Cosmological inference using gravitational waves and normalizing flows

Federico Stachurski,* Christopher Messenger, and Martin Hendry
SUPA, School of Physics and Astronomy, University of Glasgow, Glasgow, United Kingdom

(Received 28 October 2023; revised 16 February 2024; accepted 3 June 2024; published 27 June 2024)

We present a machine learning approach using normalizing flows for inferring cosmological parameters from gravitational wave events. Our methodology is general to any type of compact binary coalescence event and cosmological model and relies on the generation of training data representing distributions of gravitational wave event parameters. These parameters are conditional on the underlying cosmology and incorporate prior information from galaxy catalogues. We provide an example analysis inferring the Hubble constant using binary black holes detected during the O1, O2, and O3 observational runs conducted by the advanced LIGO/VIRGO gravitational wave detectors. We obtain a Bayesian posterior on the Hubble constant from which we derive an estimate and 1σ confidence bounds of $H_0 = 74.51^{+14.80}_{-13.63}\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$. We are able to compute this result in $\mathcal{O}(1)$ s using our trained normalizing flow model.

DOI: 10.1103/PhysRevD.109.123547

I. INTRODUCTION

The idea of using gravitational wave (GW) detections to gain insights into the cosmological properties of our Universe was first proposed by Schutz [1]. These "standard sirens" give us information about the calibrated luminosity distance of the event without the use of the cosmological distance ladder or any prior knowledge or assumptions of the Universe. By adding the information of redshifts from galaxy catalogs and using the relationship between luminosity distance and redshift [2], one can infer cosmological parameters.

At present, the measurements of H_0 are in tension with each other, of approximately 4.4σ [3]. The Planck experiment [4] estimated the Hubble constant to be $H_0 = 67.4 \pm 0.5\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ (1σ confidence interval), using measurements from the cosmic microwave background radiation. The SH0ES experiment [5] measured $H_0 = 73.04 \pm 1.04\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ (1σ confidence interval), making use of the measured distances of type 1a supernova standard candles. This suggests that either one of the experiments might be subject to unknown systematic errors or maybe an indication of some "new" underlying physics causing this discrepancy [3]. Therefore, standard siren measurements of H_0 using GWs offers a promising avenue for addressing the existing discrepancies observed in the measurements of the Hubble constant.

Since the LIGO and Virgo detectors [6–8] were activated, they have revealed a significant number of GWs events—totaling 90 thus far. These events span a range of phenomena, from binary black hole (BBH) mergers to neutron star black holes, as well as binary neutron star coalescences [9,10].

The most recent results of the Hubble constant estimation using GWs are from [11], where two independent analysis methods were employed. In the first, `gwcosmo` [12], they provide an estimated value of $H_0 = 68^{+12}_{-8}\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ (1σ highest density interval) when combined with the H_0 measurement from GW170817 and its electromagnetic radiation counterpart. In the second analysis [13], no galaxy catalog prior information was used. We also note Hubble value estimates of $H_0 = 68^{+26.0}_{-6.2}\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ (1σ equal-tailed interval) obtained using spatial cross-correlation between GW sources and the photometric galaxy surveys [14–16].

In this work, we see how our new analysis, `CosmoFlow`, using normalizing flows (NFs), a machine learning driven process, allows us to define expressive probability distributions [17], over cosmological parameters using GW posterior samples as inputs. NFs are widely used in the GW community, from performing fast and reliable GW parameter estimation [18–20] to population studies [21]. Other relevant studies, such as those reported by [22], have used NFs to calculate the local covariance at any point within the parameter space. This calculation allows the inverse, known as the Fisher matrix, to be used as a local metric. Employing these techniques, the study analyzed parameter posteriors from the dark energy survey and Planck Cosmic microwave background (CMB) lensing, where they have derived an estimate of the Hubble constant of $H_0 = 73.8 \pm 7.5\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ (1σ highest density interval).

*Contact author: f.stachurski.1@research.gla.ac.uk

In this work, we demonstrate that a NF can model the electromagnetic (EM) dependent prior of the GW parameters. We can then account for selection effects and uncertainties in the detected GW parameters to evaluate the likelihood of a cosmological parameter set. We train a NF model prior to the detection of GW events, and for each event, we input parameter estimation results and compute the cosmological parameter likelihood. The time taken by the NF is ∼1 ms per set of cosmological parameters. This allows us to compare our analysis to that of `gwcosmo` [11] and compute an overall combined posterior distribution over the Hubble constant with the 42 BBHs observed during the O1, O2, and O3 observational runs in the order of $\mathcal{O}(1)$ s.

II. BAYESIAN FRAMEWORK

We define the posterior distribution of the cosmological parameters, Ω, conditioned by the GW strain data of an ensemble of n events, $\mathbf{h} = [h_1, h_2, \ldots, h_n]$ as

$$p(\Omega|\mathbf{h}, \mathbf{D}, I) = p(\Omega|I) \prod_i \frac{p(h_i, D_i|\Omega, I)}{p(h_i, D_i|I)}, \quad (1)$$

where $\mathbf{D} = [D_1, D_2, \ldots, D_n]$ is the binary state of detection, 1 for detected and 0 for not detected for the ith event, and I represents all other assumed information. The detectability of a GW signal serves to account for the selection effects that arise when applying signal-to-noise ratio (SNR) thresholds on candidate events, and we must make sure that the likelihood term is properly normalized, such that $\sum_{D_i=0,1} \int p(h_i, D_i|\Omega, I) dh_i = 1$. However, since encoded within I is the information that only events exhibiting an SNR greater than some threshold $\rho_{\rm th}$ are considered, we find that

$$p(h_i, D_i|\Omega, I) = \frac{p(h_i|\Omega, I)}{p(D_i|\Omega, I)}, \quad (2)$$

since by definition, $p(D_i = 0|\Omega, I) = 0$ for all detected events. This result allows us to write our cosmological parameter posterior as

$$p(\Omega|\mathbf{h}, \mathbf{D}, I) = p(\Omega|I) \prod_i \frac{\int p(h_i|\theta_i, I) p(\theta_i|\Omega, I) d\theta_i}{p(h_i, D_i|I) p(D_i|\Omega, I)}, \quad (3)$$

where we have marginalized numerator of Eq. (2) over the GW parameters, θ.

It becomes apparent when we discuss the generation of training data for our NF approach that it is more practical to deal with a GW parameter prior that is conditional on detection. Hence, via Bayes' theorem we obtain

$$p(\theta_i|\Omega, D_i, I) = \frac{p(D_i|\theta_i, I) p(\theta_i|\Omega, I)}{p(D_i|\Omega, I)}, \quad (4)$$

where we have used the fact that the detectability of an event is independent of Ω if the GW parameters are given.

With some rearrangement and using Eq. (4) and replacing the GW likelihood with the ratio of its posterior and prior, we are able to write our cosmological parameter posterior as

$$p(\Omega|\mathbf{h}, \mathbf{D}, I) \propto p(\Omega|I) \prod_i \int \frac{p(\theta_i|h_i, I) p(\theta_i|D_i, \Omega, I)}{p(D_i|\theta_i, I) p(\theta_i|\Omega_0, I)} d\theta_i, \quad (5)$$

where $p(\theta|\Omega_0, I)$ represents the GW parameter priors used in the parameter estimation assuming a fixed cosmology Ω_0 [10].

We can then approximate the integral over the GW parameters as a Monte Carlo summation giving us

$$p(\Omega|\mathbf{h}, \mathbf{D}, I) \propto p(\Omega|I) \prod_i \left\langle \frac{p(\theta_i|D_i, \Omega, I)}{p(D_i|\theta_i, I) p(\theta_i|\Omega_0, I)} \right\rangle_{\theta_i \sim p(\theta_i|h_i, I)}, \quad (6)$$

as our final result.

The numerator within Eq. (6) represents the GW parameter priors conditional on detection and the cosmological parameters. We use a NF to model this prior and incorporate galaxy catalogue information within the training procedure. For the two terms in the denominator, we use the survival function of a noncentral χ-squared distribution with a noncentral parameter the SNR squared and d.o.f. $k = 2n$, where n is the number of detectors used for the detection of the event, to compute $p(D_i|\theta, I)$; instead, the term $p(\theta|\Omega_0, I)$ is the prior on the GW parameters used in the parameter estimations process conditioned by a fixed cosmology [10].

III. NORMALIZING FLOWS

NFs have the capacity to efficiently evaluate and sample from complex probability distribution functions. They operate by transforming a simple data distribution, such as a multivariate normal distribution, through a series of affine transformations, ultimately generating a more intricate output distribution [17,23]. Assuming x is a random variable sampled from a distribution $p_x(x|\omega)$ and that another random variable y sampled from distribution $y \sim p_y(y|\omega)$ is related with x following the relation $y = g(x)$ and $x = f(y)$, where ω is a conditional statement, then a flow can be constructed with a conditional statement [24], which ultimately allows for the computation of likelihood-like terms, such as the one in Eq. (6) [25]. Therefore, using the change of variable equation becomes

$$\log\left(p_{\mathbf{x}}(\mathbf{x}|\omega)\right) = \log\left(p_{\mathbf{y}}(f^{-1}(\mathbf{x},\theta|\omega))\right) + \log\left(\det\left|\frac{df^{-1}(\mathbf{x},\theta|\omega)}{d\mathbf{x}}\right|\right), \quad (7)$$

where θ are trainable parameters for the function f. The θ parameters are optimized with a loss function, using the Kullback–Leibler divergence between the left-hand side and the right-hand side of Eq. (7).

Therefore, a set of nonlinear functions can be composed together to construct more complicated functions, allowing to go from a simple distribution to a more complicated data distribution [17]. Using Eq. (7), it is possible to construct a flow to compute the numerator term in Eq. (6), where x are the GW parameters and $\omega = [H_0, D_i]$.

IV. DATA GENERATION

In order to train our NF, we must provide it with training data representing the GW parameter distribution we wish to model. We must also provide the corresponding cosmological parameters upon which that distribution is conditioned. The data generation process and results presented in this paper have been limited to the case where inference is performed only on the Hubble constant. We have restricted the analysis for the sole purpose of direct validation of our results with those presented in [11].

We begin by sampling cosmological values Ω, where in this case $\Omega \equiv H_0$, which serves as the conditional input for the NF. We consider a specific cosmology characterized by a FlatΛCDM model with fixed cosmological parameters $\Omega_m = 0.3$ and $w_0 = -1$, as used in [11,26] and vary the expansion rate parameter H_0 uniformly within the range $[20, 140]$ km s^{-1} Mps^{-1}.

We use the sampled values of H_0 to compute the Schechter functions $\phi(L|\Omega)$ from which we sample the luminosities of host galaxies. The Schechter functions are those described in the GLADE+ catalog [27], using the K-band fitting parameters given in [11]. We perform k-corrections and color-evolution corrections [11] and introduce luminosity weighting, favoring galaxies in proportion to their luminosities as hosting GW events [26], hence $p(L|\Omega) \propto L\phi(L|\Omega)$.

We next sample the redshifts of galaxies hosting BBH events from a prior uniform in comoving volume and comoving time, multiplied by the merger rate component, described in Appendix A.2 in [11], where the fitting parameters from [10] were used [28]. The redshifts are then combined with their corresponding cosmological parameters to obtain the luminosity distances of galaxies hosting GW events. With the sampled galaxy luminosity, we can then compute the apparent magnitude of the galaxy.

At this point, we sample uniformly over the two-spheres to generate sky locations. To incorporate redshift and sky location information from the local Universe into our training data, we use the GLADE+ [27] catalog. To select galaxies from the catalog, we compare the apparent magnitude and location of the sampled galaxy with the magnitude threshold map of the sky [29]; we determine if such a galaxy would have been contained within the catalog for the corresponding sky pixel location. If false, we retain the sampled galaxy attributes and record that the host is outwith the catalog. However, if true, we substitute the previously sampled galaxy with a randomly sampled galaxy (weighted by $(1+z)^{-1}$ and in proportion with luminosity) from that pixel in the catalog. The sample will retain its sampled cosmological parameter, but apparent magnitude and sky location will be taken from the selected catalog galaxy. The catalog redshift is interpreted as an uncertain measurement of the redshift which is generated by sampling from a Gaussian distribution centred on the measured value with standard deviation taken from the catalog redshift uncertainty. This redshift is combined with the cosmological parameter to compute the luminosity distance.

With host galaxy parameters selected, we proceed to sample the GW parameters of the event. The parameters we sample are m_1 and m_2, primary and secondary masses, using a power law plus peak model for m_1 and a power law for m_2 conditioned on m_1 [10,11]; the values for the hyperparameters describing the power law plus peak model are as follows: $\alpha = 3.78$, $\beta = 0.81$, $m_{\mathrm{max}} = 112.5M_\odot$, $m_{\mathrm{min}} = 4.98M_\odot$, $\delta_m = 4.8M_\odot$, $\mu_g = 32.27M_\odot$, $\sigma_g = 3.88M_\odot$, $\lambda_g = 0.03$; a_1 and a_2, primary and secondary spins, from a uniform distribution between [0, 0.99] for both parameters [9]; θ_1, θ_2, and θ_{JN}, primary and secondary axis orientation and the orbital plane inclination angle, respectively, from uniform distributions between $[-1, 1]$ in $\cos(\theta)$ [9]; ϕ_{JL}, ϕ_{12} and ψ, from a uniform distributions between $[0, 2\pi]$ [9]. Lastly, we sample geocentric time of arrival from a uniform distribution over one sidereal day, [0.0000, 86164.0905] s. This comprises our set of 11 intrinsic GW parameters (omitting a reference phase) plus the extrinsic sky position and luminosity distance parameters shared with the host galaxy.

After sampling GW parameters, we employ the Bilby package [30] to simulate event-specific SNR to determine whether an event would be detected. We produce datasets that correspond to various detector configurations, considering the relevant power spectral densities employed during the O1, O2, O3a, and O3b observational periods. Each data set is then used to train an individual flow for that specific detector setup and observational period.

Under the assumption of well behaved Gaussian detector noise, we draw samples of matched-filter SNR from a noncentral χ distribution with a noncentrality parameter equal to the optimal network SNR and with $2n$ degrees of freedom where n is the number of detectors in the network. If the sampled SNR is $\geq$ than the SNR threshold ($\rho_{\mathrm{th}} = 11$ [11]), then the event is retained. Otherwise, the sample is discarded, and we begin the sampling procedure again from the start but retaining the original sample of the cosmological parameters. This latter choice ensures that all sampled parameters will be conditional upon detection with the exception of our cosmological parameters.

We note that a direct computation of the optimal SNR becomes a significant data generation bottleneck due to one-by-one parameter processing. By setting an SNR threshold, many generated events are discarded, further slowing the process. Addressing this, we have developed and applied a multilayer perceptron (MLP) neural network model, with a structure of 8 layers and 128 neurons per layer, to act as an accurate and efficient function approximator for the SNR computation. Training the MLP using 13 GW parameters as input (omitting phase) and training with a targeted output of SNR multiplied by luminosity distance, we enable efficient SNR prediction for vectorized parameter sets, enhancing data generation speed by a factor of 20000. This MLP implementation uses the `poplar` module [31] with the widely used PyTorch package.

V. TRAINING THE NORMALIZING FLOW

The data used to train a NF model for the numerator in Eq. (6) is a combination of 10^6 samples for each observing run and detector configuration. The model considers only the luminosity distance, primary and secondary masses, and sky location, α and δ, with the conditional parameter $\omega = H_0$. The remaining nine parameters are marginalized, aligning with the analysis of [11], which includes only GW parameters influenced by specific cosmology or EM catalog information.

This study uses a CouplingNF model, which uses parametrized splines to model transformations, for the best performance. Implemented using `glasflow` [32], the model has 3 block transforms, 6 layers, and 120 neurons per layer. It is trained over 500 epochs with a learning rate of 0.0005 and takes approximately 5 hours to train using a NVIDIA GeForce RTX 2080 Ti GPU.

VI. RESULTS

We present our results in Fig. 1, where we show the posterior distributions over the Hubble constant for each of the BBHs observed during observing runs O1 through to O3b assuming a flat prior on H_0. The Jensen-Shannon (JS) divergence between the distributions obtained from `CosmoFlow` and those from `gwcosmo` are also shown in each plot In order to compare with results presented in [11], the events were selected (and the NF models trained) based on a multidetector SNR threshold $\rho_{\mathrm{th}} = 11$. We used 10,000 samples from the posterior distributions of the Gravitational-wave Transient catalog [9,33] to evaluate the expectation value within Eq. (6). Our individual events results are shown in comparison to the `gwcosmo` results from [11].

The posterior distributions over the 42 BBH events computed using `CosmoFlow` show good agreement with

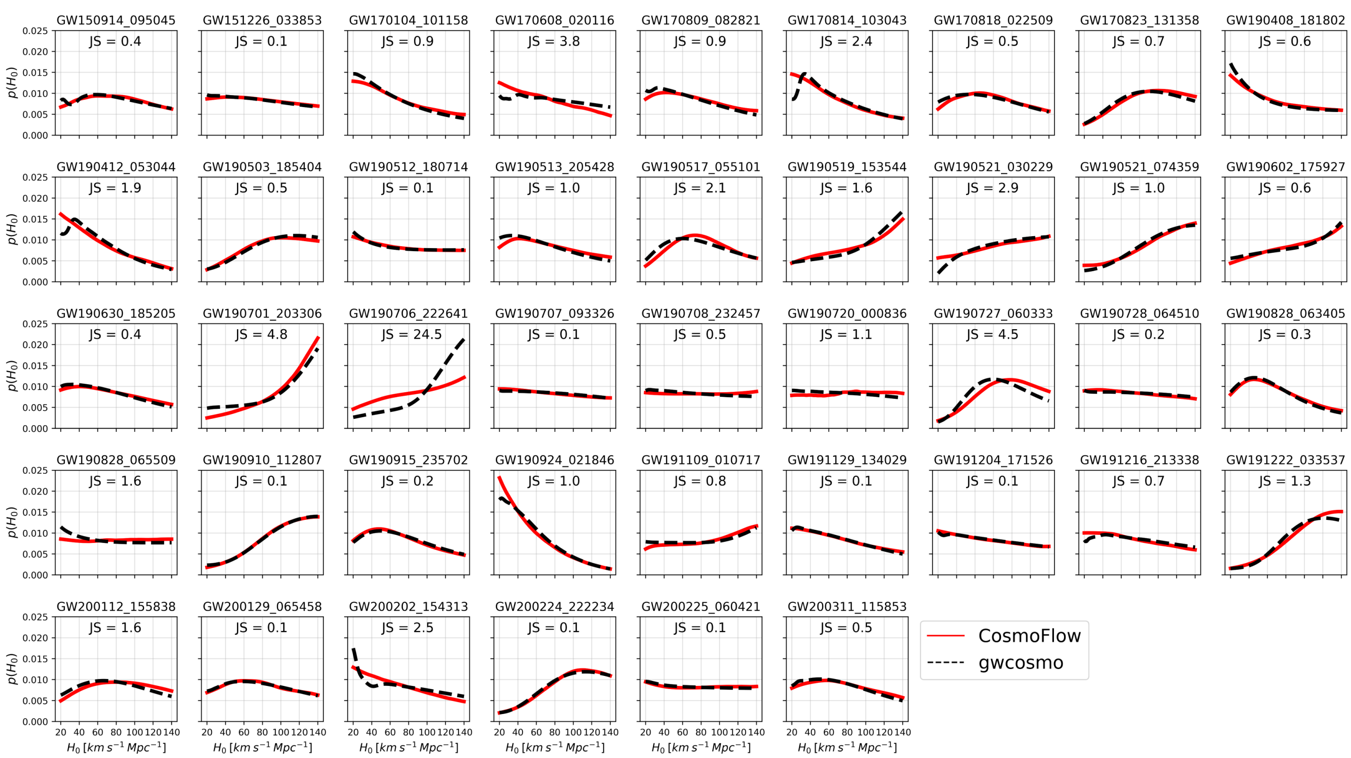

FIG. 1. Posterior distributions of H_0 for 42 BBH events detected during O1, O2, and O3 detection eras using `CosmoFlow` (solid red), compared with those from [11], using `gwcosmo` (dashed black). In all cases, a uniform prior is assumed on H_0. In each of the plots, the JS divergence between the two distributions is displayed at the top center, measured in millinats.

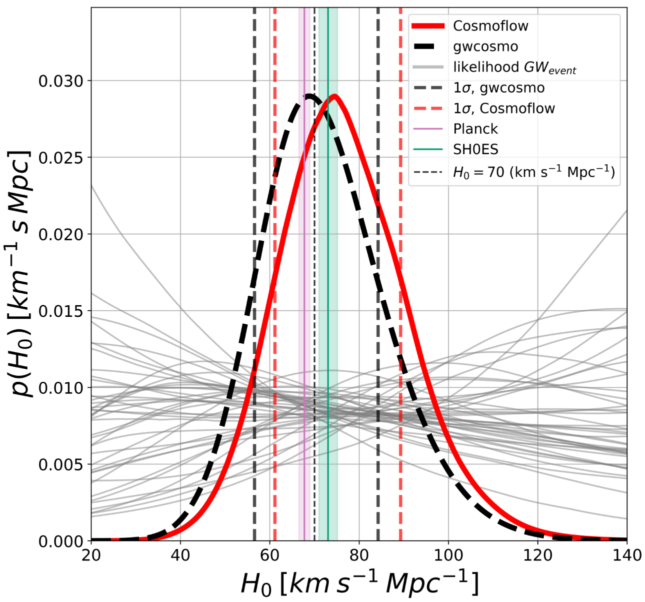

FIG. 2. Combined posteriors from Fig. 1, comparing CosmoFlow (solid red) with gwcosmo (dashed black). The single event likelihoods for CosmoFlow are plotted in the background in gray. Current Planck [4] and SH0ES [5] estimates of the Hubble constant are also plotted in pink and green, respectively, with 3σ uncertainties. The 1σ boundaries are also plotted for both CosmoFlow and gwcosmo.

the gwcosmo analysis results. Our analysis very effectively captures the features within the posteriors that are driven by population information, e.g., component mass distributions, comoving volume, and merger rate. There are a handful of events whose parameters allow for a reasonable probability that they originate from a galaxy within the catalog. In these cases the gwcosmo posteriors exhibit structure at low H_0 values caused by corresponding structure within the catalog. The CosmoFlow analysis does not show this structure as prominently and indicates that the NF could be developed and trained further to capture this in more detail.

By using Eq. (6) to combine the likelihoods from all 42 BBH events, and assuming a flat prior, we obtain the full combined posterior distribution on H_0 for both methodologies. This is shown in Fig. 2. Results show a value of $H_0 = 74.51^{+14.80}_{-13.63}$ km s^{-1} Mpc^{-1} (1σ highest density interval) computed using CosmoFlow, compared to gwcosmo, which gives a measurement $H_0 = 68.72^{+15.64}_{-12.41}$ km s^{-1} Mpc^{-1}. It is clear that the agreement between the two independent analyses is driven by the fact that both results are informed primarily through population information and the common assumptions in our models related to volumes of space with the galaxy catalog. Other studies, such as [22,34], show similarity with our results.

VII. DISCUSSION

This paper presents a methodology for using a NF via carefully generated training data to model an EM-informed and cosmology dependent prior on GW parameters. This fast-to-evaluate prior function can be combined with computationally costly precomputed posterior samples from individual GW events to evaluate the likelihood of the GW measurement given a specific cosmology. Results from individual events can then be used to provide a combined posterior distribution on cosmological parameters. We also provide an example analysis mirroring that of the existing standard approach where we show good agreement in our results on the posterior distribution of the Hubble constant using 42 BBH detections from the advanced GW detectors.

Our results exhibit good agreement with the established gwcosmo pipeline, a methodology employed in the advanced GW detector O3 analysis for cosmological parameter inference. The analysis showcases strong agreement in posteriors largely informed by the population characteristics of the posterior distribution for cosmological parameter inference, as depicted by the consistent matching trends seen in Fig. 1. This consistency leads naturally to a strong agreement between the combined posteriors over the 42 BBH events where the majority of inferred cosmological information stems from the population characteristics of the events.

We note some discrepancies between our results and those of [11] for events that exhibit posterior structure at low values of H_0. These features, which are present in the gwcosmo posterior distributions, are not as prominent in the CosmoFlow posterior distributions. Reconstructing these features is challenging and demands a larger training dataset and additional development of the NF structure to more accurately capture the galactic clustering within the GLADE+ catalog. Despite this challenge, since the discrepant posterior structure does not overlap with the area of combined posterior support (at $\sim$70 km s^{-1} Mpc^{-1}) we obtain very good agreement with the corresponding gwcosmo result. We also note that while we have used the gwcosmo analysis as a benchmark, it is not fair to treat it as an absolute standard since it contains (as do most computational analyses including our own) a finite number of minor approximations. This method additionally introduces flexibility in selecting the number of cosmological and/or population parameters for inference as well as the number of GW parameters to use as input. Our method requires only that training data consist of GW parameter samples, conditional on cosmological and or population parameter samples, and that a suitably large number of these samples be generated prior to training. This enables the extension of the analysis into inference of multidimensional cosmological and/or population parameter spaces, where subsequent comparisons with recent improvements to gwcosmo [35] can be conducted. Future work will also address the complexity of astrophysical model choices, e.g., the redshift evolution of the BBH mass distribution parameters [36].

As is to be expected with graphics processing unit (GPU) powered machine learning approaches, we can also highlight the inherent computational speed of `CosmoFlow`. The evaluation of the likelihood [the expectation value in Eq. (6)] for a single conditional parameter and 1000 GW posterior samples takes ∼1 ms on a single GPU, enabling swift computation over a range of cosmological parameter values.

It is important to address the issue of combining joint posteriors from individual events in the case of multidimensional posterior distributions over cosmological and population parameters. For our one-dimensional analysis, evaluating the likelihood for each BBH on a common vector of H_0 values was sufficient. For higher dimensional cases, traditional Bayesian sampling methods can be used, e.g., nested sampling [37] and Markov chain Monte Carlo [38]. However, the inherent parallelism within the `CosmoFlow` approach makes inefficient rejection sampling approaches a feasible and simple alternative.

ACKNOWLEDGMENTS

The authors thank Michael J. Williams and Christian Chapman-Bird for insightful discussions about training normalizing flows and insights for the optimization of the data generation. The authors also thank the members of the Data Analysis Group of the Institute of Gravitational Research and the LVK Cosmology group for helpful discussions. F. S. is supported by the University of Glasgow CoSE. C. M. and M. H. are supported by the Science and Technology Research Council [ST/V005634/1]. This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation.

[1] B. F. Schutz, Nature (London) **323**, 310 (1986).
[2] D. W. Hogg, arXiv:astro-ph/9905116.
[3] W. L. Freedman, Nat. Astron. **1**, 0121 (2017).
[4] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo *et al.*, Astron. Astrophys. **641**, A6 (2020).
[5] A. G. Riess, W. Yuan, L. M. Macri, D. Scolnic, D. Brout, S. Casertano, D. O. Jones, Y. Murakami, G. S. Anand, L. Breuval, T. G. Brink, A. V. Filippenko, S. Hoffmann, S. W. Jha, W. D. Kenworthy, J. Mackenty, B. E. Stahl, and W. Zheng, Astrophys. J. Lett. **934**, L7 (2022).
[6] The LIGO Scientific, The Virgo, and The KAGRA Collaborations, Living Rev. Relativity **23**, 3 (2020).
[7] J. Aasi *et al.* (LIGO Scientific Collaboration), Classical Quantum Gravity **32**, 074001 (2015).
[8] F. Acernese *et al.* (Virgo Collaboration), Classical Quantum Gravity **32**, 024001 (2015).
[9] The LIGO Scientific, The Virgo, and The KAGRA Collaborations, Phys. Rev. X **13**, 041039 (2023).
[10] The LIGO Scientific, The Virgo, and The KAGRA Collaborations, Phys. Rev. X **13**, 011048 (2023).
[11] The LIGO Scientific, The Virgo, and The KAGRA Collaborations, Astrophys. J. **949**, 76 (2023).
[12] R. Gray, I. M. n. Hernandez, H. Qi, A. Sur, P. R. Brady, H.-Y. Chen, W. M. Farr, M. Fishbach, J. R. Gair, A. Ghosh, D. E. Holz, S. Mastrogiovanni, C. Messenger, D. A. Steer, and J. Veitch, Phys. Rev. D **101**, 122001 (2020).
[13] S. Mastrogiovanni, G. Pierra, S. Perriès, D. Laghi, G. C. Santoro, A. Ghosh, R. Gray, C. Karathanasis, and K. Leyde, Astron. Astrophys. **682**, A167 (2024).
[14] S. Mukherjee, A. Krolewski, B. D. Wandelt, and J. Silk, arXiv:2203.03643.
[15] S. Mukherjee, B. D. Wandelt, S. M. Nissanke, and A. Silvestri, Phys. Rev. D **103**, 043520 (2021).
[16] C. C. Diaz and S. Mukherjee, Mon. Not. R. Astron. Soc. **511**, 2782 (2022).
[17] I. Kobyzev, S. Prince, and M. Brubaker, IEEE Trans. Pattern Anal. Mach. Intell. **43**, 3964 (2020).
[18] M. J. Williams, J. Veitch, and C. Messenger, Phys. Rev. D **103**, 103006 (2021).
[19] M. J. Williams, J. Veitch, and C. Messenger, Mach. Learn. Sci. Technol. **4**, 035011 (2023).
[20] M. Dax, S. R. Green, J. Gair, J. H. Macke, A. Buonanno, and B. Schölkopf, Phys. Rev. Lett. **127**, 241103 (2021).
[21] D. Ruhe, K. Wong, M. Cranmer, and P. Forré, arXiv:2211.09008.
[22] T. Dacunha, M. Raveri, M. Park, C. Doux, and B. Jain, Phys. Rev. D **105**, 063529 (2022).
[23] G. Papamakarios, E. Nalisnick, D. J. Rezende, S. Mohamed, and B. Lakshminarayanan, J. Mach. Learn. Res. **22**, 2617 (2021).
[24] G. Papamakarios, T. Pavlakou, and I. Murray, arXiv:1705.07057.
[25] C. Winkler, D. Worrall, E. Hoogeboom, and M. Welling, arXiv:1912.00042.
[26] R. Gray, I. M. Hernandez, H. Qi, A. Sur, P. R. Brady, H.-Y. Chen, W. M. Farr, M. Fishbach, J. R. Gair, A. Ghosh *et al.*, Phys. Rev. D **101**, 122001 (2020).
[27] G. Dá lya, G. Galgóczi, L. Dobos, Z. Frei, I. S. Heng, R. Macas, C. Messenger, P. Raffai, and R. S. de Souza, Mon. Not. R. Astron. Soc. **479**, 2374 (2018).
[28] In general, this distribution is dependent on the cosmological parameters Ω with the exception of the Hubble constant.
[29] R. Gray, https://theses.gla.ac.uk/82438/ (2021).

[30] G. Ashton *et al.*, Astrophys. J. Suppl. Ser. **241**, 27 (2019).
[31] C. E. A. Chapman-Bird, C. P. L. Berry, and G. Woan, Mon. Not. R. Astron. Soc. **522**, 6043 (2023).
[32] M. J. Williams, J. Mcginn, F. Stachurski, and J. Veitch, uofgravity/glasflow: v0.2.0 (2023).
[33] https://dcc.ligo.org/LIGO-P2000318/public (2021).
[34] A. Palmese, C. R. Bom, S. Mucesh, and W. G. Hartley, Astrophys. J. **943**, 56 (2023).
[35] R. Gray, F. Beirnaert, C. Karathanasis, B. Revenu, C. Turski, A. Chen, T. Baker, S. Vallejo, A. E. Romano, T. Ghosh, A. Ghosh, K. Leyde, S. Mastrogiovanni, and S. More, J. Cosmol. Astropart. Phys. 12 (2023) 023.
[36] C. Karathanasis, S. Mukherjee, and S. Mastrogiovanni, Mon. Not. R. Astron. Soc. **523**, 4539 (2023).
[37] J. Skilling, AIP Conf. Proc. **735**, 395 (2004).
[38] W. K. Hastings, Biometrika **57**, 97 (1970).