Omega_m}/\sigma_{\Omega_m} \sim 1.2$. This corresponds to an average reduction of a factor of 0.7 for H_0 and 0.6 for Ω_m .

We also show the mitigation in the case of $N_{\text{events}} = 30,000$, in Fig. 11, obtained from single set of $\{\hat{\mu}_M\}$. Indeed, given the larger sample, we do not average over multiple values of the mitigated bias. In this case, even if we still observe a reduction of the systematic bias for both the cosmological parameters, in any realisation we are unable to find a mitigation that makes the bias compatible with zero or with a significance equal or less than 1σ . Indeed b_{H_0} and b_{Ω_m} remains incompatible with zero at a significance of more than 3σ . This is also due to the fact that in this scenario we have very small error bars on the cosmological parameters, thus more accurate methods are needed for correcting the biased induced by lensing.

6.1 Additional considerations

We also consider alternative approaches that, alone or in combination, can help with the mitigation of the bias (figures are not included).

We try to remove from our catalogue the highly magnified events from the tail of the distribution. We suppose that events that fall into the strong lensing regime can be recognised and thus not used for the cosmological parameters' inference. Data analysis tools have been developed and they allow to carry strong lensing searches in GW data. These searches aim at successfully pairing different images of the same event (Haris et al. 2018; Lo & Magaña Hernandez 2023) or looking for lensing induced waveform distortions of the single image (Wang et al. 2021; Janquart et al. 2021; Ezquiaga et al. 2021). For our fiducial scenario, we find that in order to meaningfully remove the bias, i.e. average bias is below 1σ , it is necessary to remove the high magnification tail for $\mu \gtrsim 1.4$. Thus, we believe that the identification of the strongly lensed signals will not suffice to correct for the systematic bias.

As another attempt, we suppose to know the lensing magnification of the events in the catalogue with a certain precision. For each

event, we draw the second value of the lensing magnification μ'_M from a Gaussian distribution centered around the true magnification μ_{true} . We vary the width of the Gaussian, $\sigma_{\mu_{\text{true}}}$, in order to assess the precision required to reduce the bias. We compare the results for one realisation of the mock catalog. Assuming a 10% precision, i.e. $\sigma_{\mu_{\text{true}}} = 0.1\mu_{\text{true}}$, we find that the bias increases of more than one order of magnitude. Only assuming that μ_{true} is known with 1% precision we recover the same systematic bias, while sub-percent precision is needed to actually reduce it. We found that, the larger the width of the Gaussian, the more the distribution of μ'_M deviates from our fiducial magnification PDF. Once again this underlines the importance of a correct modelling of the magnification PDF.

Finally, we consider the impact of the choice of the SNR threshold. Increasing ρ_{lim} to 12, while keeping all the assumptions of the fiducial scenario unchanged, we notice a worsening of the bias, whose average significance becomes greater than 2σ for H_0 and 3σ for Ω_m . This is expected since, with increasing ρ_{lim} , demagnified, high redshift events are the first to tend to fall below the threshold, leaving a high redshift tail of magnified events. On the other hand, when the threshold is high enough the mock catalogue only consists of small redshift events, for which the magnification PDF is sharply peaked around $\mu = 1$. Clearly, in this limit no lensing bias is expected. For our mock catalog, we find that this condition is met if we impose $\rho_{\text{lim}} = 30$. In this case, the maximum redshift in the mock catalogue reduces to $z \sim 0.6$, and the bias becomes negligible. Although future investigations are needed, we can suppose that the analysis of the high-SNR subgroup can serve as a sanity check for the results obtained by the full catalogue of detection. We also point out that, with increasing ρ_{lim} , much more statistics is needed, since a smaller fraction of events survive the SNR cut. Thus, we considered a smaller number of realisations of the mock catalog. We expect that the details of this analysis might change with a more complete study, but not the overall conclusions.

7 DISCUSSION

We showed that gravitational lensing can induce a systematic bias in the estimation of the cosmological parameters from distances inferred through GWs. We reported the averages over multiple realisations of the mock catalog, since we found that for $N_{\text{events}} = 3,000$, the value of the systematic bias fluctuates over more than one order of magnitude. Even though in some scenarios, e.g. when restricting $\iota \lesssim 20^\circ$, we found that the average bias has a significance smaller than 1σ , it is important to highlight that a few realisations of the mock catalogue still show big fluctuations, up to 2σ level. It is then of crucial importance to be aware of this in the analysis of real data.

As anticipated, we expect this bias to be sourced by two distinct effects: selection effects and the magnified (or demagnified) luminosity distance. We found that the former effects is dominant for a population of NS binaries, detected by a third-generation GW detector. Previous studies on lensing selection effects were presented in Cusin et al. (2019); Cusin & Tamanini (2021); Shan et al. (2021). They define a detector-dependent magnification PDF, that accounts for the finite sensitivity of the detector, and estimate the average magnification expected at a given redshifts. This does not give the probability in absolute terms of observing an event from the given redshift. To do so, we found it more convenient to directly include the magnification in the calculation of the SNR. The SNR cut keeps only the events above the detector's threshold and thus tends to select magnified events whose SNR is increased by a factor $\sqrt{\mu}$. Cusin & Tamanini present an unbiased estimator for the luminosity distance,

which accounts for the average magnification induced by lensing selection effects. We leave for future works to check whether this estimator can mitigate the bias obtained in our analysis. In Shan et al., a forecast for ET is reported with the conclusion that no lensing selection effects are expected. The reason is that almost all GWs from the black holes population that they consider can be detected, even without the magnifying effect of lensing. This is not the case in our study, where we consider the NS population accompanied by an EM counterpart; in fact, with increasing redshift, many more events are found close to the SNR threshold.

Another point of comparison can be made with lensing of Type Ia Supernovae (SNe). Indeed, similarly to GWs, calibrated SNe are used as standard candles to test the distance-redshift relation. In Holz & Linder (2005), the lensing degradation of the power of SNe in the estimation of cosmological parameters is investigated. They highlight that lensing can induce a systematic bias, as a consequence of the non-Gaussian shape of the magnification PDF. If it is not well sampled, the mode of the PDF can show an offset from the mean, taken to be unity. In our analysis, this corresponds to the case in which we do not include the magnification in the calculation of the SNR. Similarly to Holz & Linder (2005), we find that lensing bias has negligible impact on the estimation of the cosmological parameters. Moreover, SNe surveys also suffer of selection effects which can spoil the cosmological parameter estimation (Kessler et al. 2019). For example, the Malmquist bias is due to the magnitude limitation of real surveys, which, near the threshold, tends to select brighter events. Lensing selection effects, analogous to the one considered in this work, are also expected for SNe survey, as highlighted in Shah et al. (2023). They investigate the Pantheon compilation (Scolnic et al. 2018), looking for bias on the cosmological parameters by de-lensing the SNe magnitudes. They find moderate evidence that magnified SNe events were preferentially selected near the magnitude threshold.

We stress that in this work we always assume that the redshift of the GW source is known, thanks to the detection of an EM counterpart. Pairing the GW with the incorrect EM transient will spoil the redshift determination of some of the events. Moreover, different methods have been proposed to estimate the redshift also in the absence of an EM counterpart. ‘‘Dark sirens’’ in fact constitute the great majority of GW events observed so far (Palmese et al. 2020; Abbott et al. 2021a; Finke et al. 2021; Palmese et al. 2023; Abbott et al. 2023). In the case of BNS and BHNS systems, the measurement of tidal effects promises to provide an accurate way to break the mass-redshift degeneracy (Messenger & Read 2012; Ghosh et al. 2022; Shiralilou et al. 2023). Alternatively, the ‘‘galaxy catalogue method’’ statistically infers the host galaxy (Del Pozzo 2012; Fishbach et al. 2019; Gray et al. 2020), and the ‘‘spectral siren method’’ exploits the relation between the source-frame and the detector-frame distribution of the parameters of the GWs population (Taylor et al. 2012; Ezquiaga & Holz 2022). We expect that lensing systematic bias might also become relevant for future dark sirens surveys, which will consist of many order of magnitudes more events than those considered here. Such black hole events can be detected at distances further away than the NS population, increasing the probability of high magnification. We leave the modelling of lensing bias of dark sirens population to future works.

Finally, we note that we focused only on the inference of the cosmological parameters. Future works might improve our analysis by inferring jointly cosmological and BNS population parameters. We hypothesize that the inclusion of the latter ones could decrease the amount of bias obtained in our analysis.

8 CONCLUSIONS

We have investigated the impact of lensing in the estimation of the cosmological parameters, using GWs emitted by merging BNS as standard sirens. Differently from previous works in the literature, we treated lensing also as a systematic error instead as only an additional source of statistical error. In the geometric optics approximation, lensing induces a variation in the amplitude of the GW, which translates in a biased estimation of the luminosity distance of the source. We accounted for selection effects due to the modification of SNR of the events and for the non-Gaussian shape of the magnification PDF.

For our framework, we assumed that 3G ground-base interferometers, in particular ET, together with EM detectors, will be able to collect in a few years of operational time a substantial number of bright sirens to allow for statistical studies. This will open the possibility of testing cosmology using gravitational waves with a precision comparable or even better than the one currently available from EM probes.

Using the magnification PDF obtained from N-body simulations of Hadzhiyska et al. (2023), we randomly assigned a value of μ to each bright siren event in our mock catalog. We allowed also for a tail at high magnifications ($\mu \gtrsim 2$), which could be due, for example, by unidentified strongly lensed events. We included the magnification in the calculation of the SNR and of the biased luminosity distance. We performed our analysis using Fisher forecasts, extended to account for systematic errors. We checked our results with a full likelihood analysis.

Our fiducial scenario consists of a sample of $N_{\text{events}} = 3,000$, with an error on the luminosity distance, measured from the GW waveform, given by the sum of the contributions of instrumental uncertainty, weak lensing, peculiar velocity of the source and redshift measurement, assumed to be made spectroscopically. The instrumental uncertainty on the d_L is set at the percent level, supposing that the EM counterpart can help breaking the degeneracy with the angles in the GW amplitude. We also use a selection function for EM observation that cuts our redshift distribution at a maximum redshift of $z = 2$. For this scenario, we found a bias on the cosmological parameters H_0 and Ω_m that fluctuates as a function of the specific realisation of the mock catalog, as can be seen in Figs. 3 and 4, ranging from significance of departure from zero below 1σ , up to 2.5σ for H_0 , and 3.5σ for Ω_m . Clearly, if we assumed that Ω_m is perfectly known, the statistical error on H_0 would decrease and the impact of the bias would worsen slightly. We also opened the parameter space to the dark energy equation of state parameter w_{DE} in Fig. 5. These results showed that a lensing bias can be expected from future GW data, but its impact on the cosmological parameters strongly depends on the real sample of events that will be collected. When more statistics are considered, i.e. $N_{\text{events}} = 30,000$, we found that the value of the systematic bias is more stable.

The assumptions adopted in our analysis play a crucial role in the determination of whether lensing will really be a worrisome source of systematic errors in the 3G epoch. Important aspects are the precision on the luminosity distance measurements and the number of events that we will be able to collect. We found that the larger the error bars on the cosmological parameters or the smaller the number of events, the less significant the systematic bias, as can be seen from Fig. 6, Appendix A and from the analysis with $N_{\text{events}} = 30,000$. While we compared the results obtained with different assumptions, future works might implement a complete forecast for the GW parameters estimation analysis and modelling of the EM emission.

Another point that may impact the analysis is the modelling of the lensing PDF, in particular its high magnification tail. To show this

we employed a simple Log-Normal model for the PDF, whose shape has differences both around the peak and the tail of the distribution, and we found that a smaller bias is introduced. We also checked that a similar effect is obtained by cutting the high μ events drawn from our base PDF. Other aspects that can affect the results are the inclination angle and the EM selection function. The former can be restricted to be beamed around zero in order to detect the GRB event. Such events are louder in SNR, thus less affected by lensing selection effects, and in turn, by smaller systematic bias. The latter introduces a selection of the farther events from which the redshift of the source can be determined. Farther events are more likely to undergo high magnifications and spoil the cosmological parameters estimation.

We also presented a simple procedure to try to alleviate the lensing bias. This procedure, even if it is not the final solution, can be regarded as a good and cheap starting point. As can be seen from Fig. 10, the level of correction that can be achieved depends on the specific realisation of the mock catalogue, and in the worst case, leave the results unchanged. A safer but less desired solution is to blow up the statistical error bars to make the systematic displacement irrelevant.

Overall, our work highlights the necessity of properly including lensing in the estimation of cosmological parameters from GW, not only as an additional source of noise but also as a possible systematic. On the other hand, we find that the level of accuracy needed in the estimation of the cosmological parameters to appreciate this effect is very high, i.e. of the order of sub-percent level for H_0 . Given that the current tension in the values of H_0 obtained with different probes is of the order of several per-cent, we can safely conclude that BNSs will still be able to give an independent contribution towards overcoming the Hubble tension. On the other hand, a future era of cosmological measurements may open the necessity of reaching such high precision, for example, to probe the nature of dark energy. Although future works are needed in order to better address the impact of lensing in alternative theories of gravity, it is already possible to hypothesize that degeneracy between lensing bias and dark energy effects can arise, making the necessity of correctly modelling lensing very compelling.

ACKNOWLEDGEMENTS

We thank Boet Hoitink for helpful discussions during his thesis project. We are grateful to Tessa Baker, Tanja Hinderer, Macarena Lagos for helpful feedback on the manuscript. Finally, we thank Jeger Broxterman, Boryana Hadzhiyska, Jia Liu and Ken Osato for help with the weak lensing simulations. This research was funded in whole, or in part, by the Dutch Research Council (NWO) 24.001.027. For the purpose of open access, a CC BY public copyright license is applied to any Author Accepted Manuscript version arising from this submission.

DATA AVAILABILITY

The files containing the mock catalogue of events used in this paper will be made publicly available at the following link: [Github/GWlensing](#).

REFERENCES

- Abbott B. P., et al., 2016, *Phys. Rev. Lett.*, 116, 061102
 Abbott B. P., et al., 2017a, *Nature*, 551, 85
 Abbott B. P., et al., 2017b, *The Astrophysical Journal Letters*, 848, L13

- Abbott B. P., et al., 2021a, *The Astrophysical Journal*, 909, 218
- Abbott R., et al., 2021b, *The Astrophysical Journal*, 923, 14
- Abbott R., et al., 2023, *The Astrophysical Journal*, 949, 76
- Amara A., Réfrégier A., 2008, *Monthly Notices of the Royal Astronomical Society*, 391, 228
- Amaro-Seoane P., et al., 2017, *arXiv e-prints*, p. arXiv:1702.00786
- Amati L., et al., 2018, *Advances in Space Research*, 62, 191
- Baker T., Trodden M., 2017, *Phys. Rev. D*, 95, 063512
- Balardo A., Garofolo A., Martinelli M., Mukherjee S., Silvestri A., 2023, *Journal of Cosmology and Astroparticle Physics*, 2023, 050
- Belgacem E., Dirian Y., Foffa S., Howell E. J., Maggiore M., Regimbau T., 2019, *Journal of Cosmology and Astroparticle Physics*, 2019, 015
- Bertacca D., Raccanelli A., Bartolo N., Matarrese S., 2018, *Physics of the Dark Universe*, 20, 32
- Borhanian S., Sathyaprakash B. S., 2022, *arXiv e-prints*, p. arXiv:2202.11048
- Boyle A., Uhlemann C., Friedrich O., Barthelemy A., Codis S., Bernardeau F., Giocoli C., Baldi M., 2021, *MNRAS*, 505, 2886
- Branchesi M., et al., 2023, *J. Cosmology Astropart. Phys.*, 2023, 068
- Cai R.-G., Yang T., 2017, *Phys. Rev. D*, 95, 044024
- Calderón Bustillo J., Leong S. H. W., Dietrich T., Lasky P. D., 2021, *Astrophys. J. Lett.*, 912, L10
- Camp J., 2017, in *AAS/High Energy Astrophysics Division #16*. p. 103.22
- Chisari N. E., et al., 2019, *The Astrophysical Journal Supplement Series*, 242, 2
- Congedo G., Taylor A., 2019, *Phys. Rev. D*, 99, 083526
- Cusin G., Tamanini N., 2021, *Monthly Notices of the Royal Astronomical Society*, 504, 3610
- Cusin G., Durrer R., Dvorkin I., 2019, *arXiv e-prints*, p. arXiv:1912.11916
- Cutler C., Flanagan E. E., 1994, *Phys. Rev. D*, 49, 2658
- Dai L., Venumadhav T., Sigurdson K., 2017, *Phys. Rev. D*, 95, 044011
- Dalal N., Holz D. E., Hughes S. A., Jain B., 2006, *Phys. Rev. D*, 74, 063006
- Das S., Ostriker J. P., 2006, *The Astrophysical Journal*, 645, 1
- Del Pozzo W., 2012, *Phys. Rev. D*, 86, 043011
- Di Valentino E., et al., 2021, *Classical and Quantum Gravity*, 38, 153001
- Diego J. M., 2019, *A&A*, 625, A84
- Diego J. M., Broadhurst T., Smoot G. F., 2021, *Phys. Rev. D*, 104, 103529
- Dobie D., Murphy T., Kaplan D. L., Hotokezaka K., Bonilla Ataides J. P., Mahony E. K., Sadler E. M., 2021, *Monthly Notices of the Royal Astronomical Society*, 505, 2647
- Euclid Collaboration et al., 2020, *A&A*, 642, A191
- Evans M., et al., 2021, *arXiv e-prints*, p. arXiv:2109.09882
- Ezquiaga J. M., Holz D. E., 2022, *Phys. Rev. Lett.*, 129, 061102
- Ezquiaga J. M., Zumalacárregui M., 2020, *Phys. Rev. D*, 102, 124048
- Ezquiaga J. M., Holz D. E., Hu W., Lagos M., Wald R. M., 2021, *Phys. Rev. D*, 103, 064047
- Finke A., Foffa S., Iacovelli F., Maggiore M., Mancarella M., 2021, *Journal of Cosmology and Astroparticle Physics*, 2021, 026
- Finn L. S., Chernoff D. F., 1993, *Phys. Rev. D*, 47, 2198
- Fishbach M., et al., 2019, *The Astrophysical Journal Letters*, 871, L13
- Fonseca J., Zazzera S., Baker T., Clarkson C., 2023, *Journal of Cosmology and Astroparticle Physics*, 2023, 050
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
- Fosalba P., Gaztañaga E., Castander F. J., Crocce M., 2014, *Monthly Notices of the Royal Astronomical Society*, 447, 1319
- Fragione G., 2021, *The Astrophysical Journal Letters*, 923, L2
- Ghosh T., Biswas B., Bose S., 2022, *Phys. Rev. D*, 106, 123529
- Gouin C., et al., 2019, *A&A*, 626, A72
- Goyal S., Haris K., Mehta A. K., Ajith P., 2021, *Phys. Rev. D*, 103, 024038
- Gray R., et al., 2020, *Phys. Rev. D*, 101, 122001
- Gray R., et al., 2023, *arXiv e-prints*, p. arXiv:2308.02281
- Hadzhiyska B., et al., 2023, *arXiv e-prints*, p. arXiv:2305.11935
- Haris K., Mehta A. K., Kumar S., Venumadhav T., Ajith P., 2018, *arXiv e-prints*, p. arXiv:1807.07062
- Hilbert S., Gair J. R., King L. J., 2011, *Monthly Notices of the Royal Astronomical Society*, 412, 1023
- Hirata C. M., Holz D. E., Cutler C., 2010, *Phys. Rev. D*, 81, 124046
- Hogg D. W., Foreman-Mackey D., 2018, *The Astrophysical Journal Supplement Series*, 236, 11
- Holz D. E., Hughes S. A., 2005, *The Astrophysical Journal*, 629, 15
- Holz D. E., Linder E. V., 2005, *ApJ*, 631, 678
- Ivezic Ž., et al., 2019, *ApJ*, 873, 111
- Janquart J., Seo E., Hannuksela O. A., Li T. G. F., Broeck C. V. D., 2021, *The Astrophysical Journal Letters*, 923, L1
- Jin S.-J., He D.-Z., Xu Y., Zhang J.-F., Zhang X., 2020, *Journal of Cosmology and Astroparticle Physics*, 2020, 051
- Kayo I., Taruya A., Suto Y., 2001, *The Astrophysical Journal*, 561, 22
- Kessler R., et al., 2019, *Monthly Notices of the Royal Astronomical Society*, 485, 1171
- Kim K., Lee J., Hannuksela O. A., Li T. G. F., 2022, *The Astrophysical Journal*, 938, 157
- Kocsis B., Frei Z., Haiman Z., Menou K., 2006, *The Astrophysical Journal*, 637, 27
- Laguna P., Larson S. L., Spergel D., Yunes N., 2010, *The Astrophysical Journal Letters*, 715, L12
- Laureijs R., et al., 2011, *arXiv e-prints*, p. arXiv:1110.3193
- Li S.-S., Mao S., Zhao Y., Lu Y., 2018, *Monthly Notices of the Royal Astronomical Society*, 476, 2220
- Lo R. K. L., Magaña Hernandez I., 2023, *Phys. Rev. D*, 107, 123015
- Madau P., Fragos T., 2017, *The Astrophysical Journal*, 840, 39
- Maggiore M., et al., 2020, *Journal of Cosmology and Astroparticle Physics*, 2020, 050
- Marra V., Quartin M., Amendola L., 2013, *Phys. Rev. D*, 88, 063004
- Martinelli M., Scarcella F., Hogg N. B., Kavanagh B. J., Gaggero D., Fleury P., 2022, *Journal of Cosmology and Astroparticle Physics*, 2022, 006
- Messenger C., Read J., 2012, *Phys. Rev. Lett.*, 108, 091101
- Mishra A., Meena A. K., More A., Bose S., Bagla J. S., 2021, *MNRAS*, 508, 4869
- Mpetha C. T., Congedo G., Taylor A., 2023, *Phys. Rev. D*, 107, 103518
- Mukherjee S., Wandelt B. D., Silk J., 2020, *Monthly Notices of the Royal Astronomical Society*, 494, 1956
- Muttoni N., Laghi D., Tamanini N., Marsat S., Izquierdo-Villalba D., 2023, *arXiv e-prints*, p. arXiv:2303.10693
- Nissanke S., Holz D. E., Hughes S. A., Dalal N., Sievers J. L., 2010, *The Astrophysical Journal*, 725, 496
- Oguri M., 2018, *Monthly Notices of the Royal Astronomical Society*, 480, 3842
- Osato K., Liu J., Haiman Z., 2021, *Monthly Notices of the Royal Astronomical Society*, 502, 5593
- Palmese A., et al., 2020, *The Astrophysical Journal Letters*, 900, L33
- Palmese A., Bom C. R., Mucesh S., Hartley W. G., 2023, *The Astrophysical Journal*, 943, 56
- Perivolaropoulos L., Skara F., 2022, *New Astron. Rev.*, 95, 101659
- Pieroni M., Ricciardone A., Barausse E., 2022, *Scientific Reports*, 12, 17940
- Planck Collaboration et al., 2020, *A&A*, 641, A6
- Punturo M., et al., 2010, *Classical and Quantum Gravity*, 27, 194002
- Regimbau T., et al., 2012, *Phys. Rev. D*, 86, 122001
- Regimbau T., Siellez K., Meacher D., Gendre B., Boër M., 2015, *ApJ*, 799, 69
- Reitze D., et al., 2019, in *Bulletin of the American Astronomical Society*. p. 35 (arXiv:1907.04833), doi:10.48550/arXiv.1907.04833
- Ronchini S., et al., 2022, *A&A*, 665, A97
- Sathyaprakash B. S., Schutz B. F., Broeck C. V. D., 2010, *Classical and Quantum Gravity*, 27, 215006
- Schutz B. F., 1986, *Nature*, 323, 310
- Scolnic D. M., et al., 2018, *The Astrophysical Journal*, 859, 101
- Sereni M., Jetzer P., Sesana A., Volonteri M., 2011, *Monthly Notices of the Royal Astronomical Society*, 415, 2773
- Shah P., Lemos P., Lahav O., 2023, *MNRAS*, 520, L68
- Shan X., Wei C., Hu B., 2021, *Monthly Notices of the Royal Astronomical Society*, 508, 1253
- Shang C., Haiman Z., 2011, *Monthly Notices of the Royal Astronomical Society*, 411, 9
- Shapiro C., Bacon D. J., Hendry M., Hoyle B., 2010, *Monthly Notices of the Royal Astronomical Society*, 404, 858

- Shiralilou B., Raaijmakers G., Duboef B., Nissanke S., Foucart F., Hinderer T., Williamson A., 2023, *The Astrophysical Journal*, 955, 149
- Stratta G., et al., 2018, *Advances in Space Research*, 62, 662
- Takahashi R., Nakamura T., 2003, *The Astrophysical Journal*, 595, 1039
- Takahashi R., Oguri M., Sato M., Hamana T., 2011, *The Astrophysical Journal*, 742, 15
- Takahashi R., Hamana T., Shirasaki M., Namikawa T., Nishimichi T., Osato K., Shiroyama K., 2017, *ApJ*, 850, 24
- Taylor S. R., Gair J. R., Mandel I., 2012, *Phys. Rev. D*, 85, 023535
- The LIGO Scientific Collaboration et al., 2021, *arXiv e-prints*, p. arXiv:2111.03606
- Wang Y., Lo R. K. L., Li A. K. Y., Chen Y., 2021, *Phys. Rev. D*, 103, 104055
- Wu Z.-F., Chan L. W. L., Hendry M., Hannuksela O. A., 2023, *Monthly Notices of the Royal Astronomical Society*, 522, 4059
- Xavier H. S., Abdalla F. B., Joachimi B., 2016, *Monthly Notices of the Royal Astronomical Society*, 459, 3693
- Yu J., et al., 2021, *The Astrophysical Journal*, 916, 54
- Zhao W., Van Den Broeck C., Baskaran D., Li T. G. F., 2011, *Phys. Rev. D*, 83, 023005

APPENDIX A: ADDITIONAL VARIATION FROM THE FIDUCIAL SCENARIO

In Subsec. 5.2, we discuss possible variations of our fiducial scenario that can affect the amount of bias on the cosmological parameters. In this Appendix, we focus more on the influence of the luminosity distance instrumental uncertainty σ_{GW} , which enters in equation (16). In our fiducial scenario we consider a value of $\sigma_{\text{GW}} = 0.02d_L$, assuming that bright standard sirens will be a well-measured small fraction of the total number of events collected by ET. This is in view of the fact that the EM counterpart can help breaking the degeneracy of the luminosity distance with the angles in the amplitude, especially ι . Assuming a less optimistic scenario of $\sigma_{\text{GW}} = 0.1d_L$, we report for 10 realisations of our mock catalogue the ratio of bias and statistical uncertainty on H_0 and Ω_m . Similarly to the case reported in Subsec. 5.2, with the choice $\sigma_{\text{GW}} = 0.1d_L$ we find that the statistical error increases to 0.5% and 3.2% respectively for H_0 and Ω_m , while the average ratios decrease to $b_{H_0}/\sigma_{H_0} = -0.4$ and $b_{\Omega_m}/\sigma_{\Omega_m} = 0.7$, as shown in the upper plot of Fig. A1. Thus, in this scenario we obtain that the significance of the bias is below 1σ for both H_0 and Ω_m , making its contribution on the inferred value of the cosmological parameters negligible. Indeed, the relative size between the statistical uncertainty and the bias is of crucial importance to determine whether the systematic displacement of the confident contours will be or not negligible for the next generation of GW detectors. Moreover, we stress that another important factor is the limiting SNR threshold of the detector. In our fiducial scenario, we assume $\rho_{\text{lim}} = 8$ obtained as the sum over the three detectors that compose the triangular configurations of ET. As highlighted in Subsec. 6.1, increasing ρ_{lim} tends to worsen the bias, since the lensing-induced selection effect becomes more severe. In the lower plot of Fig. A1, we show the results obtained assuming $\rho_{\text{lim}} = 12$ and $\sigma_{\text{GW}} = 0.1d_L$. In this case, we find that the average ratios increase to $b_{H_0}/\sigma_{H_0} = -1.1$ and $b_{\Omega_m}/\sigma_{\Omega_m} = 1.7$, thus comparable to the results obtained in our fiducial scenario, e.g. Fig. 3.

This paper has been typeset from a \LaTeX file prepared by the author.

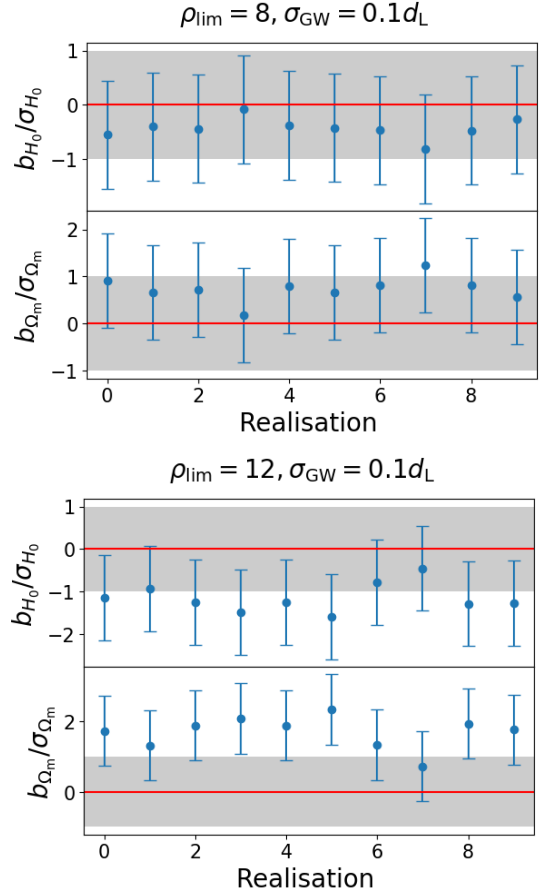


Figure A1. Ratio of the bias and statistical uncertainty on H_0 and Ω_m for 10 different realisations of the mock catalog. We vary our fiducial scenario by assuming $\sigma_{\text{GW}} = 0.1d_L$ with $\rho_{\text{lim}} = 8$ (upper plot) and $\rho_{\text{lim}} = 12$ (lower plot).