Journal of Cosmology and Astroparticle Physics
An IOP and SISSA journal

Will cosmic gravitational wave sirens determine the Hubble constant?

Arman Shafieloo,[a,b] **Ryan E. Keeley**[a] **and Eric V. Linder**[a,c,d]

[a]Korea Astronomy and Space Science Institute,
Daejeon 34055, Korea
[b]University of Science and Technology,
Yuseong-gu 217 Gajeong-ro, Daejeon 34113, Korea
[c]Berkeley Center for Cosmological Physics & Berkeley Lab, University of California,
Berkeley, CA 94720, U.S.A.
[d]Energetic Cosmos Laboratory, Nazarbayev University,
Astana, 010000 Kazakhstan

E-mail: shafieloo@kasi.re.kr, rkeeley@kasi.re.kr, evlinder@lbl.gov

Received November 25, 2019
Revised January 21, 2020
Accepted February 11, 2020
Published March 6, 2020

Abstract. Incorrect assumptions about the background expansion history of the Universe can induce significant biases when estimating the Hubble constant H_0 and other key cosmological parameters from cosmic ($z \gtrsim 0.1$) gravitational wave standard sirens, even with electromagnetic counterpart redshifts. Future gravitational wave experiments such as the Einstein Telescope can provide us with a compilation of gravitational wave sirens that can be used to determine these cosmological parameters with very high precision. In such a future, the statistical precision can reach to the level of 1% uncertainty on H_0. However, such datasets would include a large number of cosmic gravitational wave sirens, and not only sources at very low redshifts of $z \lesssim 0.1$. We show that wrong assumptions about the background expansion history of the Universe (e.g. form of dark energy) can introduce substantial bias in estimation of the Hubble constant and the other key parameters. Such biases would occur in non-ΛCDM cosmologies that can be degenerate with the standard ΛCDM model. To avoid model-dependent biases, statistical techniques that are appropriately agnostic about model assumptions need to be employed.

Keywords: dark energy experiments, dark energy theory

ArXiv ePrint: 1812.07775

 https://doi.org/10.1088/1475-7516/2020/03/019

JCAP03(2020)019

Contents

1 Introduction

ΛCDM, so named for its cosmological constant (Λ) and cold dark matter (CDM) components, has been the standard model of cosmology for the past generation. It has successfully explained a variety of cosmological observables from the cosmic microwave background at high redshift to supernova (SN) distance modulii at low redshift. Typically, modern cosmology seeks to constrain the parameters of the ΛCDM model or a small space of extensions to this model (e.g. $w_0 w_a$). In this paper, we aim to show that such assumptions can lead to biases in parameter estimation (when parametric assumptions are not appropriate) and that novel statistical techniques are required to avoid such biases. We imagine future gravitational wave (GW) datasets to illustrate the point that, for models presently allowed by cosmological datasets, biases can arise in future high precision distance datasets.

Gravitational waves emitted from inspiral and coalescence of binary compact objects can be used to measure a dimensional quantity — the time or frequency associated with the wave form. By modeling the expected wave form within general relativity these events can be standard sirens, measuring dimensional cosmic distances. Since most cosmic measurements involve dimensionless quantities (often ratios of distances), this makes standard siren distances potentially useful in a distinct way. In particular, they have been proposed to measure the absolute distance scale of the universe, or Hubble constant H_0 [1–4]. This is an exciting prospect. In the era of precision cosmology, it is becoming increasingly important to be appropriately agnostic about model uncertainties, about whether ΛCDM truly models the expansion of the Universe. Specifically, when using cosmic distance measurements, one must allow sufficient flexibility in the nature of dark energy to avoid biasing the inference of cosmological parameters. We demonstrate this point by analyzing an idealized and optimistic mock gravitational wave dataset.

Locally, at very low redshifts $z \lesssim 0.1$, the source distance is related linearly to the distance through the Hubble law, $d = H_0^{-1} z$. This means that the redshift to the source must also be determined, but it is not uniquely provided by the GW observations. The most straightforward way to obtain the redshift is to use GW systems with electromagnetic (EM) counterpart events (e.g. X-ray or optical flashes associated with the merger), where the redshift comes from the EM measurement. Crosscorrelation with redshift surveys is an alternative area of investigation, e.g. [5, 6]. This article differs from those already in the literature by focusing on potential bias of cosmological parameters, rather than precision of one particular parameter such as H_0.

However, very low (local) redshift means a small volume in which events can occur, hence small numbers of observed GW+EM systems (e.g. binary neutron star (BNS) mergers) and relatively poor precision on H_0. Some papers have held them out as the means to measure the Hubble constant to 1% or better precision (see, e.g. [7]). However, astrophysical systematics such as binary orbit inclination effects, peculiar velocities, and signal to noise (Malmquist-like) biases can also affect the use of BNS GW sirens to constrain H_0 (see, e.g., [6–8]). For local sirens this could include coherent velocity flows (cf. [9, 10]; see [11] for a general review of GW detectors). We are primarily interested in cosmological systematics, rather than astrophysical ones, therefore, we do not consider these local sirens further here.

Dark sirens from the mergers of binary black holes would be detectable at much higher distances. At large, cosmological distances, dark sirens would be in the Hubble flow, mitigating astrophysical systematics. However, to build a Hubble diagram from these sources requires crosscorrelating with redshift surveys [2, 5]. We leave the investigation of the effect of potential redshift uncertainties or biases that may arise to future work.

Upcoming gravitational wave detectors — more sensitive runs of LIGO-Hanford and LIGO-Livingston [12], Virgo [13], and new interferometers in India and Japan [14, 15] — will be able to detect GW events from cosmic sirens out to higher redshift, $z \approx 0.5$, with further generations including space based detectors such as LISA [16–18] and the Einstein Telescope [19–24] reaching $z \approx 1$ or beyond, even for BNS sources that have redshifts measured from EM counterparts. These proposed telescopes will detect significantly more events.

We look at a cosmic sample of GW sirens with an optimistic redshift distribution (i.e. volume limited) so as to understand the effects of the accuracy on the best precision H_0. We should note that the distribution of the cosmic sirens versus redshift that we considered in our analysis is similar to the expectations from the Einstein Telescope or a similar experiment [19]. We argue that to use these cosmic sirens we have to consider the bias that can arise due to background cosmology assumptions. Regardless of any specific configuration of networks, so long as they can detect sources up to high redshifts (with EM counterparts such as the case of BNS sources) our argument holds.

Obtaining a 1% measurement of H_0 from moderately low redshift observables can be important for cosmology because of the discordance between the inference of H_0 from observations of the cosmic microwave background (CMB) [25] and the value observed locally from calibrating the SN distance ladder with Cepheids [26, 27]. An independent measurement of H_0 would offer keen insight into this tension (e.g. [28]). If H_0 from GW preferred the Planck value, then that might indicate an unaccounted for systematic with Cepheid calibrations of SN distances. If it agrees with the Cepheid+SN measurement, then that might lend more credence to the idea this discordance is explained by new physics.

At redshifts $z \gtrsim 0.1$, though, the distance does not depend solely on H_0 but on an integral over the expansion history $H(z)$, with all components of energy density in the universe — in particular matter and dark energy — contributing. Without knowledge of these components one cannot cleanly separate out the Hubble constant.

Our focus here is two-fold: investigation of the precision with which H_0 and the matter density Ω_m can be constrained in a background more general than ΛCDM — and indeed whose form may be unknown, and investigation of the accuracy, i.e. the bias suffered when the background is not accounted for correctly. [7] have put forth the exciting prospect of 1–2% measurement of H_0 from GW but within a calculation assuming not only ΛCDM but a perfectly known matter density. While [29] have looked beyond a ΛCDM background, this was only for non-GW probes, implementing the GW data as purely a H_0 prior from [7].

[30] have included a full dynamical dark energy background, but for far future GW data sets of 1000 sirens out to $z = 5$ and in combination with other probes already giving strong cosmology constraints. Our approach is to examine the role of the background expansion model on both the precision and accuracy of the H_0 and Ω_m determination from mid and moderately long term GW experiments.

In summary, we focus here on cosmic ($z \gtrsim 0.1$) sources, keeping all cosmological parameters free, and examine bias due to more restrictive assumptions.

In section 2 we present the framework for the analysis, including the GW+EM datasets corresponding to future, and far future generation GW experiments and our simulation methodology. Precision on H_0 is treated in section 3, where we examine how it degrades with greater freedom for the expansion history: first including just the matter density and a cosmological constant, then allowing for dynamical dark energy with assumption of the standard w_0-w_a time dependence. Accuracy is the focus in section 4, where we show how assuming a ΛCDM cosmology can significantly bias the results in the regime of 1% precision. In section 5 we conclude and discuss the statistical techniques needed to infer $H(z)$ without bias from high precision datasets.

2 Model assumptions and data simulations

As stated in the introduction, GW standard sirens do not directly measure H_0 but rather measure luminosity distances D_L throughout the cosmic volume to which the detectors are sensitive. Each event has a cosmic redshift associated with it, which must be obtained from EM counterparts. The distance is then related to the cosmic expansion rate through

$$D_L(z) = (1+z)\frac{c}{H_0}\int_0^z \frac{dz'}{h(z')}\,, \tag{2.1}$$

where $h(z)$ is the Hubble rate scaled to the present value, $H(z)/H_0$, and a spatially flat universe is assumed. The assumptions about the background expansion model $h(z)$ have a direct impact on extraction of H_0. Even under the assumption of flat ΛCDM the matter density must also be known: $h^2(z) = \Omega_\mathrm{m}(1+z)^3 + 1 - \Omega_\mathrm{m}$.

To be concrete about how uncertainties in the expansion history can affect model dependent inferences using GWs, we generate mock luminosity distance datasets from a given background cosmology. We consider two alternative possibilities: 1) a flat ΛCDM model with $\Omega_\mathrm{m} = 0.3$ and $h \equiv h(z=0) = 0.69$, and 2) time varying dark energy with assumption of $w(z) = w_0 + w_a z/(1+z)$ which provides more flexibility to the expansion history. Additional classes of dark energy models could offer similar, interesting results [31].

For the latter we choose two models consistent with current data, specifically lying on the 68% confidence contour of the Pantheon supernovae plus Planck CMB plus SDSS BAO plus HST H_0 combined data fit of [32].

To generate a realistic mock dataset of GW luminosity distances, we sample the event redshift distribution based on the assumption that the GW events have a constant rate per comoving volume. That is,

$$N(z) = \int dz \frac{dN}{dz} = \int dz \frac{dN}{dV_c}\frac{dV_c}{dz}, \tag{2.2}$$

where we assume dN/dV_c is constant and calculate dV_c/dz from our fiducial input cosmology.

We perform this sampling for two cases. One is a case where we use a maximum redshift $z = 0.5$ and draw 120 events from this distribution (which represent a sample of events with somehow accurately determined redshifts), with distance errors normally distributed with a 1σ precision of 13%, roughly following [7]. This number comes from their result that the precision on H_0 scales as $13\%/\sqrt{N}$ where N is the number of GW detections. This result would be consistent with the case if the average uncertainty for each GW distance was 13%, though there are potential non-Gaussianities in the measurement of these distances which could further confuse results. Such a case may correspond roughly to the ∼2026 HLVJI array of Hanford and Livingston detectors of LIGO, and Virgo, KAGRA (Japan), and LIGO-India detectors or it may correspond to a detector configuration further in the future. Whichever configuretion does achieve such a dataset, that is when our results will hold.

The other is a far future case (possibly corresponding to 3rd generation detectors such as the Einstein Telescope [19, 22] or LISA [16–18] though we prefer to remain agnostic about the specifics of potential future detectors) where we take the maximum redshift to be $z = 1.0$ and draw 600 events from the distribution, with 7% distance precision. This 7% precision was chosen to be illustrative of a case where 1% precision on H_0 is achieved, even when marginalizing over Ω_m. One realization for each of these cases can be seen in figure 1.

Again we emphasize that any projection of the distance information from these redshifts to $H_0 = H(z = 0)$ requires the assumption of an uncertain background model. Testing whether or not this model is true is of course one aim of any analysis of cosmological datasets, along with H_0.

Our approach is similar to the approach from [21] where the authors use GW sirens with known redshifts out to z=5. They are similarly agnostic about how that redshift information is obtained, whether it is from an EM counterpart or from cross-correlation with galaxy surveys. We proceed in a similar manner and construct a mock GW dataset out to cosmological redshifts. The differences between our papers is what questions we are trying to answer with the future cosmic GW datasets. [21] discuss how precisely we can estimate the key cosmological parameters assuming we know the true model of the Universe (e.g. ΛCDM). In our analysis, we discuss the case of ΛCDM being assumed when some other model of the Universe's expansion history is the true model. We show how this wrong assumption about the true background model of the Universe can result in biasing the estimation of key cosmological parameters including the Hubble constant and matter density. While these biases may not be a concern with present cosmological datasets, we show they will become important for future high precision datasets, such as those from 3rd generation GW detectors.

3 Precision vs cosmological model

Combining multiple cosmological probes within a given model can give tight constraints on parameters, including H_0 and Ω_m. As a rough rule of thumb, note that the CMB already tightly constrains the combination $\Omega_\mathrm{m}h^3$, to about 0.3% [25], fairly independently of late time physics (though still power-law form of the primordial power spectrum is assumed). Thus $\delta h/h \approx (1/3)\delta\Omega_\mathrm{m}/\Omega_\mathrm{m}$ so a prior of 0.03 on Ω_m from large scale structure probes gives a 3% constraint on h. Adding other probes such as supernovae would tighten this further. Thus we want another, individual probe at the level of $\sim 1\%$ on h. Let us explore under what conditions GW sirens can provide this in themselves.

Table 1 shows the constraints on H_0 from future GW data, either alone or with external priors. For this table alone, the numbers come from Fisher information computation; all

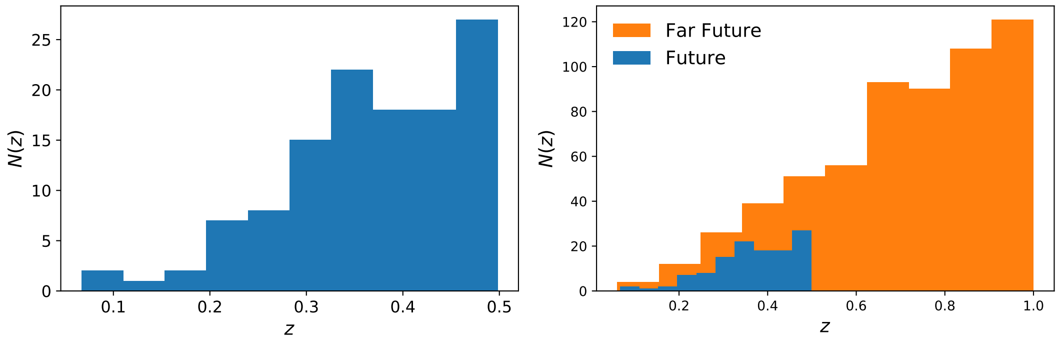

Figure 1. One realization of the redshift distribution of GW+EM events for the future case of 120 total events observed out to maximum redshift $z = 0.5$ (left panel) and for the far future case of 600 total events observed out to maximum redshift of $z = 1.0$ (right panel). The future events are overplotted on the lower panel for comparison.

Background	Prior	$\sigma(h)$	$\sigma(h)/h$
ΛCDM	none	0.036	5%
ΛCDM	$\sigma(\Omega_{\rm m}) = 0.03$	0.010	1.4%
ΛCDM	fix $\Omega_{\rm m}$	0.0083	1.2%
w_0–w_a*	none	0.039*	6%*
w_0–w_a*	$\sigma(\Omega_{\rm m}) = 0.03$	0.039*	6%*
w_0–w_a*	fix $\Omega_{\rm m}$	0.039*	6%*

Table 1. Constraints on H_0 from the future GW+EM set are given under various backgrounds and priors. An asterisk denotes a broad prior of 1 on both w_0 and w_a, since the Fisher information approach is inaccurate for extended degeneracies. We see that GW siren constraints are sensitive to the input model.

other numbers in the article are from MCMC. They are in good agreement where they overlap. When the background is fixed to ΛCDM, and furthermore the matter density is perfectly known, then the uncertainty is $\sigma(h) = 0.0083$ or approximately 1.2%. When external information is used at the level of a prior on matter density of $\sigma(\Omega_{\rm m}) = 0.03$ then the H_0 precision has a modest increase to 1.4%. However, such combination of probes can be done as well between non-GW probes (e.g. supernovae, strong lenses, large scale structure, CMB), so we should look at what GW sirens themselves deliver. For GW alone, even with fixing the background to ΛCDM, the uncertainty is $\sigma(h) = 0.036$ or 5%. This will not allow them to make a statistically significant statement on the tension between the Planck value (which used the same assumption of ΛCDM) and the local distance ladder value.

To illustrate the effect of the matter density covariance with H_0 within the ΛCDM model, we perform a Monte Carlo simulation of future GW data. That is, we generate realizations of the data in a particular ΛCDM case with $\Omega_{\rm m} = 0.3$ and $h = 0.69$, and fit for these two parameters under the assumption we know the background is ΛCDM. The 1D and 2D joint confidence contours appear in figure 2.

The covariance between $\Omega_{\rm m}$ and h is clear, showing that — just as with other distance probes — external data to break parameter degeneracies is necessary. Only if $\Omega_{\rm m}$ is well constrained does the uncertainty on h reduce to the 1–2% level (still under the assumption

JCAP03(2020)019

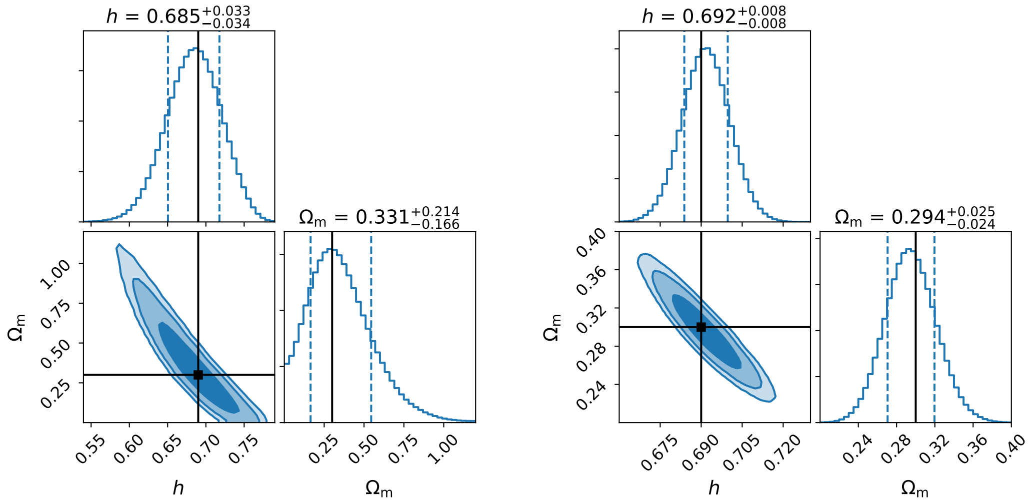

Figure 2. Forecast posterior for mock data generated from a ΛCDM cosmology for future (left) and far future (right) sensitivities. The 2D posterior for h and Ω_m shows the 68.3%, 95.4%, 99.7% confidence regions in increasingly lighter shades of blue. The 1D posteriors show the 68.3% confidence regions in dashed blue. The input values for Ω_m and h are indicated with solid black lines. Note the scales for the parameter ranges are different.

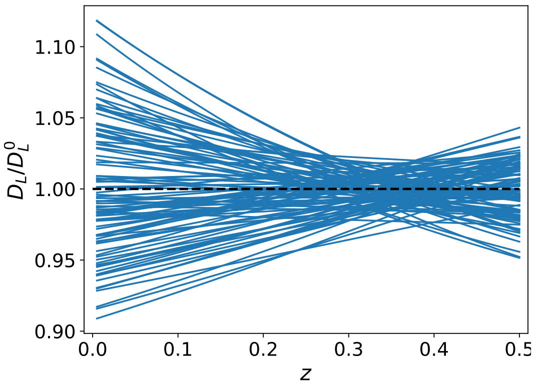

Figure 3. The distance uncertainty is shown as a scatter plot from a posterior predictive distribution sampling of the distances (relative to the input cosmology), for future GW data.

of ΛCDM). Of course, if the external data has systematics that shift the value of Ω_m, then the value of h derived from the GW plus this data will be biased.

Another, intuitive way of seeing the difficulty in GW sirens (or any cosmic distance) in determining cleanly H_0 is provided in figure 3. We plot realizations of the ratio of GW luminosity distances for ΛCDM cosmologies with parameters drawn from the posterior relative to the input cosmology. The size of the scatter at different redshifts indicates at which redshifts the distance is better constrained. Both the parameter covariances and the number of events at a given redshift enter into the scatter. Unfortunately, redshift zero and hence $D_L(z \ll 1) = H_0^{-1} z$ has large uncertainty. Thus we do not expect H_0 to be constrained near the 1% level when we are appropriately agnostic about the expansion history.

If we allow for the standard w_0-w_a dynamical dark energy freedom in the background we find an even bleaker picture. Now, even with the matter density perfectly known, GW

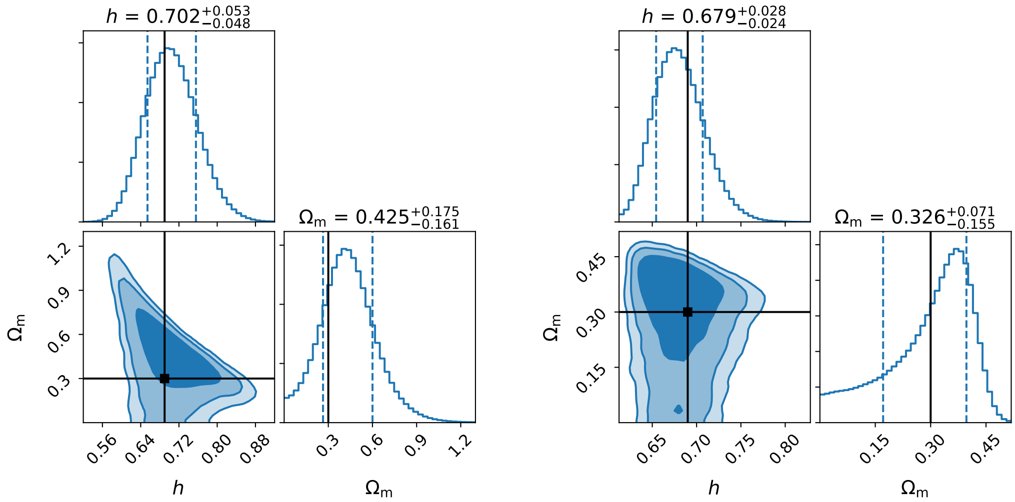

Figure 4. Forecast posterior for mock data generated from a ΛCDM cosmology, but marginalizing over w_0 and w_a in the MCMC fit, for future (left) and far future (right) sensitivities. Note the scales for the parameter ranges are different.

sirens deliver only $\sigma(h) = 0.135$ — 16 times worse than the ΛCDM case — due to the degeneracy with w_0 and w_a. If we add extra Gaussian priors on w_0 and w_a of width 1 to cut the degeneracy, then $\sigma(h) = 0.039$, i.e. 6%, not better than current estimates. These Fisher information constraints are also shown in table 1, however all our actual numbers come from the full MCMC analysis. With Ω_m as a fit parameter, the covariance increases the MCMC uncertainty on h, even in the far future generation case. The MCMC contours in the Ω_m-h plane are shown in figure 4. Thus the cosmological model assumed (e.g. freedom beyond ΛCDM) plays a critical role in the constraints from GW sirens at cosmic distances.

4 Bias vs cosmological model

Apart from the issue of precision on H_0 and Ω_m, if insufficient freedom is given to the background cosmology fit then the resulting value of H_0 (and Ω_m) will be biased. We explore the magnitude of this effect by choosing two alternative w_0-w_a model points on the 68% confidence contour of the current joint probe analysis in [32] and generating GW data sets in these cosmologies. If these are then analyzed within the ΛCDM model, parameter biases ensue. The two cosmologies used are $(w_0,w_a)=(-0.90,-0.75)$ and $(-1.14,0.35)$, with $\Omega_\mathrm{m} = 0.3$, $h = 0.69$, and by construction are consistent with current combined data sets. Should future probes further constrain the w_0-w_a parameter space towards ΛCDM then by the time future and far future detectors come online, then any potential cosmological bias would be less.

After generating our mock luminosity distance dataset from GW+EM observations, we then use MCMC sampling to infer the 1D parameter fits and 2D joint confidence contours for a ΛCDM model. This allows us to assess the bias induced by the incorrect background model assumption, and its significance relative to the statistical precision. The data quality will be particularly important for this last question: as the precision improves a given bias becomes more important. Therefore we study both future and far future GW data sets.

JCAP03(2020)019

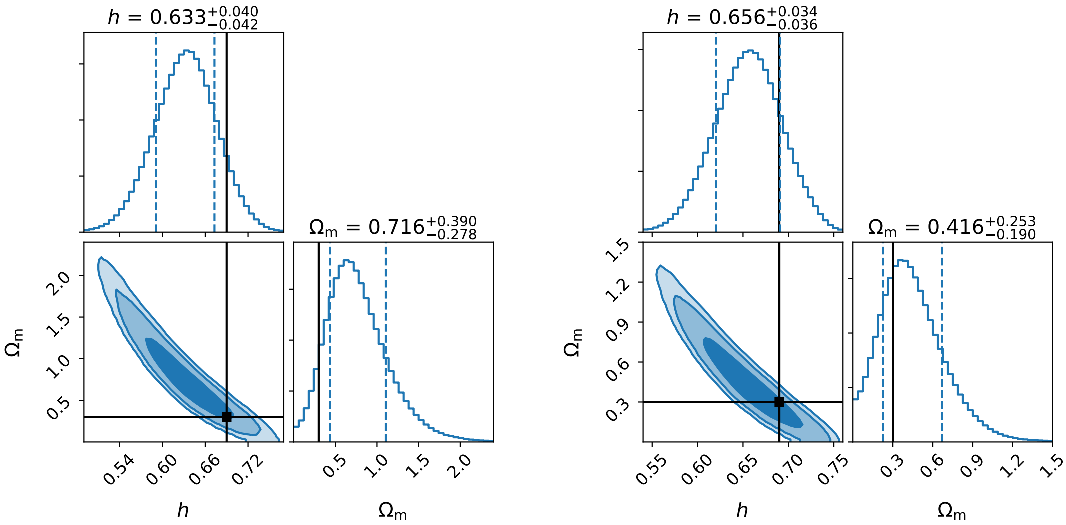

Figure 5. Forecast posteriors of the ΛCDM parameters for the future case, where the distances to the GW+EM events are generated from currently viable (w_0,w_a) cosmologies. The left panel is for $(w_0, w_a) = (-0.9, -0.75)$ and the right panel is for $(-1.14, 0.35)$.

We present the results for the future case in figure 5. The precision obtained for the two input models is comparable, with ∼6% uncertainty on H_0 and ∼50% on Ω_m. Clearly future GW alone will not give the desired constraint. The bias, due to misassuming ΛCDM, shifts the fit contours so that the true input values are at the edge of 68% confidence contour in each case.

Figure 6 repeats the analysis for the far future data case. The parameter precision is now strongly improved, to 1.1% on H_0 and 8% on Ω_m. However, the bias is much more severe, with the true values lying outside the 99.7% joint confidence contour, i.e. roughly 3σ bias (although the 1D values do not accurately show this tension). Note that knowing the actual value of matter density here from other observations could increase the bias in estimation of $H0$ due to our wrong assumption of the background expansion model.

The biases are evident even in a simple constant w extension to ΛCDM. If we take $w = -1.10$ from the 68% confidence contour of the current joint data constraint of [32], figure 7 shows the input values of (h, Ω_m) have been biased to outside the 99.7% joint confidence contour in the ΛCDM analysis (indeed to more than the equivalent of 5σ).

From eq. (2.1) and figure 3 we can see that the bias must exist if the cosmological framework used in the fit does not cover the true cosmological model. For GW data to constrain primarily H_0, then $h(z)$ should be indistinguishable from 1 at the level of the statistical precision. Similarly, if we go beyond H_0 to include Ω_m, then to fit these parameters without bias the $h(z)$ for the ΛCDM cosmology fit should be indistinguishable from the true (potentially non-ΛCDM) cosmology, again at the level of the statistical precision.

The direction and magnitude of the parameter biases is a combination of the data properties (e.g. redshift distribution), cosmological parameter covariances, and parameter values. We have verified the MCMC results through the analytic Fisher bias formalism [33, 34], which makes these dependencies more explicit. The bias on a parameter (such as h or

JCAP03(2020)019

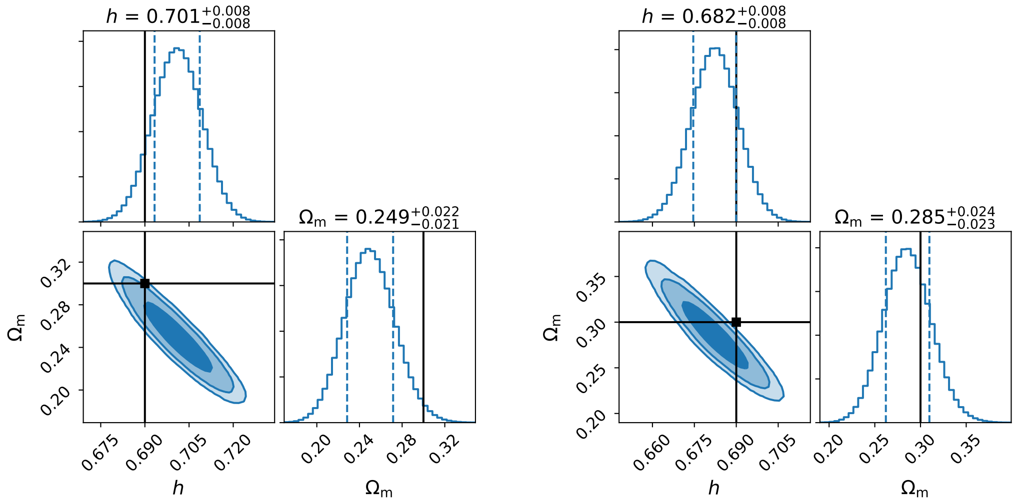

Figure 6. As figure 5 but for far future data.

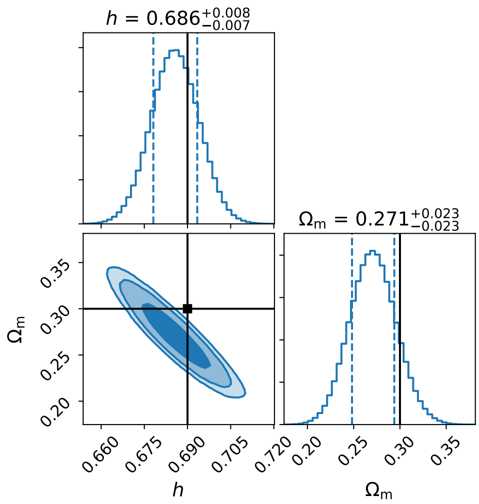

Figure 7. As figure 6 but for distances generated from a constant $w = -1.1$ cosmology.

$\Omega_{\rm m}$) is given by

$$\delta p_i = \left(F_{\rm sub}^{-1}\right)_{ij} \sum_k \frac{\partial \mathcal{O}_k}{\partial p_j} \frac{1}{\sigma_k^2} \Delta\mathcal{O}_k \tag{4.1}$$

$$= \left(F_{\rm sub}^{-1}\right)_{ij} \left[(1+w_0)F_{jw_0} + w_a F_{jw_a}\right] , \tag{4.2}$$

where $\Delta\mathcal{O}_k$ is the difference between the true distance and in the assumed model. Here the second line holds when the assumed cosmological model is a subset of a cosmological model with additional parameters, as ΛCDM is a subset of a w_0-w_a cosmology. This parameter bias estimation is in good agreement with the full Monte Carlo analysis we use.

5 Discussions and conclusions

Gravitational wave sirens in conjunction with electromagnetic counterparts provide a new distance measure for the Universe. New and improved detectors with better sensitivities and further redshift reach are exciting developments that will deliver abundant science from a significant numbers of events. While GW provide an absolute distance measurement they are not a panacea — covariance between the evolution of the expansion $H(z)$ and the absolute scale today H_0 still exists.

We have quantified how assuming that the expansion history $H(z)$ is purely of the ΛCDM form to infer the value of H_0 and Ω_m from cosmic GW luminosity distance data (and not only GW sources at $z \lesssim 0.1$) will yield inaccurate results should the true cosmology be different than ΛCDM. This holds even if the deviation from ΛCDM cosmology is modest, within the 68% confidence level constraint from the current combination of data from several probes. Indeed we find that even a constant w model within the current 68% joint confidence contour can deliver almost a 5σ bias if inappropriately analyzed within ΛCDM. Fixing the value of the matter density Ω_m within ΛCDM, as is sometimes adopted in predicting 1% precision on H_0 from GW, is further problematic for accuracy.

Future GW events will probe deeper into the universe, so all the freedom that enters into $H(z)$, from the imperfectly known matter density Ω_m and dark energy properties, will both dilute the precision and open up the potential for bias as we try to project distances to the very low redshift behavior involving only the Hubble constant H_0. Being properly agnostic about the expansion history translates into uncertainties in H_0 that are well above 1%. To quantify this, we carry out a Monte Carlo analysis simulating the GW+EM event distance data for future generation and far future generation experiments.

Future generation data reaches 1.2% precision on H_0 only if both ΛCDM is assumed and Ω_m is perfectly known, with a degradation to 1.4% if ΛCDM is assumed and an external prior on Ω_m is used. For cosmic GW themselves, the precision is 5% when restricted to ΛCDM. Allowing for uncertainty in the cosmological model by including dynamical dark energy such as with w_0, w_a dilutes the precision to 7% — barely more constraining than the single local GW binary neutron star event already measured [35, 36]. Thus cosmic GW data can clearly not be implemented as a pure prior on key cosmological parameters such as H_0.

This is in no way a failing of GW data. Any cosmic distance measurement has the same issues with covariances (and note strong lensing time delays involve H_0 in a similar way to GW), and potential biases if unduly restricted to the wrong expansion model.

The biases exist if fixing to the wrong Ω_m, the wrong constant w, or in general assuming a wrong model or form of dark energy. (Note that the w_0-w_a form does fit the distances out to $z = 1$ to 0.1% in a wide variety of viable models [37].) This is also the case when we use a wrong prior on H_0 trying to estimate the other cosmological quantities such as matter density or equation of state of dark energy. In this work we show that for the case of far future generation GW data we can have more than 3σ bias in estimation of H_0 while precision of the estimation can be very tight at 1.1%. One safe approach to advocate is to carry out the analysis with model independent reconstruction techniques.

If one could achieve substantial samples of GW+EM events at $z \lesssim 0.05$ then most of the parameter covariance vanishes and one does get purer determination of H_0 independent of background cosmological model (modulo issues of peculiar velocities and coherent flows). This would be an exciting prospect, though uncertainty in event rates could impact the leverage on H_0.

Acknowledgments

We thank KIAS, where this project was discussed among the co-authors, and especially Stephen Appleby for hospitality. A.S. would like to acknowledge the support of the National Research Foundation of Korea (NRF- 2016R1C1B2016478). E.L. is supported in part by the Energetic Cosmos Laboratory and by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award DE-SC-0007867 and contract no. DE-AC02-05CH11231.

The datasets generated and analysed during the current study are available in the GWCosmology repository, https://github.com/rekeeley/GWCosmology.

References

[1] B.F. Schutz, *Determining the Hubble constant from gravitational wave observations*, Nature **323** (1986) 310 [INSPIRE].

[2] W. Del Pozzo, *Inference of the cosmological parameters from gravitational waves: application to second generation interferometers*, Phys. Rev. **D 86** (2012) 043011 [arXiv:1108.1317] [INSPIRE].

[3] D.E. Holz and S.A. Hughes, *Using gravitational-wave standard sirens*, Astrophys. J. **629** (2005) 15 [astro-ph/0504616] [INSPIRE].

[4] N. Dalal, D.E. Holz, S.A. Hughes and B. Jain, *Short grb and binary black hole standard sirens as a probe of dark energy*, Phys. Rev. **D 74** (2006) 063006 [astro-ph/0601275] [INSPIRE].

[5] P. Zhang, *Accurate redshift determination of standard sirens by the luminosity distance space-redshift space large scale structure cross correlation*, arXiv:1811.07136 [INSPIRE].

[6] LIGO Scientific, Virgo collaborations, *A standard siren measurement of the Hubble constant from GW170817 without the electromagnetic counterpart*, Astrophys. J. **871** (2019) L13 [arXiv:1807.05667] [INSPIRE].

[7] H.-Y. Chen, M. Fishbach and D.E. Holz, *A two per cent Hubble constant measurement from standard sirens within five years*, Nature **562** (2018) 545 [arXiv:1712.06531] [INSPIRE].

[8] D.J. Mortlock, S.M. Feeney, H.V. Peiris, A.R. Williamson and S.M. Nissanke, *Unbiased Hubble constant estimation from binary neutron star mergers*, Phys. Rev. **D 100** (2019) 103523 [arXiv:1811.11723] [INSPIRE].

[9] L. Hui and P.B. Greene, *Correlated fluctuations in luminosity distance and the (surprising) importance of peculiar motion in supernova surveys*, Phys. Rev. **D 73** (2006) 123526 [astro-ph/0512159] [INSPIRE].

[10] A. Cooray and R.R. Caldwell, *Large-scale bulk motions complicate the hubble diagram*, Phys. Rev. **D 73** (2006) 103002 [astro-ph/0601377] [INSPIRE].

[11] KAGRA, LIGO Scientific, VIRGO collaborations, *Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA*, Living Rev. Rel. **21** (2018) 3 [arXiv:1304.0670] [INSPIRE].

[12] LIGO Scientific collaboration, *Advanced LIGO*, Class. Quant. Grav. **32** (2015) 074001 [arXiv:1411.4547] [INSPIRE].

[13] VIRGO collaboration, *Advanced Virgo: a second-generation interferometric gravitational wave detector*, Class. Quant. Grav. **32** (2015) 024001 [arXiv:1408.3978] [INSPIRE].

[14] B. Iyer et al., *LIGO-India, proposal of the consortium for indian initiative in gravitational-wave observations (IndIGO)*, *Proposal* (2011) M1100296-v2.

[15] KAGRA collaboration, *Interferometer design of the KAGRA gravitational wave detector*, Phys. Rev. **D 88** (2013) 043007 [arXiv:1306.6747] [INSPIRE].

[16] N. Tamanini, C. Caprini, E. Barausse, A. Sesana, A. Klein and A. Petiteau, *Science with the space-based interferometer eLISA. III: probing the expansion of the Universe using gravitational wave standard sirens*, JCAP **04** (2016) 002 [arXiv:1601.07112] [INSPIRE].

[17] P. Amaro-Seoane et al., *eLISA/NGO: astrophysics and cosmology in the gravitational-wave millihertz regime*, *GW Notes* **6** (2013) 4 [arXiv:1201.3621] [INSPIRE].

[18] LISA collaboration, *Laser interferometer space antenna*, arXiv:1702.00786 [INSPIRE].

[19] B. Sathyaprakash et al., *Scientific objectives of Einstein telescope*, Classical Quant. Grav. **29** (2012) 124013.

[20] S.R. Taylor and J.R. Gair, *Cosmology with the lights off: standard sirens in the Einstein telescope era*, Phys. Rev. **D 86** (2012) 023502 [arXiv:1204.6739] [INSPIRE].

[21] R.-G. Cai and T. Yang, *Estimating cosmological parameters by the simulated data of gravitational waves from the Einstein telescope*, Phys. Rev. **D 95** (2017) 044024 [arXiv:1608.08008] [INSPIRE].

[22] E. Belgacem, Y. Dirian, S. Foffa and M. Maggiore, *Modified gravitational-wave propagation and standard sirens*, Phys. Rev. **D 98** (2018) 023510 [arXiv:1805.08731] [INSPIRE].

[23] J.-Z. Qi, S. Cao, Y. Pan and J. Li, *Cosmic opacity: cosmological-model-independent tests from gravitational waves and Type Ia Supernova*, Phys. Dark Univ. **26** (2019) 100338 [arXiv:1902.01702].

[24] J.-F. Zhang, M. Zhang, S.-J. Jin, J.-Z. Qi and X. Zhang, *Cosmological parameter estimation with future gravitational wave standard siren observation from the Einstein telescope*, JCAP **09** (2019) 068 [arXiv:1907.03238] [INSPIRE].

[25] PLANCK collaboration, *Planck 2018 results. VI. Cosmological parameters*, arXiv:1807.06209 [INSPIRE].

[26] A.G. Riess et al., *A 2.4% Determination of the local value of the Hubble constant*, Astrophys. J. **826** (2016) 56 [arXiv:1604.01424] [INSPIRE].

[27] A.G. Riess, S. Casertano, W. Yuan, L.M. Macri and D. Scolnic, *Large Magellanic cloud cepheid standards provide a 1% foundation for the determination of the Hubble constant and stronger evidence for physics beyond* ΛCDM, Astrophys. J. **876** (2019) 85 [arXiv:1903.07603] [INSPIRE].

[28] S.M. Feeney et al., *Prospects for resolving the Hubble constant tension with standard sirens*, Phys. Rev. Lett. **122** (2019) 061105 [arXiv:1802.03404] [INSPIRE].

[29] E. Di Valentino, D.E. Holz, A. Melchiorri and F. Renzi, *The cosmological impact of future constraints on* H_0 *from gravitational-wave standard sirens*, Phys. Rev. **D 98** (2018) 083523 [arXiv:1806.07463] [INSPIRE].

[30] M. Du, W. Yang, L. Xu, S. Pan and D.F. Mota, *Future constraints on dynamical dark-energy using gravitational-wave standard sirens*, Phys. Rev. **D 100** (2019) 043535 [arXiv:1812.01440] [INSPIRE].

[31] M.J. Mortonson, W. Hu and D. Huterer, *Hiding dark energy transitions at low redshift*, Phys. Rev. **D 80** (2009) 067301 [arXiv:0908.1408] [INSPIRE].

[32] D.M. Scolnic et al., *The complete light-curve sample of spectroscopically confirmed SNe Ia from Pan-STARRS1 and cosmological constraints from the combined pantheon sample*, Astrophys. J. **859** (2018) 101 [arXiv:1710.00845] [INSPIRE].

[33] L. Knox, R. Scoccimarro and S. Dodelson, *The impact of inhomogeneous reionization on cosmic microwave background anisotropy*, *Phys. Rev. Lett.* **81** (1998) 2004 [astro-ph/9805012] [INSPIRE].

[34] E.V. Linder, *Biased cosmology: pivots, parameters and figures of merit*, *Astropart. Phys.* **26** (2006) 102 [astro-ph/0604280] [INSPIRE].

[35] LIGO Scientific, Virgo, 1M2H, Dark Energy Camera GW-E, DES, DLT40, Las Cumbres Observatory, VINROUGE, MASTER collaborations, *A gravitational-wave standard siren measurement of the Hubble constant*, *Nature* **551** (2017) 85 [arXiv:1710.05835] [INSPIRE].

[36] LIGO Scientific, Virgo, Fermi GBM, INTEGRAL, IceCube, AstroSat Cadmium Zinc Telluride Imager Team, IPN, Insight-Hxmt, ANTARES, Swift, AGILE Team, 1M2H Team, Dark Energy Camera GW-EM, DES, DLT40, GRAWITA, Fermi-LAT, ATCA, ASKAP, Las Cumbres Observatory Group, OzGrav, DWF (Deeper Wider Faster Program), AST3, CAASTRO, VINROUGE, MASTER, J-GEM, GROWTH, JAGWAR, CaltechNRAO, TTU-NRAO, NuSTAR, Pan-STARRS, MAXI Team, TZAC Consortium, KU, Nordic Optical Telescope, ePESSTO, GROND, Texas Tech University, SALT Group, TOROS, BOOTES, MWA, CALET, IKI-GW Follow-up, H.E.S.S., LOFAR, LWA, HAWC, Pierre Auger, ALMA, Euro VLBI Team, Pi of Sky, Chandra Team at McGill University, DFN, ATLAS Telescopes, High Time Resolution Universe Survey, RIMAS, RATIR, SKA South Africa/MeerKAT collaborations, *Multi-messenger observations of a binary neutron star merger*, *Astrophys. J.* **848** (2017) L12 [arXiv:1710.05833] [INSPIRE].

[37] R. de Putter and E.V. Linder, *Calibrating dark energy*, *JCAP* **10** (2008) 042 [arXiv:0808.0189] [INSPIRE].

JCAP03(2020)019