

A dark siren measurement of the Hubble constant using gravitational wave events from the first three LIGO/Virgo observing runs and DELVE

V. Alfradique   1★ C. R. Bom  1,2 A. Palmese, 3 G. Teixeira, 1 L. Santana-Silva  1 A. Drlica-Wagner, 4,5,6 A. H. Riley, 7 C. E. Martínez-Vázquez  8 D. J. Sand, 9 G. S. Stringfellow  10 G. E. Medina  11 J. A. Carballo-Bello, 12 Y. Choi, 13 J. Esteves, 14 G. Limberg  5,6,15 B. Mutlu-Pakdil  16 N. E. D. Noël, 17 A. B. Pace  3 J. D. Sakowska  17 and J. F. Wu  18,19

Affiliations are listed at the end of the paper

Accepted 2024 January 8. Received 2023 December 19; in original form 2023 September 29

ABSTRACT

The current and next observation seasons will detect hundreds of gravitational waves (GWs) from compact binary systems coalescence at cosmological distances. When combined with independent electromagnetic measurements, the source redshift will be known, and we will be able to obtain precise measurements of the Hubble constant H_0 via the distance–redshift relation. However, most observed mergers are not expected to have electromagnetic counterparts, which prevents a direct redshift measurement. In this scenario, one possibility is to use the dark sirens method that statistically marginalizes over all the potential host galaxies within the GW location volume to provide a probabilistic source redshift. Here we presented H_0 measurements using two new dark sirens compared to previous analyses using DECam data: GW190924_021846 and GW200202_154313. The photometric redshifts of the possible host galaxies of these two events are acquired from the DECam Local Volume Exploration Survey (DELVE) carried out on the Blanco telescope at Cerro Tololo. The combination of the H_0 posterior from GW190924_021846 and GW200202_154313 together with the bright siren GW170817 leads to $H_0 = 68.84^{+15.51}_{-7.74} \text{ km s}^{-1} \text{ Mpc}^{-1}$. Including these two dark sirens improves the 68 percent confidence interval (CI) by 7 percent over GW170817 alone. This demonstrates that the addition of well-localized dark sirens in such analysis improves the precision of cosmological measurements. Using a sample containing 10 well-localized dark sirens observed during the third LIGO/Virgo observation run, without the inclusion of GW170817, we determine a measurement of $H_0 = 76.00^{+17.64}_{-13.45} \text{ km s}^{-1} \text{ Mpc}^{-1}$.

Key words: gravitational waves – catalogues – surveys – cosmology: observations.

1 INTRODUCTION

The advent of gravitational wave measurements opened a new era of multimessenger observation, shedding light on the properties of our Universe. Standard sirens, a term introduced by Schutz (1986), provide a way to measure cosmological parameters by restricting the distance–redshift relation. The gravitational wave detections provide a direct measure of luminosity distance without any additional distance calibrator, justifying the name ‘standard sirens’ in analogy with standard candles. If a source has an electromagnetic counterpart, its redshift (z) can be directly measured, and we referred to them as ‘bright standard sirens’. The first bright standard siren measured was the binary neutron star (BNS) merger GW170817 (Abbott et al. 2017b), whose electromagnetic gamma-ray burst counterpart was detected by the Fermi Gamma-ray Burst Monitor (Goldstein et al. 2017) and the anticoincidence shield of the gamma-ray spectrometer on-board INTErnational Gamma-Ray Astrophysics Laboratory (Savchenko et al. 2017) within 0.1–0.647 s, and later complementing with the identification of the optical kilonova (e.g. Arcavi et al. 2017;

Chornock et al. 2017; Coulter et al. 2017; Cowperthwaite et al. 2017; Evans et al. 2017; Kasliwal et al. 2017; Nicholl et al. 2017; Pian et al. 2017; Smartt et al. 2017; Soares-Santos et al. 2017; Tanvir et al. 2017; Valenti et al. 2017) detected about 11 h after the merger. This event produced the first direct and independent measure of H_0 , $H_0 = 70^{+12}_{-8} \text{ km s}^{-1} \text{ Mpc}^{-1}$ (Abbott et al. 2017d).

After a 3-yr hiatus during which improvements in the sensitivity of the detectors were made, the upcoming fourth run of the LIGO, Virgo, and KAGRA collaboration will be able to observe a larger fraction of the universe than previous observing runs and projected to detect an estimated ~ 90 gravitational wave events per year with an ~ 85 percent improvement in sky localization expected by the end of the run if all detectors are operating at their target sensitivities (Abbott et al. 2018). With more interferometers in operation (like the Einstein Telescope, Sathyaprakash et al. 2012; Cosmic Explorer, Abbott et al. 2017a; and the LISA space interferometer, Amaro-Seoane et al. 2017), it is possible that in the next years more standard sirens will be identified which can lead to an H_0 measurement with precision in the same order as what is achieved with other cosmological probes such as the cosmological microwave background (CMB; Planck Collaboration 2020) and the Cepheid (Riess et al. 2021) or Red Giant Branch (Freedman et al. 2019)–

* E-mail: vivianeapa@cbpf.br

calibrated type Ia supernovae. This new independent measurement of the Hubble constant can enable a way to clarify the origin of the observed current $4 - 6\sigma$ tension (Verde, Treu & Riess 2019; Di Valentino et al. 2021). Despite these improvements, detection of the events electromagnetic counterparts remains a challenge, requiring dedicated follow-up campaigns and strategies (Bom et al. 2023), particularly for those events involving black hole companions, which may have no electromagnetic signature emitted or be associated to a flare (Bom & Palmese 2023; Rodríguez-Ramírez et al. 2023).

A prime example of the challenge to localize and identify the electromagnetic emission of an event involving a black hole is the GW190814 event (Abbott et al. 2020b). GW190814 was the result of the coalescence of a $23.2M_{\odot}$ black hole with a compact object of mass $2.5 - 2.67 M_{\odot}$. Since the secondary mass lies in the mass gap region, this object was either the massive neutron star or the lightest black hole ever seen in a binary system. Due to its excellent sky localization (23 deg^2), this event becomes a great candidate to provide the first detection of the counterpart of a binary system involving at least one black hole, hoping to shed light on the nature of this compact system. Several electromagnetic follow-ups, from gamma rays to radio, were started by different groups (e.g. Dobie et al. 2019; Gomez et al. 2019; Ackley et al. 2020; Andreoni et al. 2020; Vieira et al. 2020; Watson et al. 2020; Alexander et al. 2021; Kilpatrick et al. 2021; Tucker et al. 2021; de Wet et al. 2021) with a continuous duration of up to more than 250 d after the merger. The properties of the electromagnetic counterpart candidates were analysed and compared with the theoretical prediction for NSBH fusion, including optical spectra, variability of radio sources, their location, photometric evolution, and redshift of possible host galaxies. Despite immense dedicated effort, no sign of a gamma-ray burst or any optical counterpart has been identified, but allowed to discard some possible types of electromagnetic transients such as: kilonova with large ejecta mass $M \geq 0.1M_{\odot}$ (Ackley et al. 2020), ‘blue’ kilonovae with $M > 0.5M_{\odot}$ (Kilpatrick et al. 2021), an AT2017gfo-like kilonova (de Wet et al. 2021), short gamma-ray burst with viewing angles less than 17° (Kilpatrick et al. 2021), and a short gamma-ray burst-like Gaussian jet with a particular configuration (Alexander et al. 2021). In view of these problems, an alternative to the lack of an electromagnetic counterpart is to use the redshifts of galaxies that are within the coalescence location volume to break the H_0 - z degeneracy and infer cosmological parameters. This methodology is known as *dark standard sirens* (see Gair et al. 2023 for a review of the method).

The dark standard sirens approach was applied to constrain the cosmology in several LIGO and Virgo detections. Fishbach et al. (2019) studied the event GW170807 and showed that the obtained precision of H_0 is about 3 times worse than the ‘bright’ siren method (Abbott et al. 2017d). Soares-Santos et al. (2019) and Palmese et al. (2020) investigated the method with the Dark Energy Survey (DES) galaxy catalogue for binary black hole (BBH) mergers (GW170814 and GW190814, respectively) and showed that a single dark siren BBH provides a measure of H_0 with a precision of 48 per cent for GW170814 and 55 per cent GW190814. Recently, Palmese et al. (2023) demonstrated that 8 dark sirens well localized in the sky are able to provide a measurement as accurate as that obtained with a single bright siren GW170817 (about 20 per cent against 18 per cent; Abbott et al. 2017d).

Chen, Fishbach & Holz (2018) predicted that 5 years of detections for LIGO, Virgo and KAGRA collaboration (at design sensitivity) could lead to a precision of ~ 5 per cent and 10 per cent of H_0 measurement for the BNS and BBH, respectively. For this result, they assumed that all events within $10\,000 \text{ Mpc}^3$ will be detected and that

complete galaxy catalogues will be available. In the next decade, the arrival of the next generation of terrestrial interferometers, such as the Einstein Telescope and the Cosmic Explorer, could rapidly increase the number of detections, allowing us to check the predictions of the percentage level of the measure of H_0 made by Muttoni et al. (2023).

The intent of this study is to investigate the ability of the dark siren events GW190924_021846 and GW200202_154313 to constrain the Hubble constant. We combine our results with that of 8 dark sirens present in Palmese et al. (2023) and perform the most precise H_0 measurement with the better localized dark sirens. The choice for these events is justified due to the small localization volume, which decreases the number of potential host galaxies to be marginalized over, and because their localization region is covered by DELVE¹ (Drlica-Wagner et al. 2021) galaxy catalogues. All the photometry redshift information is provided by the second release of DELVE data (DELVE DR2², Drlica-Wagner et al. 2022), the galaxy photometric redshift was estimated using the Mixture Density Network (MDN, Bishop 1994), a machine learning technique that provides the probability density function (PDF) of the photo- z . This technique uses magnitudes and colour information to train the various Gaussian distributions that will be combined into the final PDF. In contrast to previous work, our work innovates by applying DELVE data to the standard siren methodology for the first time for two new events from the third observing run (O3), implementing a more refined artificial neural network technique for photo- z measurements instead of the commonly used random forest algorithms (Zhou et al. 2020; Mucesh et al. 2021). The results of this study may provide insight into the potential of dark sirens as a cosmological probe, computing the precision level of H_0 measurement that this methodology can achieve with realistic photometric uncertainty and sky coverage.

This paper is organized as follows: in Section 2 we describe the data used in the dark sirens methodology that is discussed in Section 3. Our results are presented and discussed in Section 4, and our final conclusions are presented in Section 5. Throughout the article, we adopt a flat Λ CDM cosmology with $\Omega_m = 0.3$ and H_0 values in the $20 - 140 \text{ km s}^{-1} \text{ Mpc}^{-1}$ range. When not otherwise stated, quoted error bars represent the 68 per cent CI.

2 DATA

2.1 LIGO and Virgo data: gravitational wave events

Here, we extend the eight-event catalogue used in Palmese et al. (2023) by adding two new events: GW190924_021846 and GW200202_154313. In total, our sample includes the 10 best localized events in the sky detected during the third LIGO/Virgo observing period. For these two added events, we used the gravitational wave data from the maps publicly available by the LIGO and Virgo collaboration in Abbott et al. (2021a) and Collaboration et al. (2021). The right ascension (RA), declination (dec), and distance probability are given in HEALPIX pixels (Górski et al. 2005), where this probability is supposed to be Gaussian along each line of sight. GW200202_154313 is the result of the merger of two black holes of approximately 7 and 10 solar masses, this is one of the best three-dimensional localizations from the second-half of the O3 (see Table 1), having a 90 per cent credible volume of 0.0034 Gpc^3 and a

¹<https://delve-survey.github.io/>

²<https://datalab.noirlab.edu/delve/>

Table 1. Luminosity distance, 90 per cent CI area, and volume of gravitational wave events and candidates used in this analysis. We also report the reference paper or GCN that reports the sky map used for each event. These events have estimated false alarm rates of fewer than 1 in 10^3 – 10^{23} years. These candidates have all recently been confirmed as gravitational wave events in Collaboration et al. (2021).

Event	d_L [Mpc]	A [deg 2]	V [Gpc 3]	Reference
GW170608	320^{+120}_{-110}	392	3×10^{-3}	Abbott et al. (2019)
GW170814	540^{+130}_{-210}	62	2×10^{-3}	Abbott et al. (2017c)
GW170818	1060^{+420}_{-380}	39	7×10^{-3}	Abbott et al. (2019)
GW190412	740^{+120}_{-130}	12	4×10^{-4}	Abbott et al. (2020a)
GW190814	241^{+26}_{-26}	19	3×10^{-5}	Abbott et al. (2020b)
S191204r	678^{+149}_{-149}	103	6×10^{-3}	LIGO Scientific Collaboration and Virgo Collaboration (2019)
S200129m	755^{+194}_{-194}	41	2×10^{-3}	LIGO Scientific Collaboration and Virgo Collaboration (2020a)
S200311bg	1115^{+175}_{-175}	34	5×10^{-3}	LIGO Scientific Collaboration and Virgo Collaboration (2020b)
GW200202_154313	410^{+150}_{-160}	167	3×10^{-3}	Collaboration et al. (2021)
GW190924_021846	570 ± 220	348	2×10^{-2}	Abbott et al. (2021a)

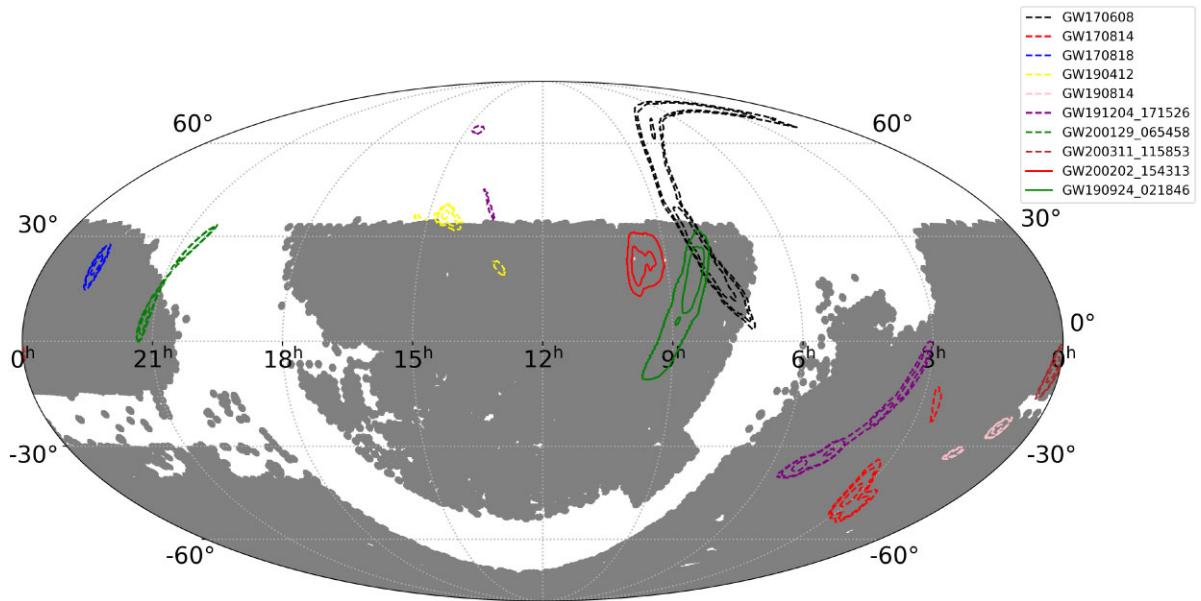


Figure 1. LIGO/Virgo GW dark standard sirens analysed in this paper, where the contours represent the 90 per cent CI localization from the sky maps. The shaded regions are those that are covered by the DELVE catalogues used in this work.

90 per cent CI area of 167 deg 2 . The location of maximum probability is centred at RA = 146.25 deg and dec = 20.98 deg. Marginalizing over all other parameters, the estimate of the luminosity distance has a mean equal to 364.3 Mpc with a standard deviation of 90.2 Mpc. The second detection, GW190924_021846 is probably the result of the merger of two lowest-mass black holes so far observed (Abbott et al. 2021a), with inferred masses equal to $m_1 = 8.9^{+7.0}_{-2.0} M_\odot$, $m_2 = 5.0^{+1.4}_{-1.9} M_\odot$ and a 90 per cent credible volume of ~ 0.02 Gpc 3 and a 90 per cent CI area of 348 deg 2 . The GW190924_021846 has a maximum probability of being located at RA, dec = (134.561, 2.687) deg. At the peak location, the luminosity distance mean is 479.4 Mpc and the standard deviation is 151.7 Mpc. The area of the sky enclosing 90 per cent CI is 348 deg 2 . Fig. 1 shows the 90 per cent CI contours of all events used in this work with the area covered by DELVE. The two new events are fully covered by DELVE, except for a small tip on GW190924_021846 containing approximately 2 per cent of the 90 per cent CI localization region.

2.2 The galaxies photo- z ’s: a deep learning algorithm for DELVE data

The DELVE is a project that combines public data (including data from DES, Dark Energy Camera Legacy Survey and DECam eROSITA Survey) with more than 126 nights of novel observations made with the Dark Energy Camera (DECam, Flaugher et al. 2015) located on the 4-meter Blanco Telescope at Cerro Tololo Inter-American Observatory in Chile. DELVE uses the combination of the large field of view (3 deg 2) and the fast readout time (27 s) of DECam with the expectation of providing complete coverage of the entire high-Galactic-latitude southern sky. In the WIDE observational program, the DELVE DR2 (Drlica-Wagner et al. 2022) covers $> 17,000$ deg 2 in the $griz$ bands out to 23.5 mag.

The photometric redshifts (or photo- z s) for the DELVE data were computed using the deep learning method called Mixture Density Network. In brief, the method is a combination of a deep neural network with the assumption that any distribution can be written as a mixture of distributions (chosen to be the normal distribution

in its traditional form). The deep neural network is trained, given some input features, to select the best parameters of the multiple distributions that will be mixed into a single distribution. The output parameters used are the mean, standard deviation, and mixing coefficients, which are the probabilistic weights of each normal distribution. In this way, the MDN is capable to reproduce the galaxy photo- z PDF, given some input features. The input features are the $griz$ magnitudes, and the $g - r, g - i, g - z, r - i, r - z, i - z$ colours. In the next sections, we use this approach to compute the photo- z s of the possible galaxy host whenever the spectroscopic redshift is not available.

The MDN was implemented with the following structure: a LMU layer with 212 units; a 2-layer Multi-Layer Perceptron with 96 units each; a Dropout layer with 20 per cent rate; and finally a `MixtureNormal` layer that returns the outputs (the mean, standard-deviation and weights of the 20 Gaussian distributions). The LMU layer was implemented using the `keras-lmu` Voelker, Kajić & Eliasmith (2019a) application; the inner Perceptron and Dropout layers, the standard DL framework and the `MixtureNormal` output layer were built within the `tensorflow` and `tensorflow-probability` libraries API³ (Abadi, Agarwal & et al. 2015). The architecture of the network also incorporates a Legendre Memory Unit (LMU, Voelker, Kajić & Eliasmith 2019b) Layer at the head of the network. This architecture was one of the networks submitted in the LSST-DESC Tomography Optimization Challenge (Zuntz et al. 2021), and it exhibited the best performance for the DELVE DR2 photo- z 's regression task. We combined the photo- z PDF estimated by the MDN output layer (also used for photometric redshift regression in the S-PLUS Survey in Lima et al. 2022) with the well-performing LMU layer to estimate the photometric redshifts. More details can be found in Teixeira et al. Following the work of Zuntz et al. (2021), the LMU layer is included to more efficiently assign galaxies to redshift bins, selecting relevant information from previous data while simultaneously removing expendable data. For the loss function, we chose the maximum likelihood, which was minimized with the Nadam Optimizer (Dozat 2016) and results in a learning rate of 0.0002.

The network was trained to maximize the PDF peak value for the spectroscopic redshifts (z_{spec}) of each galaxy. The spectroscopic information came from a crossmatch between DELVE DR2 and the data available in different large sky surveys (Colless et al. 2001; Mortlock, Madgwick & Lahav 2001; Wilson et al. 2006; Jones et al. 2009; Bacon et al. 2010; Drinkwater et al. 2010; Holwerda, Blyth & Baker 2011; Cooper et al. 2012; Mao et al. 2012; McLure et al. 2012; Bradshaw et al. 2013; Fèvre et al. 2013; Newman et al. 2013; Baldry et al. 2014; Treu et al. 2015; Wirth et al. 2015; Bayliss et al. 2016; Momcheva et al. 2016; Nanayakkara et al. 2016; Tasca et al. 2017; McLure et al. 2018; Scudiego, M. et al. 2018; Masters et al. 2019; Ahumada et al. 2020; Newman et al. 2020; Pharo et al. 2020; Mao et al. 2021; Mercurio et al. 2021), which resulted in approximately 4.5 million galaxies with z_{spec} measurements. We also added the z_{spec} 's available from the DECals DR9 Catalogue (Dey et al. 2019) also by doing a crossmatch with the DELVE DR2 data. All the matches were made considering a maximal separation of 0.972 arcsec.

In order to guarantee high quality photometric data used to train and test the model, we apply the following constraints on the colours, signal-to-noise ratio (SNR), and the limit of z_{spec} :

- (i) $\text{SNR} > 3$ for g

- (ii) $\text{SNR} > 5$ for riz
- (iii) $-1 < g - r < 4$
- (iv) $-1 < r - i < 4$
- (v) $-1 < i - z < 4$
- (vi) $g < 22.5$
- (vii) $0.01 < z_{\text{spec}} < 1$

The SNR cuts were used to eliminate spurious sources, bad measurements, and very faint galaxies. The g mag limitation serves to reinforce the exclusion of faint galaxies. The colour cuts were made in order to eliminate nonphysical (extremely blue and extremely red) objects (see Drlica-Wagner et al. 2018), thus the majority of the objects in our sample populate the colour-colour diagram in the regions $-0.5 \leq r - i \leq 1.5$ and $-0.5 \leq r - i \leq 0.8$. We restricted our z_{spec} interval to avoid spurious detections of low surface brightness galaxies located at high redshift. We also used the MODEST_CLASS criteria (Drlica-Wagner et al. 2018) to remove contaminant stars by choosing the objects that lie in the classes 1 (high-probably galaxy) and 3 (ambiguous classification).

To account for the lack of a band on our data, we decided to train 3 different MDN's for each different observation scenario: (1) full coverage, with optical data in $griz$ bands and partial coverage when we are missing a band—coverage only in (2) gri bands or (3) grz bands. Each MDN was trained with the magnitude (and colours) appropriate to the different scenarios. We used all of them to predict the z_{phot} 's. The objects with full coverage were assigned to the flag `model_GRIZ`, and the same was made for `gri` and `grz` coverages with the flags `model_GRI` and `model_GRZ`, respectively. For example, on the GW200202.154313 event we have approximately 3.4M objects with estimated z_{phot} 's, being ~ 62 per cent, ~ 30 per cent and ~ 8 per cent of the objects covered by $griz$, grz , gri bands, respectively.

After the selection cuts, our training sample contains about one million objects distributed at redshift $z < 1$. There are 31 252 and 6367 of these galaxies in the 90 per cent probability region of GW190924_021846 and GW200202_154313, respectively. Fig. 2 shows this final distribution, where we plotted the redshift distribution dN/dz subtracted from uniform number density (dN/dz)_{com} assuming $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ to emphasize the presence of overdensities along the line of sight.

To evaluate the performance of our MDN method, we performed a complete analysis, evaluating the point statistics and PDF's metric for the validation sample (represented by the $2.3 \times 10^5 z_{\text{spec}}$'s that have not been used for training the photo- z 's). The predicted photo- z 's as a function of the measured spectroscopic redshifts is shown in the left panel of Fig. 3. We can see that the majority of data points lie close to the diagonal, thus pointing to the accuracy of the predicted redshifts. Additionally, we can see the presence of outliers in every redshift interval. However, the outlier fraction (which is defined as $|\Delta z| > 0.15 \times (1 + z_{\text{spec}})$) results indicate that these data points only represent a minimum fraction (< 4 per cent) of the entire sample over the redshift range of interest. In order to avoid any systematic biases in DELVE galaxy distribution and their photo- z , we select three different areas with the same size of LIGO 90 per cent probability region and analyse the photo- z quality for these regions (solid lines and shadows in the right panel in Fig. 3). The right panel of Fig. 3 shows the median photo- z bias in photo- z bins of size 0.025 for DELVE and LEGACY-DR9 measurements. The results for DELVE full spectroscopic sample (dashed red line) and DELVE limited areas (dashed red line) revealed that the photo- z bias is under control at $z_{\text{phot}} = 0.5$, having median bias values smaller than 0.01 for each photo- z bins and when considering the complete sample,

³Tensorflow v2.9.1; Tensorflow Probability v0.17.0; keras-lmu v0.5.0

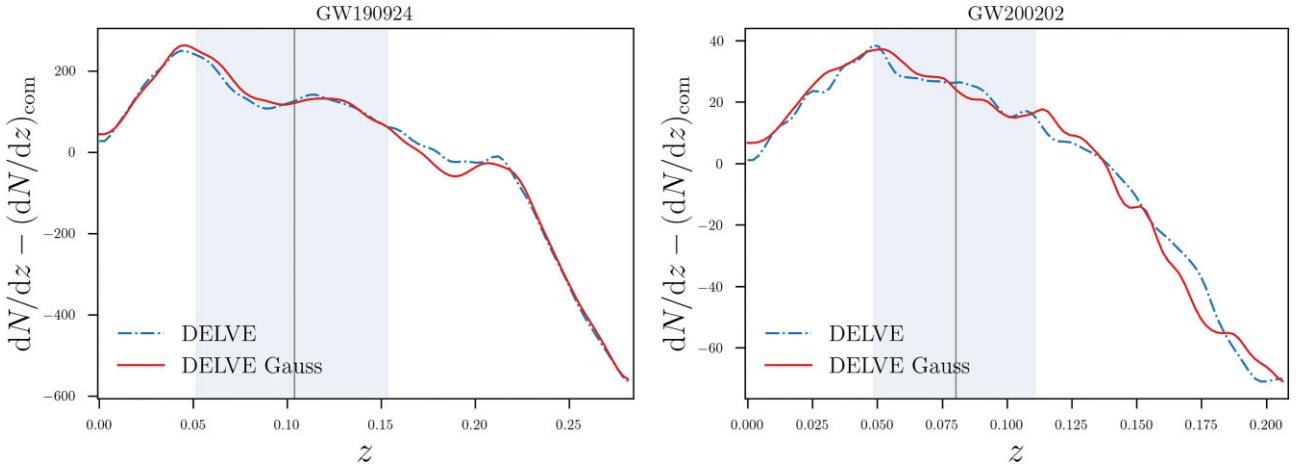


Figure 2. Redshift distribution of galaxies in the 90 per cent CI area of the dark siren events analysed in this work. The distribution is subtracted with a dN/dz with uniform number density to highlight the presence of overdensities and underdensities along the line of sight. The dashed blue line shows the distribution using the photometric redshift point estimates from the DELVE, the red line shows the same redshifts when their uncertainty is considered as a Gaussian error. The grey vertical lines represent the luminosity distance of each GW event marginalized over the entire sky, assuming an H_0 of $70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, and the shaded regions are the 1σ uncertainties considering the same H_0 ; these regions are only shown for reference.

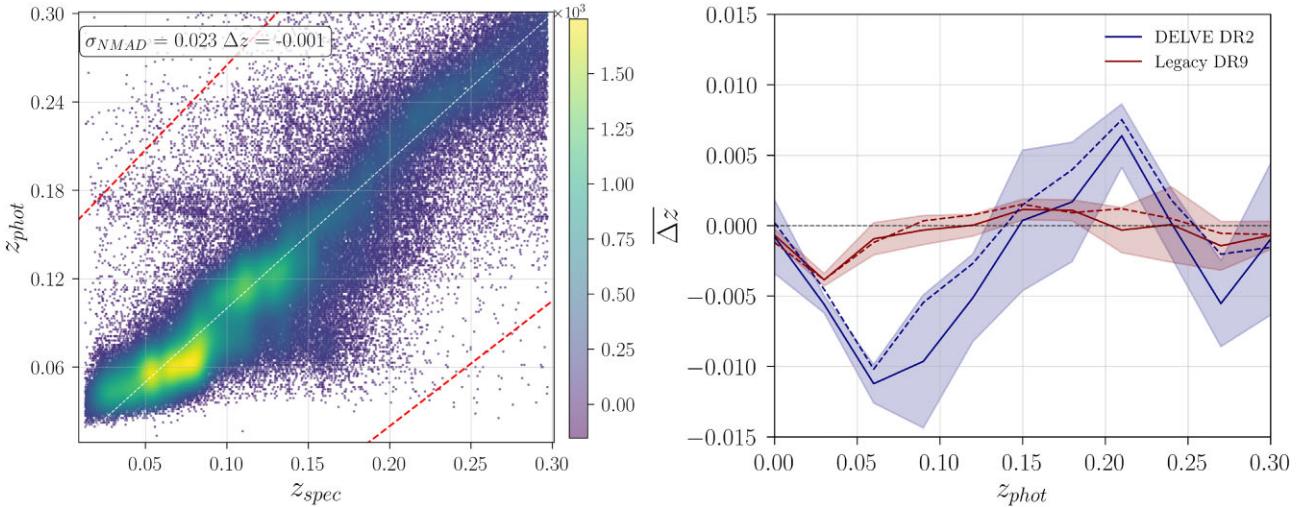


Figure 3. Photometric redshift quality assessment plots using the testing sample from the available spectroscopic data. Left: density plot of galaxies in the validation sample, showing the predicted z_{phot} 's (PDF peaks) as a function of spectroscopic redshift. The red dotted lines represent the outliers' limits, where outliers $z_{\text{phot}}^{\text{out}}$ are defined as $|z_{\text{phot}}^{\text{out}} - z_{\text{spec}}| > 0.15/(1 + z_{\text{spec}})$. Right: median value of the bias distribution $\Delta z = z_{\text{phot}} - z_{\text{spec}}$ in bins of photometric redshift for our model used in the DELVE DR2 and the Legacy Survey DR9 photometric redshifts. The median was calculated taking three different non-correlated regions in the sky covered by our test sample. The regions were chosen to have the same area as GW200202-154313. Both panels are plotted with $r < 21 \text{ mag}$, $z_{\text{spec}} < 0.3$ and $z_{\text{phot}} < 0.3$.

the value reduces to -0.001. Thus, the measurements are uniform over the DELVE footprint. In contrast, the photo-z results from the LEGACY full spectroscopic sample (blue dashed line) appear to outperform DELVE, displaying median bias values consistently below 0.005. This difference in quality could be attributed to the fact that LEGACY measurements benefit from uniform coverage across all bands and also leverage the advantages of infrared bands in their Spectral Energy Distributions (SEDS). The scatter of z_{phot} predictions was quantified with the normalized median absolute deviation, defined as $\sigma_{\text{NMAD}} = 1.48 \times \text{median}(|\Delta z|/(1 + z_{\text{spec}}))$, and the 68th percentile width of the bias distribution about the median (σ_{68}). Data for DELVE objects brighter than $r < 21$ yields $\sigma_{\text{NMAD}} = 0.023$ for all the galaxies in $0 < z_{\text{spec}} < 0.3$ and the σ_{68} is less than 0.04 for all the photo-z bins. These results are in agreement with previous

works that use similar techniques to measure photometric redshift in large sky surveys (see, Lima et al. 2022). In order to validate the individual photo-z PDF as a whole, we use two different metrics: the probability integral transform (PIT) distribution and the Odds value. The PIT distribution for DELVE data has a positive skewness, which indicates that our deep learning methods overestimate the z_{phot} . The Odds value represents the fraction of the photo-z PDF that is contained in the interval $z_{\text{spec}} \pm 0.06$, its distribution reveals that our MDN models produce narrow photo-z PDFs with their values centred near 1.

As shown previously (Soares-Santos et al. 2019; Palmese et al. 2020; Palmese et al. 2023), to overcome the fact that our sample is magnitude-limited, we have to ensure that it is volume-limited. For that, we follow the same steps used in Palmese et al. (2023). For

each GW event, we start by computing the maximum redshift after converting the higher 90 per cent CI bounds in luminosity distance into the redshift, adopting the largest value of H_0 we considered in the prior. The next step is to find an absolute magnitude threshold value that corresponds to the apparent magnitude limit at the maximum redshift. Finally, we exclude all galaxies in our sample that have an absolute magnitude above this threshold (-19.39 and -20.32 for GW200202_154313 and GW190924_021846, respectively). The galaxies that survive this cut represent the most luminous galaxies. In this work we assume that the GW hosts trace the large-scale structure similarly to these most luminous galaxies. Thus, if the GW event occurred in a galaxy below our magnitude cut, the fainter galaxies follow the same matter distribution as the most luminous galaxies. The formation channel of binary systems has an influence on the specific properties of the GW hosts and whether it is able to pass through this cut. There are studies (e.g. Rauf et al. 2023) in the literature that point to the fact that more massive galaxies are indeed more likely to host BBH mergers, in favour of the chosen cuts. However, more work on this front is needed to expand the discussion for the different formation channels and their parameter space (e.g. the assumptions that go in the binary population synthesis).

3 METHOD

In this work, we used the Bayesian formalism, described in detail in Chen, Fishbach & Holz (2018) and adapted into Soares-Santos et al. (2019) and Palmese et al. (2020, 2023), to estimate the posterior probability of H_0 for the dark siren method. The H_0 posterior for a gravitational wave measurement d_{GW} and electromagnetic data d_{EM} for a galaxy survey is written via Bayes' theorem as

$$p(H_0|d_{\text{GW}}, d_{\text{EM}}) \propto p(d_{\text{GW}}, d_{\text{EM}}|H_0) p(H_0), \quad (1)$$

where $p(H_0)$ is the prior on H_0 and $p(d_{\text{GW}}, d_{\text{EM}}|H_0)$ is the joint GW-EM likelihood. Assuming that the GW and EM measurements are independent, the joint likelihood can be written as $p(d_{\text{GW}}, d_{\text{EM}}|H_0) = p(d_{\text{GW}}|H_0)p(d_{\text{EM}}|H_0)$. By marginalizing over the true redshift z , the sky position of the GW source, the photo- z bias Δz and over all the possible galaxy hosts, the H_0 posterior can be written as in Palmese et al. (2023):

$$p(H_0|d_{\text{GW}}, d_{\text{EM}}) \propto \frac{p(H_0)}{\beta(H_0)} \sum_i \frac{1}{\mathcal{Z}_i} \int p(d_{\text{GW}}|d_L(z, H_0), \hat{\Omega}_i) \times p_i(d_{\text{EM}}|z, \Delta z) p(\Delta z) \frac{r^2(z)}{H(z)} dz d\Delta z, \quad (2)$$

where $r(z, H_0)$ is the comoving distance, $H(z) = H_0(\Omega_m(1+z)^3 + 1 - \Omega_m)^{1/2}$ is the Hubble parameter in a Flat Λ CDM model, $p(\Delta z)$ is the prior on the photometric redshift bias which is measured with the method described in the Section 2.2 (see Fig. 3), $\beta(H_0)$ is the selection function responsible for normalizing the likelihood, and \mathcal{Z} is the evidence term defined as $\mathcal{Z} = \int dz_i p(d_{\text{EM}}|z_i) r^2(z_i) / H(z_i)$. The term $p(d_{\text{GW}}|d_L(z, H_0), \hat{\Omega}_i)$ is the marginal GW likelihood computed at the solid angle $\hat{\Omega}_i$ and the redshift of the observed galaxy i , where we assume that this follow a Gaussian function according to Singer et al. (2016a). The second term in the integral represents the marginal EM likelihood of the galaxy shifted by the photo- z bias Δz , which is written as a product of galaxies photo- z PDFs computed using the deep learning algorithm described in Section 2.2. The above expression includes the assumption that the source of the GW is located in one of the galaxies present in the galaxy catalogue, making it a function of the solid angle $\hat{\Omega}_i$ and the redshift of each galaxy.

The above posterior has two important ingredients: the selection effect defined by the β function and the photo- z bias Δz . The first is associated with the selection effects adopted in the measurement process (of the electromagnetic counterparts and the detection of gravitational waves), the $\beta(H_0)$ function is computed following the same steps described in Chen, Fishbach & Holz (2018) and Palmese et al. (2023). For the electromagnetic emission selection effects, we used galaxies from the DELVE DR2 catalogue distributed up to the known absolute magnitudes for each of the GW events in the analysis, where we consider only those in $z < 0.5$. As reported in Chen, Fishbach & Holz (2018) despite this being a simplification of the real EM selection effect, since it disregards any sky accessibility, weather, and observing conditions, it is still a coherent approximation for estimating the observation of the real-time electromagnetic follow-up. On a large scale, we assume that galaxies are isotropically distributed across the sky. By marginalizing over the entire sky, the selection function can be written as

$$\beta(H_0) = \int p_{\text{sel}}^{\text{GW}}(d_L(z, H_0)) p(z) dz, \quad (3)$$

where $p(z)$ is the distribution of possible host galaxies and $p_{\text{sel}}^{\text{GW}}(d_L(z, H_0))$ is the probability of a source located at d_L being detected. This term quantifies the GW selection effect introduced by detector sensitivity and detection conditions. For the computation of $\beta(H_0)$ we follow the same steps as Palmese et al. (2023): first we simulate 70 000 BBH mergers for 20 different values of H_0 within our prior range $[20, 140] \text{ km s}^{-1} \text{ Mpc}^{-1}$. The BBH population is distributed through the redshift distribution $p(z)$ which is a function of the merger rate evolution and the cosmology-dependent comoving volume element. For simplicity, we assume that the merger rate follow the Madau-Dickinson star formation rate (Madau & Dickinson 2014). The mass of the black holes is distributed according to a power-law with index 1.6 (in agreement with the results found in Abbott et al. 2021b). We draw spins from a uniform distribution between $(-1, 1)$. The GW signals were generated using the BAYESTAR software (Singer & Price 2016; Singer et al. 2016a, b) using the frequency domain approximant IMRPhenomD. Finally, we assume the O3 sensitivity curves for LIGO and Virgo⁴, use a matched-filter analysis, and calculate the SNR of each event. We assume, as a detection condition, that the network SNR is above 12 and at least 2 detectors have a single-detector SNR above 4.

Another important effect considered in our analysis is the photo- z bias correction. When we are dealing with simulated data, the machine learning algorithm used for photometric redshift estimates can provide a biased redshift probability distribution function. The non-uniform training samples can cause systematic biases in the photo- z , causing the peak of the distribution to be shifted by Δz from the true value of z . In order to consider this effect, we use the photo- z bias computation⁵ for the DELVE DR2 catalogue (see the detailed description of this calculation in Section 2.2) in different values of z that enter on H_0 posterior through $p(d_{\text{EM}}|z, \Delta z)$.

This methodology can be extended to a sample of multiple events j with a combined data $\{d_{\text{GW},j}\}$ and d_{EM} , if we assume that the events are independent of each other and that they share the same galaxy catalogue. The Hubble constant posterior can be written as the product of the single event j likelihoods:

$$p(H_0|\{d_{\text{GW},j}\}, d_{\text{EM}}) \propto p(H_0) p(d_{\text{EM}}|H_0) \prod_j p(d_{\text{GW},j}|H_0). \quad (4)$$

⁴ Available at <https://dcc.ligo.org/LIGO-P1200087/public>

⁵ <https://datalab.noirlab.edu/delve/photoz.php>

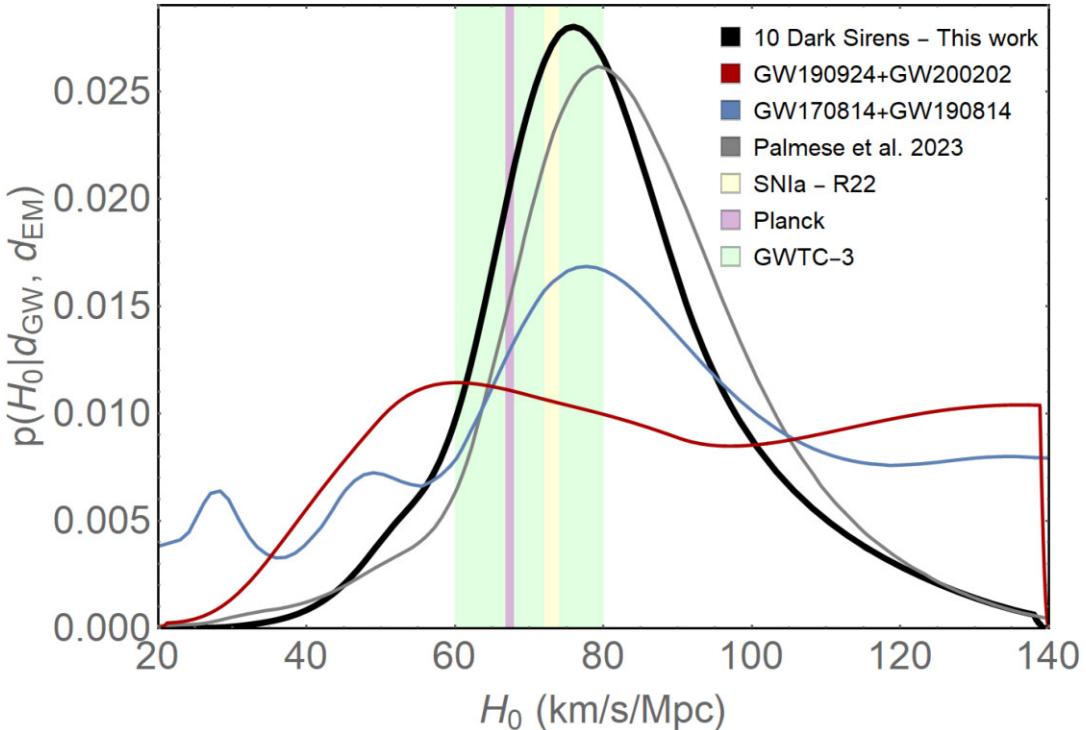


Figure 4. Hubble constant posterior distributions for the dark sirens considered in this work and previous works. The black line is the result of the combination of the two dark sirens considered here with the posteriors of all the eight dark sirens (grey line) found in Palmese et al. (2023). The combination H_0 posteriors of GW190924_021846 and GW200202_154313 is shown in red, and for comparison we also show the combination result of GW170814 and GW190814 (blue line) presented in Palmese et al. (2020). For comparison, we show the 1σ constraints on H_0 found by Planck Collaboration (2020), Riess et al. (2021) (R22), and Abbott et al. (2023) (GWTC-3) as the vertical shaded regions.

Table 2. Hubble constant measurements found with the dark sirens from the three LIGO/Virgo runs and the bright siren GW170817. All priors are flat in the range [20, 140]. The uncertainty from the flat prior is derived by assuming the same H_0 maximum found in the analysis. Quoted uncertainties represent 68 per cent HDI around the maximum of the posterior. The ‘ $\sigma_{H_0}/\sigma_{\text{prior}}$ ’ column shows the 68 per cent CI from the posterior divided by 68 per cent CI of the prior width.

Event	H_0 (km s $^{-1}$ Mpc $^{-1}$)	σ_{H_0} (km s $^{-1}$ Mpc $^{-1}$)	$\sigma_{H_0}/\sigma_{\text{prior}}$	Reference
GW190924_021846	$70.4^{+54.7}_{-15.1}$	34.9	85 %	This work
GW200202_154313	$51.2^{+61.6}_{-11.8}$	36.7	90 %	This work
GW170817—bright	$68.8^{+17.30}_{-7.63}$	12.3	30 %	Nicolaou et al. (2020)
GW190814	78^{+57}_{-13}	35	86 %	Palmese et al. (2020)
GW190924 + GW200202	$60.33^{+55.79}_{-13.61}$	34.7	85 %	This work
GW190814 + GW170814	77^{+41}_{-22}	31.5	77 %	Palmese et al. (2020)
8 dark sirens	$79.8^{+19.1}_{-12.4}$	15.8	39 %	Palmese et al. (2023)
10 dark sirens	$76.00^{+17.64}_{-13.45}$	15.55	38 %	This work

Note that the above expression does not apply to the two dark sirens studied here, since they do not share the same catalogue of galaxies because they are located in distinct regions of the sky and distance.

4 RESULTS AND DISCUSSION

We now use the DELVE photo- z ’s in the dark siren methodology to produce the H_0 posterior for GW190924_021846 and GW200202_154313. Then we combine the results for these two new GW events with those for eight dark siren events (GW170608, GW170818, GW190412, S191204r, S200129m and S200311bg, GW170814, and GW190814) from Palmese et al. (2020, 2023). The first five events were found in Palmese et al. (2023) using

the DESI Legacy Survey galaxies’ redshifts, and the last two are presented in Palmese et al. (2020) with the photo- z catalogue from DES. Fig. 4 shows the H_0 posterior from the combination of these two new dark sirens (dark red curve) and the final result (black curve) after combining the posterior of all the ten dark siren events. For comparison, we also show the results (blue curve) found in Palmese et al. (2020) with the dark sirens GW170814 and GW190814 and for the eight well-localized events (dark grey curve) found in Palmese et al. (2023). The two new events reduce the 68 per cent CI of the H_0 prior to values close to those found in Palmese et al. (2023) (see Table 2): GW190924_021846 is able to reach the value of 85 per cent and GW200202_154313 achieve the constraint of 90 per cent. The H_0 posterior distributions for GW190924_021846 and GW200202_154313 are presented in Fig. 5, we can see that both dark sirens display an evident peak at a low value of H_0

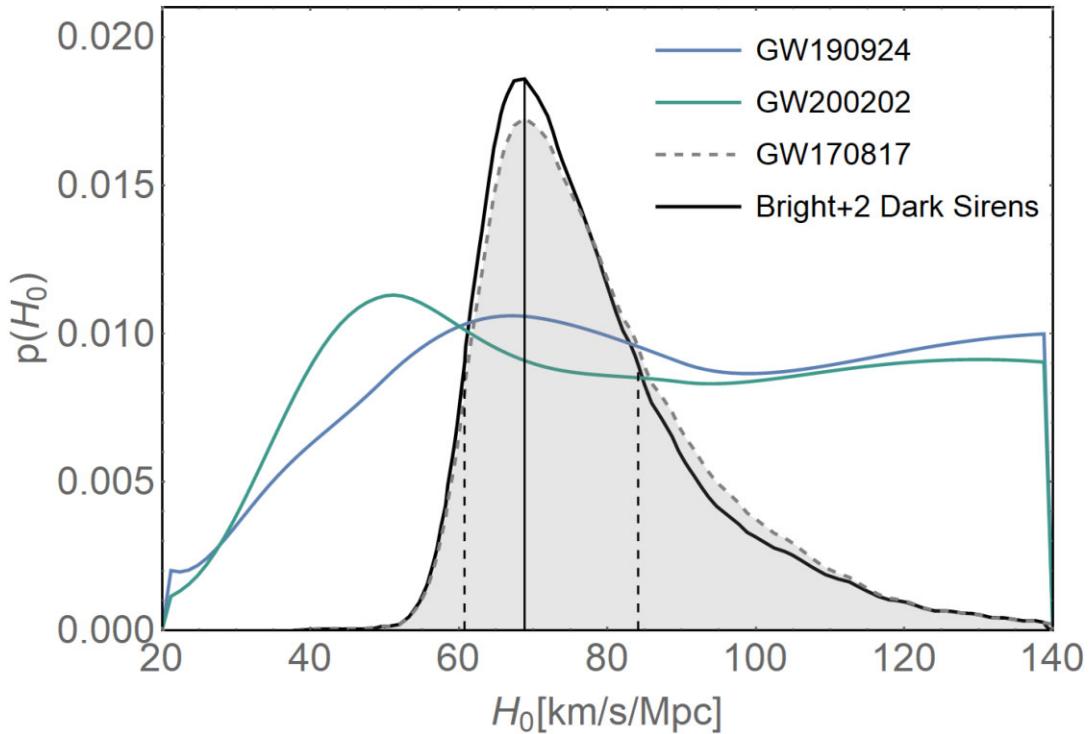


Figure 5. Hubble constant posterior distributions found using the DELVE galaxies for GW200202_154313 (green line) and GW190924_021846 (blue line). The dashed line represents the GW170817 bright siren result adapted from Nicolaou et al. (2020), which includes the peculiar velocity corrections for the galaxy host NGC 4993. The black line is the result from the combination of the two dark sirens from this work with GW170817, and the vertical dashed lines show the 68 per cent region for this posterior. Posteriors are arbitrarily rescaled only for visualization purposes.

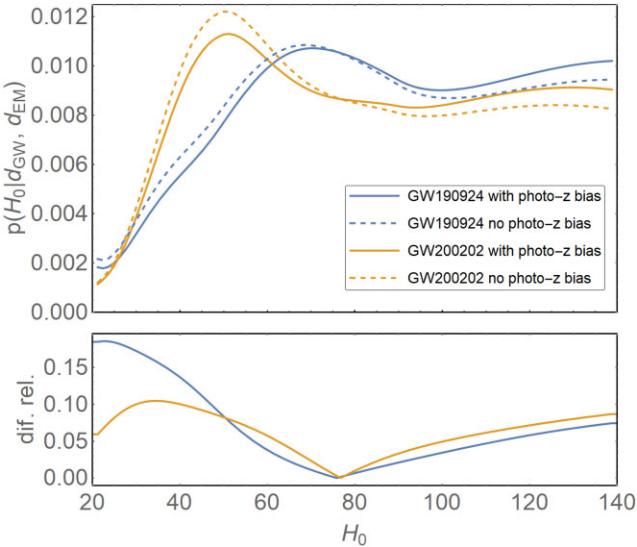


Figure 6. Photo-z bias effect on the Hubble constant posterior distributions for GW190924_021846 (blue lines) and GW200202_154313 (yellow lines). Solid curves are the results considering the inclusion of the photo-z bias correction, and dashed curves ignore this correction. The subplots present the relative difference between these two curves.

(near $70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ for GW190924_021846 and $\sim 51 \text{ km s}^{-1} \text{ Mpc}^{-1}$ for GW200202_154313) that is a consequence of the notable overdensity of galaxies (see Fig. 2) around redshift 0.05 to 0.1 and 0.05 for GW190924_021846 and GW200202_154313, respectively. As a result of better localization volume, which corresponds to a marginalization over a smaller number of galaxies, we can see that

the posterior of GW200202_154313 has a narrower peak, but the presence of a secondary peak at $H_0 \sim 114 \text{ km s}^{-1} \text{ Mpc}^{-1}$ makes it flatter than GW190924_021846 (its kurtosis value is lower, ~ 1.79 , than that produced by GW190924_021846, ~ 1.95). The analysis of the skewness showed that event GW190924_021846 produces a slightly more asymmetric posterior (a relative difference of approximately 58 per cent) than GW200202_154313. The individual posteriors shown in Fig. 5 present a high probability at the high H_0 end. The same result was observed for the H_0 posterior of GW190814 and GW170814 in Palmese et al. (2020), as explained by the authors, this is a characteristic of the dark siren method once the GW analysis only provides a d_L that is coherent with high values of H_0 . Here we decided not to adopt a wide H_0 prior, allowing to find a vast amount of galaxies distributed in redshifts that correspond to high values of H_0 . Fig. 2 reveals that the galaxies distributions for the two dark sirens are more uniformly distributed in comoving volume at high z , which results in a less informative EM likelihood and implies the recovery of the flat prior. The cut on the prior range does not bias the final result, as it only changes the redshift range considered in the dark siren analysis and not the posterior behaviour.

The combined result of all 10 dark sirens is shown in black in Fig. 4. The mode of the final posterior and the 68 per cent CI is $H_0 = 76.00^{+17.64}_{-13.45} \text{ km s}^{-1} \text{ Mpc}^{-1}$. The addition of the two new dark sirens causes a reduction of ~ 1 per cent over the 68 percent confidence region found in Palmese et al. (2023). Table 2 summarizes our results and compares the performance with other standard siren analyses.

Fig. 6 illustrates the photo-z bias effect on the H_0 posterior distribution for the events GW190924_021846 and GW200202_154313. We see that the effect is a little more significant for GW200202_154313, with a relative difference of $\sim 0.1(0.09)$ for low(high) H_0 values. For the GW190924_021846 dark siren, these values vary from 0.0002 to 0.18. This is different than what was discussed in Palmese et al.

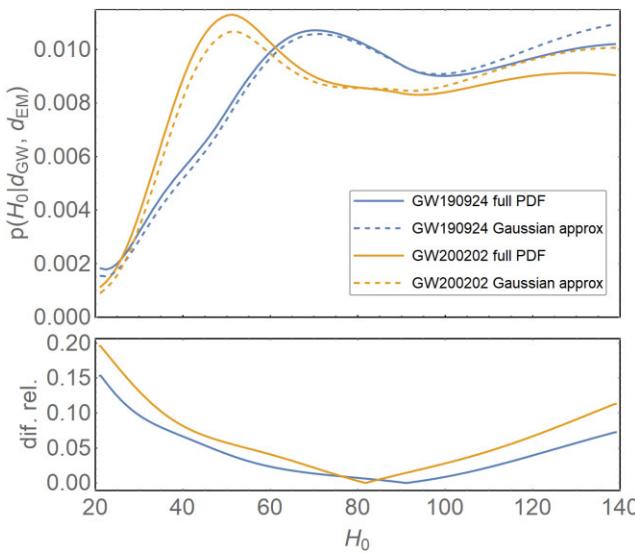


Figure 7. Comparison between the Hubble constant posterior distributions for GW190924_021846 (blue curves) and GW200202_154313 (yellow curves) obtained by using the full galaxies redshift PDFs and a Gaussian approximation. The subplots present the relative difference between the curves.

(2020), which showed that the effect of marginalization over the photo- z bias is minimum for all values of H_0 .

Another correction applied here is the full redshift PDF instead of a Gaussian approximation. The effect of this correction is shown in Fig. 7, where it is almost the same for the two dark sirens, with the relative difference varying between 0.0001 and ~ 0.2 for the entire interval of H_0 .

In order to understand the impact of dark sirens on the precision of H_0 , we combine our results with the bright siren GW170817 from Nicolaou et al. (2020). Fig. 5 presents the combined H_0 posterior. The combination of GW190924_021846, GW200202_154313, and GW170817 gives $H_0 = 68.84^{+15.51}_{-7.74} \text{ km s}^{-1} \text{ Mpc}^{-1}$. We also combine the 10 dark sirens with GW170817, it gives $H_0 = 71.54^{+10.96}_{-6.61} \text{ km s}^{-1} \text{ Mpc}^{-1}$ representing an improvement of 6 per cent in the precision of the GW170817 measurement. This result highlights the improvement obtained when well-localized GW events at redshifts well covered by galaxy catalogues are incorporated into the analysis. Our results are in agreement with the recently presented results in Abbott et al. (2023), $H_0 = 68^{+8}_{-6} \text{ km s}^{-1} \text{ Mpc}^{-1}$, that used 47 dark sirens (43 BBH, 2 BNS, and 2 NSBH) from the third LIGO–Virgo–KAGRA GW transient catalogue (Collaboration et al. 2021) with GLADE + galaxy catalogues (Dálya et al. 2018, 2022). This constraint represents a reduction of ~ 7 per cent in the 68 per cent CI of the H_0 measurement found with only GW170817. Although our measurements is less precise than Abbott et al. (2023) (as expected given the smaller number statistics), we note that we expect our result to be less sensitive to the black hole population assumptions. As noted in Abbott et al. (2023), these assumptions, and specifically the shape of the mass distribution, strongly dominate the inference on H_0 . A possible cause of this dependency is the lack of completeness of GLADE + catalogue at the redshifts of interest.

5 CONCLUSIONS

In this work, we investigate the dark siren method to constrain H_0 , and present a new measure of H_0 provided by two GW events detected by LIGO/Virgo, GW190924_021846 and GW200202_154313, with the

redshifts of the potential host galaxies derived using DELVE DR2 data. The estimation of galaxies photo- z ’s was performed using the deep learning technique Mixture Density Network. Our analyses implement the full redshift PDF of the galaxies instead of the Gaussian approximation. The main result of this study includes the measurement of the Hubble constant of $70.4^{+54.7}_{-15.1}$ and $51.2^{+61.6}_{-11.8} \text{ km s}^{-1} \text{ Mpc}^{-1}$ for GW190924_021846 and GW200202_154313, respectively, which is consistent with previous measurements of H_0 . The combination of GW190924_021846 and GW200202_154313 together with GW170817 bright siren leads to $H_0 = 68.84^{+15.51}_{-7.74} \text{ km s}^{-1} \text{ Mpc}^{-1}$, i.e. the addition of the two dark sirens reduces the 68 per cent CI interval by ~ 7 per cent, which is comparable to the ~ 12 per cent found in Palmese et al. (2020) when they add GW190814 and GW170814. This result demonstrates the power of well-localized dark siren events in better constraining the determination of the Hubble constant using deep imaging photometry obtained from surveys performing wide-sky coverage.

In addition, we also present the Hubble constant using only the dark standard siren method. We combine the H_0 posteriors found here with the posteriors of the eight well-localized dark siren events (GW170814, GW190814, GW170608, GW170818, GW190412, S191204r, S200129m, and S200311bg) presented by Palmese et al. (2020, 2023). The H_0 measurement found is $76.00^{+17.64}_{-13.45} \text{ km s}^{-1} \text{ Mpc}^{-1}$, which has a precision of 20 per cent and the 68 per cent CI interval is ~ 38 per cent of the prior width. Our result indicates that a sample with ten well-localized dark sirens and a complete galaxy catalogue can provide a significant constraint on the Hubble constant that is equivalent to that achieved with a standard siren, providing complementary information to the standard method.

Our results provide an indication of the dark siren potential as a precision cosmological probe. After a period of sensitivity upgrades, over the past few months, the LIGO/Virgo/KAGRA collaboration has returned to operation and is expected to make ~ 90 detections of mergers per year (Abbott et al. 2018). With the increase in GW observations and the arrival of deeper and wider surveys, like the forthcoming Vera C. Rubin Legacy Survey of Space and Time (LSST; Ivezić et al. 2019), it is possible that in the next few years, dark sirens will provide a measure of H_0 at the several percentage level (Del Pozzo 2012). In this regime, we highlight the need for a more robust analysis, which takes into account potential systematics neglected in the methodology adopted here. Likely, some of the most significant sources of systematics will be the galaxy catalogue selection effects, galaxy catalogue completeness, the dependence of host galaxy properties on the BBH formation channels, and the use of a Gaussian approximation for the GW likelihood instead of its full asymmetric distribution. In future work, we intend to improve the dark siren methodology in order to consider these corrections.

ACKNOWLEDGEMENTS

For the analysis, we use the PYTHON programming language, along with the following package: ASTROPY (Robitaille et al. 2013; Price-Whelan et al. 2018), MATPLOTLIB (Hunter 2007), NUMPY (Harris et al. 2020), SCIPY (Virtanen et al. 2020), TENSORFLOW (Abadi et al. 2015), and LIGO.SKYMAP (Singer & Price 2016; Singer et al. 2016a, b; Kasliwal et al. 2017).

GT and CB acknowledge the financial support from CNPq (402577/2022-1). CB acknowledges the financial support from CNPq (316072/2021-4), FAPERJ (grants 201.456/2022 and 210.330/2022), and the FINEP contract 01.22.0505.00 (ref. 1891/22). The authors made use of Sci-Mind servers machines developed by the CBPF AI LAB team and would like to thank P. Russano and M. Portes de Albu-

querque for all the support in infrastructure matters. CEM-V is supported by the International Gemini Observatory, a program of NSF's NOIRLab, which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation, on behalf of the Gemini partnership of Argentina, Brazil, Canada, Chile, the Republic of Korea, and the United States of America. GEM acknowledges support from the University of Toronto Arts & Science Postdoctoral Fellowship program. Time-domain research by DJS is supported by NSF grants AST-1821987, 1813466, 1908972, 2108032, and 2308181 and by the Heising-Simons Foundation under grant #2020-1864. JAC-B acknowledges support from FONDECYT Regular N 1220083.

The DELVE project is partially supported by the NASA Fermi Guest Investigator Program Cycle 9 No. 91201 and NSF awards AST-2108168 and AST-2307126. This work is partially supported by Fermilab LDRD project L2019-011.

This project used data obtained with the Dark Energy Camera (DECam), which was constructed by the Dark Energy Survey (DES) collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo a Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Cientifico e Tecnologico and the Ministerio da Ciencia, Tecnologia e Inovacao, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenossische Technische Hochschule (ETH) Zurich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciencias de l'Espai (IEEC/CSIC), the Institut de Fisica d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universitat Munchen and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

Based on observations at Cerro Tololo Inter-American Observatory, NSF's NOIRLab (NOIRLab Prop. ID 2019A-0305; PI: Alex Drlica-Wagner), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

BASS is a key project of the Telescope Access Program (TAP), which has been funded by the National Astronomical Observatories of China, the Chinese Academy of Sciences (the Strategic Priority Research Program 'The Emergence of Cosmological Structures' Grant # XDB09000000), and the Special Fund for Astronomy from the Ministry of Finance. The BASS is also supported by the External Cooperation Program of Chinese Academy of Sciences (Grant #

114A11KYSB20160057), and Chinese National Natural Science Foundation (Grant # 11433005).

This manuscript has been authored by Fermi Research Alliance, LLC, under contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Abadi M. et al., 2015, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, <https://www.tensorflow.org/>

Abbott B. P. et al., 2017a, *Class. Quantum Gravity*, 34, 18

Abbott B. P. et al., 2017b, *Phys. Rev. Lett.*, 119, 161101

Abbott B. P. et al., 2017c, *Phys. Rev. Lett.*, 119, 141101

Abbott B. et al., 2017d, *Nature*, 551, 85

Abbott B. et al., 2018, *Living Rev. Relativ.*, 21, 3

Abbott B. et al., 2019, *Phys. Rev. X*, 9, 031040

Abbott R. et al., 2020a, *Phys. Rev. D*, 102, 043015

Abbott R. et al., 2020b, *ApJ*, 896, L44

Abbott R. et al., 2021a, *Phys. Rev. X*, 11, 021053

Abbott R. et al., 2021b, *ApJ*, 913, L7

Abbott R. et al., 2023, *ApJ*, 949, 37

Ackley K. et al., 2020, *A&A*, 643, 48

Ahumada R. et al., 2020, *ApJS*, 249, 3

Alexander K. D. et al., 2021, *ApJ*, 492, 13

Amaro-Seoane P. et al., 2017, Laser Interferometer Space Antenna , preprint (arXiv:1702.00786)

Andreoni I. et al., 2020, *ApJ*, 884, 16

Arcavi I. et al., 2017, *Nature*, 551, 64

Bacon R. et al., 2010, in McLean I. S., Ramsay S. K., Takami H.eds, SPIE Proceedings, SPIE,Bellingham, Washington, 773508

Baldry I. K. et al., 2014, *MNRAS*, 441, 2440

Bayliss M. B. et al., 2016, *ApJS*, 227, 3

Bishop C., 1994, Aston University, p. 26

Bom C. R., Palmese A., 2023, preprint (arXiv:2307.01330)

Bom C. R. et al., 2023, preprint (arXiv:2302.04878)

Bradshaw E. J. et al., 2013, *MNRAS*, 433, 194

Chen H.-Y., Fishbach M., Holz D. E., 2018, *Nature*, 562, 545

Chornock R. et al., 2017, *ApJ*, 848, L19

Collaboration T. L. S., the Virgo Collaboration, the KAGRA Collaboration, 2023, GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run, *Phys. Rev. X*, 13, 041039

Colless M. et al., 2001, *MNRAS*, 328, 1039

Cooper M. C. et al., 2012, *MNRAS*, 425, 2116

Coulter D. A. et al., 2017, *Science*, 358, 1556

Cowperthwaite P. S. et al., 2017, *ApJ*, 848, L17

Dálya G. et al., 2018, *MNRAS*, 479, 2374

Dálya G. et al., 2022, *MNRAS*, 514, 1403

Del Pozzo W., 2012, *Phys. Rev. D*, 86, 043011

Dey A., Schlegel D. J., Lang D. et al., 2019, *AJ*, 157, 168

Di Valentino E., Mena O., Pan S. et al., 2021, *Class. Quantum Grav.*, 38, 153001

Dobie D., Stewart A., Murphy T. et al., 2019, *ApJ*, 887, L8

Dozat T., 2016, in Proceedings of the 4th International Conference on Learning Representations (ICLR Workshop). p. 2013–2016, <https://openreview.net/forum?id=OM0jvwB8Jp57ZJjtNEZ>

Drinkwater M. J., Jurek R. J., Blake C. et al., 2010, *MNRAS*, 401, 1429

Drlica-Wagner A., Sevilla-Noarbe I., Rykoff E. S. et al., 2018, *ApJS*, 235, 33

Drlica-Wagner A., Carlin J. L., Nidever D. L. et al., 2021, *ApJ*, 256, 22

Drlica-Wagner A., Ferguson P. S., Adamów M. et al., 2022, *ApJ*, 261, 18

Evans P. A., Cenko S. B., Kennea J. A. et al., 2017, *Science*, 358, 1565

Fèvre O. L., Cassata P., Cucciati O. et al., 2013, *A&A*, 559, A14

Fishbach M., Gray R., Magaña Hernandez I. et al., 2019, *ApJ*, 871, L13

Flaugh B., Diehl H. T., Honscheid K. et al., 2015, *AJ*, 150, 150

Freedman W. L., Madore B. F., Hatt D. et al. 2019, *ApJ*, 882, 34

Gair J. R. et al., 2023, *AJ*, 166, 22

Goldstein A., Veres P., Burns E. et al. 2017, *ApJ*, 848, 14

Gomez S., Hosseinzadeh G., Cowperthwaite P. S., Villar V. A. et al. 2019, *ApJ*, 884, L8

Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, *ApJ*, 622, 759

Harris C. R., Millman K. J., van der Walt S. J. et al., 2020, *Nature*, 585, 357

Holwerda B. W., Blyth S.-L., Baker A. J., 2011, *Proc. Int. Astron. Union*, 7, 496

Hunter J. D., 2007, *Comput Sci Eng*, 9, 90

Ivezic Z., Kahn S. M., Tyson J. A. et al., 2019, *ApJ*, 873, 111

Jones D. H., Read M. A., Saunders W. et al. 2009, *MNRAS*, 399, 683

Kasliwal M. M., Nakar E., Singer L. P. et al., 2017, *Science*, 358, 1559

Kilpatrick C. D., Coulter D. A., Arcavi I. et al., 2021, *ApJ*, 923, 26

LIGO Scientific Collaboration and Virgo Collaboration, 2019, GRB Coordinates Network, 26334, 1

LIGO Scientific Collaboration and Virgo Collaboration, 2020a, GRB Coordinates Network, 26926, 1

LIGO Scientific Collaboration and Virgo Collaboration, 2020b, GRB Coordinates Network, 27382, 1

Lima E., Sodré L., Bom C. et al., 2022, *Astron. Comput.*, 38, 100510

Madau P., Dickinson M., 2014, *ARA&A*, 52, 71

Mao M. Y., Sharp R., Norris R. P. et al., 2012, *MNRAS*, 426, 3334

Mao Y.-Y., Geha M., Wechsler R. H. et al., 2021, *ApJ*, 907, 85

Masters D. C., Stern D. K., Cohen J. G. et al., 2019, *ApJ*, 877, 81

McLure R. J., Pearce H. J., Dunlop J. S. et al., 2012, *MNRAS*, 428, 1088

McLure R. J., Pentericci L., Cimatti A. et al., 2018, *MNRAS*, 479, 25

Mercurio A., Rosati P., Biviano A. et al., 2021, *A&A*, 656, A147

Momcheva I. G., Brammer G. B., van Dokkum P. G. et al., 2016, *ApJS*, 225, 27

Mortlock D. J., Madgwick D. S., Lahav O., 2001, *PASA*, 18, 192

Mucesh S., Hartley W. G., Palmese A. et al., 2021, *MNRAS*, 502, 2770

Muttoni N., Laghi D., Tamanini N., Marsat S., Izquierdo-Villalba D., 2023, *Phys. Rev. D*, 108, 043543

Nanayakkara T., Glazebrook K., Kacprzak G. G. et al., 2016, *ApJ*, 828, 21

Newman J. A., Cooper M. C., Davis M. et al., 2013, *ApJS*, 208, 5

Newman A. B., Rudie G. C., Blanc G. A. et al., 2020, *ApJ*, 891, 147

Nicholl M., Berger E., Kasen D. et al., 2017, *ApJ*, 848, L18

Nicolaou C., Lahav O., Lemos P., Hartley W., Braden J., 2020, *MNRAS*, 495, 90

Palmese A., deVicente J., Pereira M. E. S. et al., 2020, *ApJ*, 900, L33

Palmese A., Bom C. R., Mucesh S., Hartley W. G., 2023, *ApJ*, 946, 12

Pharo J., Malhotra S., Rhoads J. E. et al., 2020, *ApJ*, 888, 79

Pian E., Avanzo D. P., Benetti S. et al., 2017, *Nature*, 551, 67

Planck Collaboration, 2020, *A&A*, 641, 67

Price-Whelan A. M., Sspócz B. M., Günther H. M. et al., 2018, *AJ*, 156, 123

Rauf L., Howlett C., Davis T. M. et al., 2023, *MNRAS*, 523, 5719

Riess A. G., Yuan W., Macri L. M. et al., 2021, *ApJ*, 934, L7

Robitaille T. P., Tollerud E. J., Greenfield P. et al., 2013, *A&A*, 558, A33

Rodríguez-Ramírez J., Bom C., Fraga B., Nemmen R., 2023, *MNRAS*, 527, 6076

Sathyaprakash B., Abernathy M., Acernese F. et al., 2012, *Class. Quantum Gravity*, 29, 18

Savchenko V., Ferrigno C., Kuulkers E. et al., 2017, *ApJ*, 848, L15

Schutz B. F., 1986, *Nature*, 323, 310

Scodellaggio M., Guzzo L., Garilli B. et al., 2018, *A&A*, 609, A84

Singer L. P., Price L. R., 2016, *Phys. Rev. D*, 93, 024013

Singer L. P., Chen H.-Y., Holz D. E. et al., 2016a, *ApJ*, 226, L10

Singer L. P., Chen H.-Y., Holz D. E. et al., 2016b, *ApJ*, 829, L15

Smartt S. J., Chen T.-W., Jerkstrand A. et al., 2017, *Nature*, 551, 75

Soares-Santos M., Holz D. E., Annis J. et al., 2017, *ApJ*, 848, L7

Soares-Santos M., Palmese A., Hartley W. et al., 2019, *ApJ*, 876, L7

Tanvir N. R., Levan A. J., Gonzalez-Fernandez C. et al., 2017, *ApJ*, 848, L27

Tasca L. A. M., Fèvre O. L., Ribeiro B. et al., 2017, *A&A*, 600, A110

Treu T., Schmidt K. B., Brammer G. B. et al., 2015, *ApJ*, 812, 114

Tucker D. L., Wiesner M. P., Allam S. S. et al., 2021, *ApJ*, 929, 26

Valenti S., Sand D. J., Yang S. et al. 2017, *ApJ*, 848, L24

Verde L., Treu T., Riess A. G., 2019, *Nat. Astron.*, 3, 891

Vieira N., Ruan J. J., Haggard D. et al., 2020, *ApJ*, 895, 18

Virtanen P., Gommers R., Oliphant T. E. et al., 2020, *Nat. Methods*, 17, 261

Voelker A. R., Kajic I., Eliasmith C., 2019a, in *Advances in Neural Information Processing Systems*. p. 15544, Curran Associates, Inc., Red Hook, New York, United States

Voelker A., Kajic I., Eliasmith C., 2019b, in *Advances in Neural Information Processing Systems*. Curran Associates, Inc., Red Hook, New York, United States, p. 29, <https://proceedings.neurips.cc/paper/2019/file/9522-85b9b7e7a1be5aa7849f32ffff05-Paper.pdf>

Watson A. M., Butler N. R., Lee W. H. et al., 2020, *MNRAS*, 492, 5916

de Wet S., Groot P. J., Bloemen S. et al., 2021, *A&A*, 649, 7

Wilson G., Muzzin A., Lacy M. et al., 2006, Clusters of Galaxies at $1 < z < 2$: The Spitzer Adaptation of the Red-Sequence Cluster Survey , preprint(arXiv:0604289)

Wirth G. D., Trumpp J. R., Barro G. et al., 2015, *ApJ*, 150, 153

Zhou R., Newman J. A., Mao Y.-Y. et al., 2020, *MNRAS*, 501, 3309

Zuntz J. et al., 2021, The LSST-DESC 3x2pt Tomography Optimization Challenge , preprint (arXiv:2108.13418)

¹Centro Brasileiro de Pesquisas Físicas, Rua Dr Xavier Sigaud 150, 22290-180 Rio de Janeiro, RJ, Brazil

²Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Rodovia Márcio Covas, lote J2, quadra J, 23812-101 Itaguaí, Brazil

³McWilliams Center for Cosmology, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA

⁴Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510, USA

⁵Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA

⁶Department of Astronomy & Astrophysics, University of Chicago, 5640 S Ellis Avenue, Chicago, IL 60637, USA

⁷Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

⁸Gemini Observatory/NSF's NOIRLab, 670 N. A'ohoku Place, Hilo, HI 96720, USA

⁹Department of Astronomy/Steward Observatory, 933 North Cherry Avenue, Room N204, Tucson, AZ 85721-0065, USA

¹⁰Center for Astrophysics and Space Astronomy, University of Colorado, 389 UCB, Boulder, CO 80309-0389, USA

¹¹Department of Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON M5S 3H4, Canada

¹²Instituto de Alta Investigación, Universidad de Tarapacá, Casilla 7D, Arica 1020000, Chile

¹³NSF's NOIRLab, 950 N. Cherry Ave, Tucson, AZ 85719, USA

¹⁴Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA

¹⁵IAG, Departamento de Astronomia, Universidade de São Paulo, SP 05508-090 São Paulo, Brazil

¹⁶Department of Physics and Astronomy, Dartmouth College, Hanover, NH 03755, USA

¹⁷Department of Physics, University of Surrey, Guildford GU2 7XH, UK

¹⁸Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

¹⁹Center for Astrophysical Sciences, Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21217, USA

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.