Astronomy & Astrophysics manuscript no. paper
October 13, 2022

GWSim: A python package to create GW mock samples for different astrophysical populations and cosmological models of binary black holes

Christos Karathanasis[1]⋆, Benoît Revenu[2,3], Suvodip Mukherjee[4]⋆⋆, and Federico Stachurski[5]

[1] Institut de Física d'Altes Energies (IFAE), Barcelona Institute of Science and Technology, Barcelona, Spain
[2] Subatech, CNRS - Institut Mines-Telecom Atlantique - Université de Nantes, France
[3] Université Paris Cité, CNRS, Astroparticule et Cosmologie, France
[4] Department of Astronomy & Astrophysics, Tata Institute of Fundamental Research, 1, Homi Bhabha Road, Colaba, Mumbai 400005, India
[5] Institute of Gravitational Research (IGR), University of Glasgow, Glasgow, United Kingdom

October 13, 2022

ABSTRACT

Precision cosmology with gravitational wave (GW) sources requires understanding the interplay between GW source population and cosmological parameters governing the dynamics of the Universe. With the fast increase of GW detections, for exploring many aspects of cosmology and fundamental physics it is necessary to develop a tool which can simulate GW mock samples for several population and cosmological models with and without a galaxy catalog. We have developed a new code called `GWSim`, allowing to make GW mock events from a large range of configurations, varying the cosmology, the merger rate, and the GW source parameters (mass and spin distributions in particular), for a given network of GW detectors. We restrict the cosmology to spatially flat universes, including models with varying dark energy equation of state. `GWSim` provides each mock event with a position in the sky and a redshift; these values can be those of random host galaxies coming from an isotropic and homogeneous simulated Universe or a user-supplied galaxy catalog. We use realistic detector configurations of the LIGO and Virgo network of detectors to show the performance of this code for the latest observation runs and the upcoming observation run.

Key words. Cosmology: observations–cosmological parameters–Gravitational waves–Galaxies

1. Introduction

The detection of the gravitational waves (GWs) sources by the LIGO-Virgo-KAGRA collaboration (Gregory 2010; Aso et al. 2013; Acernese et al. 2014; Aasi et al. 2015; Abbott et al. 2016a,b, 2018; Tse et al. 2019; Abbott et al. 2019a; Akutsu et al. 2020; Abbott et al. 2021b,e,c) has opened the era of GW astronomy capable to bring plethora of scientific discoveries in astrophysics, cosmology and fundamental physics. Such discoveries will require understanding different sources of statistical and systematic uncertainties as well as the interplay between the astrophysical population of GW sources, the cosmological models, the fundamental physics, and searches beyond the standard model (Abbott et al. 2019b, 2021a,f,d). To understand these, we need realistic simulations of GW events for different astrophysical populations and cosmological models, including realistic detector sensitivities.

These simulations are produced by `GWSim` which simulates a sample of mock GW sources for different (i) mass, (ii) spin, and (iii) merger rate models, as well as for flat-ΛCDM and flat w_0-w_aCDM cosmological models. The code assigns host galaxies to GW events in two cases: assuming a constant in comoving volume distribution of galaxies or using a simulated galaxy catalog, such as `MICECATv2` (Fosalba et al. 2014; Crocce et al. 2015; Fosalba et al. 2015). The code also considers different detector sensitivities and observation periods. Then, the `Bilby` (Ashton et al. 2019) pipeline is used to generate the actual GW strains corresponding to these GW events and perform parameter estimation to finally obtain mock GW samples. Having such sets of samples allows us to estimate the performances of the analysis algorithms. For one given GW event, its associated samples consist of a list of values obtained by an MCMC algorithm of all parameters describing the binary system, in particular the luminosity distance of the source, the masses of the two black holes, and their spins.

The paper is structured as follows. In Sec. 2, we describe different models included in the current version of `GWSim`. In Sec. 3 and Sec. 4, we describe the structure of the code and validated different setups. Finally, in Sec. 5, we discuss the use and future developments of this code.

2. Models of cosmology and GW sources astrophysical population

Inference of the cosmological results using binary black holes (BBHs) will require understanding the dependencies on the astrophysical source parameters for different techniques such as statistical host identification (Schutz 1986; MacLeod & Hogan 2008; Del Pozzo 2012; Arabsalmani et al. 2013; Gray et al. 2020; Fishbach et al. 2019; Abbott et al. 2021a; Soares-Santos et al. 2019; Finke et al. 2021; Abbott et al. 2020b; Palmese et al. 2021), cross-correlation technique after taking into account GW

⋆ ckarathanasis@ifae.es
⋆⋆ suvodip@tifr.res.in

source population uncertainty (Oguri 2016; Mukherjee & Wandelt 2018; Mukherjee et al. 2020; Calore et al. 2020; Mukherjee et al. 2021a; Bera et al. 2020; Scelfo et al. 2020; Mukherjee et al. 2021b; Cañas-Herrera et al. 2021; Scelfo et al. 2022; Cigarrán Díaz & Mukherjee 2022; Mukherjee et al. 2022), or BBH mass-distribution (Taylor et al. 2012; Farr et al. 2019; Mastrogiovanni et al. 2021b; You et al. 2021; Mancarella et al. 2021; Mukherjee 2022; Leyde et al. 2022; Ezquiaga & Holz 2022).

We describe here the cosmological models that can be used in `GWSim` and the models describing the population of BBH mergers (merger rates, mass distributions, and spin distributions).

2.1. Cosmological models

The code considers only the cosmological models with spatial curvature zero (flat) and for three scenarios namely, Lambda Cold Dark Matter (ΛCDM), w_0CDM, which takes into account variation of the dark energy equation of state without redshift dependence, and w_0-w_aCDM, which takes into account variation of the dark energy equation of state with redshift dependence (Chevallier & Polarski 2001; Linder 2003). In Tab. 1, we present these models. For a particular choice of the cosmological parameters, the code computes the luminosity distances $d_l(z_m)$ to the merger situated at a redshift z_m using

$$d_l(z_m) = \frac{c(1+z_m)}{H_0}\int_0^{z_m}\frac{\mathrm{d}z}{E(z)}, \tag{1}$$

with $E^2(z) \equiv \Omega_m(1+z)^3 + \Omega_{DE}\exp[\int_0^z \mathrm{d}\ln(1+z')(1+\omega(z'))]$.

In the part of the code which is used for assigning the redshift to the GW sources from a galaxy catalog, there we calculate the luminosity distance to the source for different cosmological models. Though this part of the code can be used to infer the cosmological parameters using the luminosity distance, it cannot be used to estimate the cosmological parameters using the spatial clustering of the GW sources. The spatial clustering of the GWs sources will also depend on the intrinsic choice of the cosmological parameters. So it can work only for the fiducial values of the cosmological parameters considered by `MICECATv2` simulations.

2.2. Merger rate models of the compact objects

The merger rate of the binary compact objects is modeled for three different scenarios which we call the (a) constant merger rate, (b) phenomenological merger rate, and (c) delay-time model. The models are described in Tab. 2 and their parameters can be seen in Tab. 3.

The first scenario with a constant merger rate model is a model with a single parameter R_0 which corresponds to a scenario with no redshift evolution. The second scenario of the model considers a merger rate varying with the redshift; it uses four phenomenological parameters (Madau & Dickinson 2014):

$$R(z) = R_0(1+z)^{\gamma}\frac{1+(1+z_p)^{-(\gamma+\kappa)}}{1+\left(\frac{1+z}{1+z_p}\right)^{\gamma+\kappa}}. \tag{2}$$

The third model we consider is a physics-driven model that models the merger rate in terms of the delay time distribution denoted by t_d (O'Shaughnessy et al. 2010; Banerjee et al. 2010; Dominik et al. 2012; Dominik et al. 2015; Mandel & de Mink 2016; Lamberts et al. 2016; Cao et al. 2018; Elbert et al. 2018; Eldridge et al. 2019; Vitale et al. 2019; du Buisson et al. 2020; Santoliquido et al. 2021). Currently the constraints are weak from individual events (Fishbach & Kalogera 2021; Karathanasis et al. 2022) and the stochastic GW background (Mukherjee & Silk 2021). In the future, by combing the properties of emission line galaxies with GW sources, a better measurement of the delay time is possible (Mukherjee & Dizgah 2021). The time delay used in this model refers to the elapsed time between the formation of the stars, which later will become black holes, and the merging of the black holes. The time delay is not the same value for all binary black holes, but it is rather following a distribution (see Eq. 4). For a star to become a black hole only a few Myrs are required, but for a black hole to find another companion, form a binary, and merge a significant amount of time is needed (in the order of Gyrs). This model assumes a star formation rate and different power-law forms of the probability distribution of the delay time $P_t(t_d|t_d^{min}, t_d^{\max}, d)$ (Mukherjee 2022; Karathanasis et al. 2022). This model has three parameters as described in Tab. 3 and the corresponding merger rate at redshift z can be written as

$$R_{\rm TD}(z) = R_0\frac{\int_z^{\infty}P_t(t_d|t_d^{min}, t_d^{\max}, d)R_{\rm SFR}(z_f)\frac{\mathrm{d}t}{\mathrm{d}z_f}\mathrm{d}z_f}{\int_0^{\infty}P_t(t_d|t_d^{min}, t_d^{\max}, d)R_{\rm SFR}(z_f)\frac{\mathrm{d}t}{\mathrm{d}z_f}\mathrm{d}z_f}, \tag{3}$$

where the probability distribution of the delay time can be written as

$$P_t(t_d|t_d^{min}, t_d^{\max}, d) \propto \begin{cases}(t_d)^{-d} & \text{, for } t_d^{\min} < t_d < t_d^{\max}\\ 0 & \text{, otherwise}\end{cases}, \tag{4}$$

For each of this merger rates $R(z)$, the number of coalescing events per unit redshift per unit observation time is calculated using

$$\frac{\mathrm{d}N_{GW}}{\mathrm{d}z\,\mathrm{d}t} = \frac{R(z)}{1+z}\frac{\mathrm{d}V_c}{\mathrm{d}z}(\theta_c), \tag{5}$$

where $\frac{\mathrm{d}V_c}{\mathrm{d}z}(\theta_c)$ is the comoving volume at redshift z which depends on the set of cosmological parameters denoted by θ_c. For a given observation time, duty cycle, and redshift range, the total number of events can be obtained by integrating the above equation. As an example, using the illustrative values of parameters shown in Tab.: 4, the number of detectable events is of the order of 150 ($\pm$10%) for a total observation time of 3 years with a network signal-to-noise ratio (SNR) higher than $\mathrm{SNR}_{\rm thr} = 12$. The network SNR is defined as the square root of the quadratic sum of the individual SNR measured in the interferometers: $\mathrm{SNR} = (\sum_i \mathrm{SNR}_i^2)^{1/2}$.

2.3. Mass models

The GW signal strongly depends on the individual masses m_1 and m_2 (with $m_1 \geqslant m_2$) of the two black holes forming the compact system. The joint distribution $p(m_1, m_2)$ and the marginalized $p(m_1)$ and $p(m_2)$ depend on the black hole formation channels. We consider broadly two scenarios of phenomenological mass models, namely the redshift-independent mass models and redshift-dependent mass models. The mass models are described in Tab. 2 and their associated parameters in Tab. 3.

2.3.1. Redshift-independent mass models

For the redshift-independent mass models, we include scenarios with only (a) a Truncated power-law model without

Cosmology

Definition of the cosmological parameters

Parameter	Description
H_0	The Hubble constant in km/s/Mpc.
Ω_m	Present-day matter density of the Universe.
$w(z) = w_0 + w_a\, z/(1+z)$	Dark energy equation of state and its redshift dependence

Cosmological models considered

Model	Parameters
Flat ΛCDM :	H_0, Ω_m, $w_0 = -1$, $w_a = 0$
Flat w_0CDM:	H_0, Ω_m, w_0, $w_a = 0$
Flat w_0-w_aCDM :	H_0, Ω_m, w_0, w_a

Table 1: Description of the cosmological models available and the parameters that govern them.

Models considered

Merger rate models

Model	Parameters
Redshift independent :	$R_0, \gamma = 0, \kappa = 0, z_p = 0$
Redshift dependent :	R_0, γ, κ, z_p

Mass models

Model	Parameters
Truncated powerlaw:	$\alpha, \beta, M_{\rm min}, M_{\rm max}$
Broken powerlaw:	$\alpha, \alpha_2, \beta, M_{\rm min}, M_{\rm max}, \delta_m, b$
Powerlaw+Gaussian:	$\alpha, \beta, M_{\rm min}, M_{\rm max}, \delta_m, \mu_g, \sigma_g, \lambda_g$
Multi peak:	$\alpha, \beta, M_{\rm min}, M_{\rm max}, \delta_m, \mu_g, \sigma_g, \lambda_g, \mu_{\rm g,low}, \sigma_{\rm g,low}, \lambda_{\rm g,low}$
Redshift dependent:	All parameters of each model from above+one ϵ for each parameter.

Spin models

Model	Parameters
Uniform:	–
Gaussian:	χ_0, σ_0
Heavy mass:	m_{th}
Correlated:	$\alpha_s, \beta_s, \chi_0, \sigma_0$

Time delay model

Model	Parameters
Time delay:	$R_0, \gamma = 2.7, \kappa = 2.9, z_p = 1.9, \alpha, \beta, M_{\rm min}, M_{\rm max}, \delta_m, M_{\rm PISN}(Z_*), \sigma_g, \lambda_g, d, t_d^{min}, \alpha_{\rm Z}, \gamma_{\rm Z}$

Table 2: Description of the population models available.

smoothing (i.e. there is a sharp cut-off at low mass values), (b) a broken power-law model with smoothing, (c) a power-law+Gaussian peak model with smoothing (d) a model with power-law+multipeak Gaussian model with smoothing. All parameters are in Tab. 3. The redshift independent models are motivated by the current studies of the rates and population of the GW sources from nearly ninety GW sources discovered by the LVK collaboration (Abbott et al. 2021b,d). For these mass models, the code creates a mock sample of the GW sources which follows the underlying distribution of both the companion masses m_1 and m_2 accordingly to the form (Talbot & Thrane 2018; Abbott et al. 2020a; LIGO et al. 2021; Abbott et al. 2021g)

$$P_s(m_1, m_2|\Phi_m) = p(m_1|\Phi_m)\, p(m_2|m_1, \Phi_m) \times S(m_1|\delta_m, M_{\rm min})\, S(m_2|\delta_m, M_{\rm min}), \quad (6)$$

where Φ_m refers to all the population parameters not related to cosmology and $p(m_2|m_1, \Phi_m)$ is the distribution of m_2, conditioned on the constraint $m_1 \geqslant m_2$, and S is the function used to smooth the lower end of the mass distributions, taken as a sigmoid. For the model without smoothing, $S = 1$. (Abbott et al. 2021d)

GW source population

Definition of the mass distribution parameters

Parameter	Description
α	Spectral index for the power-law of the primary mass distribution.
α_2	Second spectral index for the power-law of the primary mass distribution in case.
β	Spectral index for the power-law of the mass ratio distribution.
b	Point in the mass distribution were the slope changes in $M_\odot$.
$M_{\rm min}$	Minimum mass of the power-law component of the primary mass distribution in $M_\odot$.
$M_{\rm max}$	Maximum mass of the power-law component of the primary mass distribution in $M_\odot$.
$\lambda_{\rm g}$	Fraction of the model in the Gaussian component.
$M_{\rm PISN}(Z_*)$	The value of the PISN mass scale in $M_\odot$ units at metallciity $Z_* = 10^{-4}$.
$\sigma_{\rm g}$	Width of the Gaussian component in the primary mass distribution in $M_\odot$.
δ_m	Range of mass tapering at the lower end of the mass distribution in $M_\odot$.
$\lambda_{\rm g,low}$	Fraction of the model in the secondary Gaussian component.
$\mu_{\rm g,low}$	Mean of the secondary Gaussian component in the primary mass distribution in $M_\odot$.
$\sigma_{\rm g,low}$	Width of the secondary Gaussian component in the primary mass distribution in $M_\odot$.
ϵ_x	Redshift evolution parameter of the x parameter.

Definition of the merger rate parameters

Parameter	Description
γ	Power law index of the merger rate model.
κ	Power law index of the merger rate model.
z_p	The redshift where the merger rate model changes slope.
R_0	Value of the merger rate at $z = 0$ in Gpc^{-3} yr^{-1}.

Definition of the time delay parameters

Parameter	Description
d	Spectral index for the power-law of the delay time distribution.
$t_d^{\rm min}$	Minimum time for the power-law of the delay time distribution in Gyrs.
$\alpha_{\rm Z}$	The parameter that captures a weak logarithmic dependence of $M_{\rm PISN}$ on the metallicity.
$\gamma_{\rm Z}$	The parameter that captures the redshift dependence of the metallicity.

Definition of the spin parameters

Parameter	Description
m_{th}	The mass threshold value below which spins are set to 0.
$\mu_{\chi,0}$	The mean of the Gaussian distribution for $\chi_{\rm eff}$.
$\sigma_{\chi,0}$	The standard deviation of the Gaussian distribution for $\chi_{\rm eff}$.
α_s	Mass-ratio correlation parameter for the expectation value.
β_s	Mass-ratio correlation parameter for the standard deviation.

Table 3: Description of the parameters that govern the population models.

2.3.2. Redshift-dependent mass models

The redshift dependence of the mass distribution of BBHs can arise due to the evolution of the stellar properties in the Universe (Spera & Mapelli 2017; Farmer et al. 2019; Renzo et al. 2020; Mukherjee 2022). Two different scenarios of the redshift-dependent mass model are considered in the analysis, namely (a) a phenomenological model which considers the redshift evolution of the mass distribution, and (b) a mixing-mass model: a model of the mass distribution which considers the dependence of the binary masses on the stellar metallicity and delay time distribution.

For the phenomenological redshift dependent mass distribution model, we consider a linear redshift dependence of the mass model that can is modeled as $\mu(z) = \mu_0 + \epsilon_\mu z$, where μ_0 denotes the expectation value of the mass of the redshift independent case (discussed previously) and ϵ_μ is the value of the parameter which controls a linear redshift evolution of this expectation value. This model is an extension of the redshift-independent model with one additional parameter for each of the redshift-independent mass model parameters. As a result, there are $2N$ parameters, if the redshift independent model has N parameters.

In the mixing-mass model (Mukherjee 2022; Karathanasis et al. 2022), the mass distributions of the compact objects get a redshift dependence due to a metallicity evolution as a function of redshift and a non-zero value of the delay time distribution. Due to the non-zero value of the delay time, black holes forming at different cosmic times originate from different parent stars with different metallicities and consequently can have different masses. Consequently, when the black holes form at different cosmic times, they merge after a different amount of delay time and the observed BH mass distribution gets a redshift depen-

Cosmology: flat ΛCDM		Merger rate		Mass distributions: Power-law+Gaussian peak	
Parameter	Value	Parameter	Value	Parameter	Value
H_0	67.8 km s^{-1} Mpc^{-1}	γ	2.7	α	3.4
Ω_m	0.3	κ	2.9	β	0.8
Ω_Λ	0.7	z_p	1.9	λ_g	0.04
		R_0	20 Gpc^{-3} yr^{-1}	μ_g	35 M$_\odot$
				σ_g	3.9 M$_\odot$
				δ_m	4.8 M$_\odot$
				M_{max}	100 M$_\odot$
				M_{min}	5 M$_\odot$

Duty factors				
Detector	Observation Run			
	O1	O2	O3	O4
H	0.6	0.6	0.75	0.75
L	0.5	0.6	0.75	0.75
V	-	-	0.75	0.75

Table 4: Tables show the fiducial value of the parameters and duty cycle used in the analysis.

dence. This redshift-dependent mass distribution $P[m(z_m)]$ can be written in terms of the window function $W_{t_d}[m(z_m)]$ and the source frame mass distribution $P_s[m(z_m)]$ of the compact objects as

$$P[m(z_m)] = P_s[m(z_m)]W_{t_d}[m(z_m)], \tag{7}$$

where the window function can be written as

$$W_{t_d}[m(z_m)] = N \int_{z_m}^{\infty} P_t(t_d|t_d^{min}, t_d^{\max}, d) \frac{\mathrm{d}t}{\mathrm{d}z_f} W[m(z_f)]\, \mathrm{d}z_f. \tag{8}$$

and N is a normalization constant. Further details about this model can be found in (Mukherjee 2022; Karathanasis et al. 2022).

2.4. Spin models

The GW strain also depends on the individual spins (amplitude and orientation) of both compact objects. The study of the spin of GW sources can shed light on the astrophysical population and the formation channel of these sources (Mandel & O'Shaughnessy 2010; Kushnir et al. 2016; Rodriguez et al. 2016; Stevenson et al. 2017; O'Shaughnessy et al. 2017; Farr et al. 2017; Gerosa et al. 2018; Fuller & Ma 2019; Bavera et al. 2020; Biscoveanu et al. 2021b,a, 2022). The spin of a black hole is defined as $\boldsymbol{S} = S\hat{\boldsymbol{S}}$, where $|\hat{\boldsymbol{S}}| = 1$ and $S = Gm^2\chi/c$, with χ being a dimensionless scalar parameter in the range of $\chi \in [0; 1]$. Denoting the total angular momentum of the binary system as $\boldsymbol{J}$, the tilt angles of the spin vectors with respect to $\boldsymbol{J}$ are defined as $\cos\theta_i = \hat{\boldsymbol{S}}_i \cdot \boldsymbol{J}/J$. Aligned spins correspond to the cases $\theta_1 = \theta_2 = 0$ or $\theta_1 = \theta_2 = \pi$. In practice, GW measurements are sensitive to the following combination of the spins, which defines the effective spin $\chi_{\rm eff}$ of the binary system:

$$\chi_{\rm eff} = \frac{m_1\chi_1\cos\theta_1 + m_2\chi_2\cos\theta_2}{m_1 + m_2} = \frac{\chi_1\cos\theta_1 + q\chi_2\cos\theta_2}{1+q}, \tag{9}$$

where q is the mass-ratio $q = m_2/m_1 (q \in [0; 1])$ and $\chi_{\rm eff}$ takes values in the range of $\chi_{\rm eff} \in [-1; 1]$. Individual spin components are not very well constrained with the current interferometers. Thus, we deduce the individual values from $\chi_{\rm eff}$ and not the other way. We are considering two families of scenarios for the spin distribution of the binary compact objects, namely the redshift independent and the redshift dependent scenarios. The spin models are described in Tab. 2 and the details of their parameters are summarised in Tab. 3.

2.4.1. Redshift independent spin models

The redshift independent case is considered in two models. The first one consists in taking, for each GW event, a random $\chi_{\rm eff}$ spin value from a uniform distribution $\mathcal{U}[-1, 1]$. Then, the individual parameters (χ_i, θ_i) are computed given $\chi_{\rm eff}$ but there is a strong degeneracy: many tuples of $(\chi_1, \theta_1, \chi_2, \theta_2)$ given q correspond to the same $\chi_{\rm eff}$ so we present here one way to get such a tuple. If the spins are requested to be aligned then the computation of (χ_i, θ_i) is performed as follows: set first both angles θ_i to 0 or π for a positive and negative value of $\chi_{\rm eff}$ respectively. Then draw both values of χ_i following:

$$\begin{aligned}
&\chi_2 \sim \mathcal{U}[a, b] \\
&\text{with } a = \max(0, ((1+q)|\chi_{\rm eff}| - 1)/q), \\
&\text{and } b = \min(1, (1 + (1+q)|\chi_{\rm eff}|)/q) \\
&\text{then } \chi_1 = (1+q)|\chi_{\rm eff}| - q\chi_2.
\end{aligned}$$

If the spin directions are requested as random, then we draw the values following:

$$\begin{aligned}
&\chi_2\cos\theta_2 \sim \mathcal{U}[a; b] \\
&\text{with } a = \max(-1, ((1+q)\chi_{\rm eff} - 1)/q), \\
&\text{and } b = \min(1, (1 + (1+q)\chi_{\rm eff})/q) \\
&\text{then } \chi_1\cos\theta_1 = (1+q)\chi_{\rm eff} - q\chi_2\cos\theta_2.
\end{aligned}$$

The degeneracy between individual spins χ_i and tilt angles θ_i is broken with:

$$\begin{aligned}
&\chi_2 = |\chi_2\cos\theta_2|^x \text{ with } x \sim \mathcal{U}[0; 1] \\
&\text{and finally } \cos\theta_2 = \chi_2\cos\theta_2/\chi_2 \\
&\chi_1 = |\chi_1\cos\theta_1|^y \text{ with } y \sim \mathcal{U}[0; 1] \\
&\text{and finally } \cos\theta_1 = \chi_2\cos\theta_1/\chi_1.
\end{aligned}$$

The second spin model independent of redshift allows a correlation between the individual masses and the spin values. We have considered two sub-classes of redshift independent spin

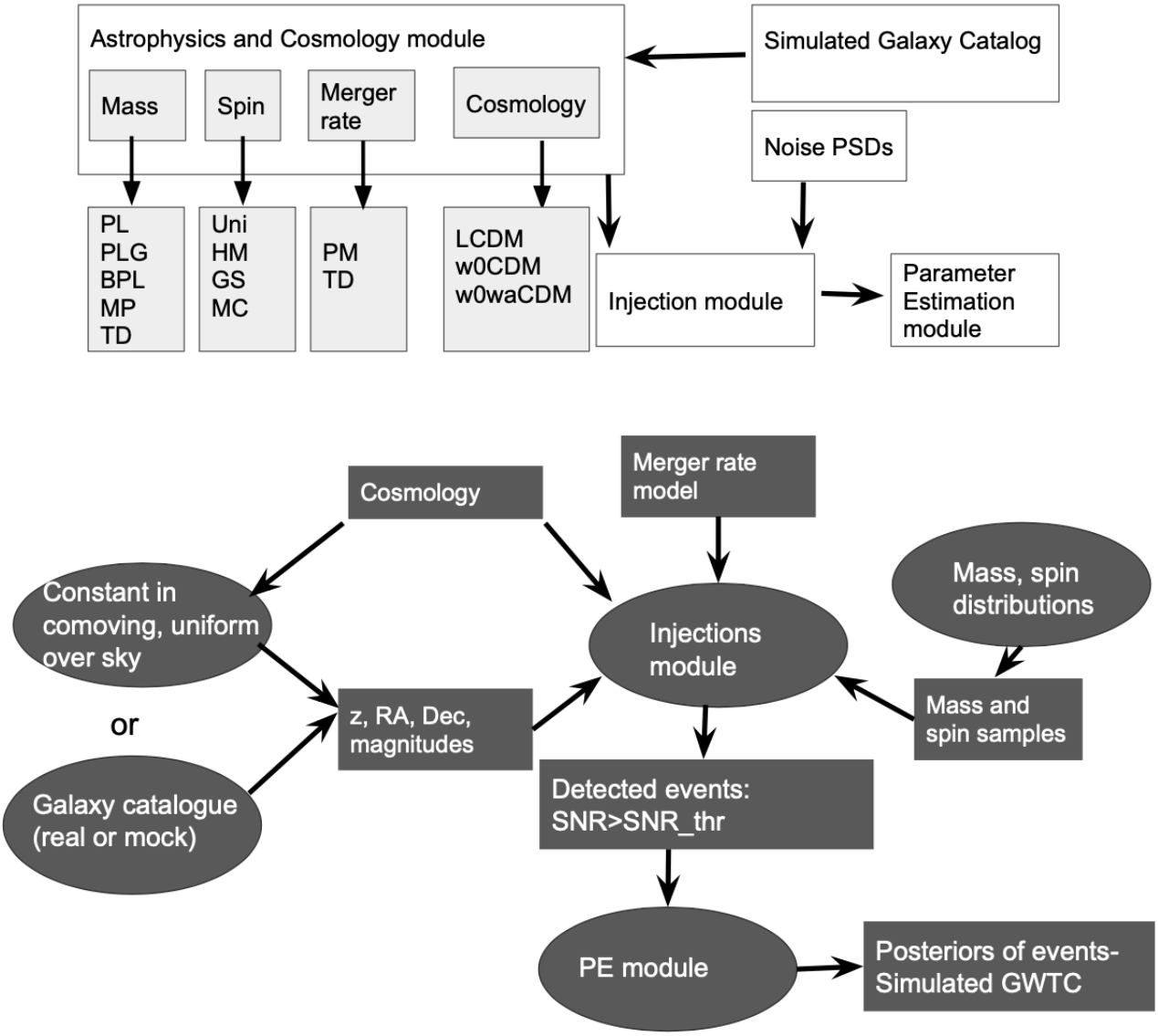

Fig. 1: (Top) Plots describes different modules which are present in the `GWSim` code. We show the different modules present under mass (Power-Law (PL), Power-Law with Gaussian (PLG), Broken Power-Law (BPL), Multi Peak (MP), and the time-delay (TD) model), spin (Uniform (Uni), Heavy-Mass (HM), Gaussian Spin (GS), and Mass-spin Correlated (MC)), merger rate (phenomenological model (PM), time-delay (TD)), and cosmology (LCDM, w_0CDM, w_0-w_aCDM)). (Bottom) In the plot, we show the flowchart describing the structure of the `GWSim` code and how the injection samples are produced.

models under: *(i)* objects with heavier masses have higher individual spin χ_i, *(ii)* the mass and spin of the black holes are (anti)correlated. In the first case, we choose a free threshold parameter of the mass values $m_{\rm th}$, above which the individual spin values are not zero and below which the spin is set to zero. In the second case, we include a correlation between the mass ratio and spin values as motivated in (Callister et al. 2021). In this model, $\chi_{\rm eff}$ follows the Gaussian distribution

$$p(\chi_{\rm eff}|q,\mu_{\chi,0},\sigma_{\chi,0},\alpha_s,\beta_s) \propto \exp\left[-\frac{\left(\chi_{\rm eff}-\mu_\chi(\mu_{\chi,0},\alpha_s,q)\right)^2}{2\sigma^2_\chi(\sigma_{\chi,0},\beta_s,q)}\right], \quad (10)$$

with an expectation value $\mu_\chi = \mu_{\chi,0} + \alpha_s(q-1/2)$ and a variance $\sigma^2_\chi(\sigma_{\chi,0},\beta_s,q)$ such that $\log_{10}\sigma_\chi(\sigma_{\chi,0},\beta_s,q) = \log_{10}\sigma_{\chi,0} + \beta_s(q-1/2)$. For $\alpha_s = \beta_s = 0$ the model is mass independent and $\chi_{\rm eff}$ follows a Gaussian. We refer to this particular model as the Gaussian spin model $\chi_{\rm eff} \sim \mathcal{N}(\mu_{\chi,0},\sigma_{\chi,0})$.

2.4.2. Redshift dependent spin models

We consider one scenario of the redshift dependent model, where the individual spin amplitude distribution is considered as Gaussian with a redshift dependent expectation value and standard deviation $P(\chi_i|z) = \mathcal{N}(\tilde{\chi}(z),\tilde{\sigma}(z))$, with $\tilde{\chi} = \chi_0 + \delta\chi\, z$, where χ_0 denotes the mean value without any redshift dependence, and $\delta\chi$ controls the redshift dependence part. Similarly, the stan-

dard deviation of the spin amplitude distribution is modeled as $\tilde{\sigma} = \sigma_0 + \delta\sigma\, z$, where σ_0 denotes the standard deviation without any redshift dependence, and the $\delta\sigma$ regulates the redshift dependence of the standard deviation. The spin orientations can be taken as aligned ($\theta_1 = \theta_2 = 0$ or $\theta_1 = \theta_2 = \pi$ randomly) or isotropically distributed.

3. Code description

GWSim is a code to make simulated GWTCs by assigning GW events to galaxies. The galaxies are taken from a galaxy catalog (or without the usage of a galaxy catalog, simply creating a redshift distribution of galaxies assuming they are distributed uniformly in comoving volume, i.e without any spatial clustering) for different astrophysical models of the binary compact objects and different cosmological models. The basic structure of the code and a flow chart showing the working principle of the code are presented in Fig. 1 respectively. We describe the structure of the code below.

The code is composed of five different modules, (i) astrophysical and cosmological module, (ii) galaxy catalog module, (iii) detector noise module, (iv) injection module, and (v) parameter estimation module. Under the astrophysical module, we have four sub-modules which include different models of the merger rate, source frame mass population, spin distribution, and cosmological models. Each of these modules contains different models which are discussed in Sec. 2. The galaxy catalog module includes functions to handle simulated galaxy catalogs which can be used to populate the GW sources into galaxies. In the next version of this code, scenarios with a real galaxy catalog will also be included. The third module is the detector module which includes the noise PSD for different detectors such as Hanford (H), Livingstone (L), Virgo (V) and different LVK observation runs such as O1, O2, O3 and the future run such as O4. These three modules are called at different stages in the fourth module called the injection module. The injection module is used to take into consideration all the aforementioned and simulate a GW strain of all GW sources for different choices of the astrophysical population models, cosmological models, detector noise, duty cycles, and observation time. The different GW strains are injected into the detector noise and the code calculates the SNR of all simulated events.

3.1. Simulated GWTC without galaxy catalog

In the injection module, the code uses a particular model of the merger rate and cosmology (specified by the user) and calculates the number of GW events (using Eq. (5)) up to a redshift $z_{\rm max}$. The value of $z_{\rm max}$ can be decided by the user. For each of the GW events, the code draws a random value from the source frame mass and spin distributions. The redshift of the event and the sky position is taken to be those of the selected galaxy. In this case, the galaxies are assumed to follow a uniform in comoving volume distribution and they are distributed uniformly over the sky. From the chosen redshifts, the luminosity distances to the GW sources are calculated using the selected cosmological model.

The final step of this process is to estimate the matched filtering SNR for a chosen network of detectors using the corresponding noise PSDs (which are available in the detector module). The GW events that have an SNR greater than the detection threshold $\rm SNR_{thr}$ chosen by a user, make up the detected *events* and are going to undergo parameter estimation using the publicly available code Bilby.

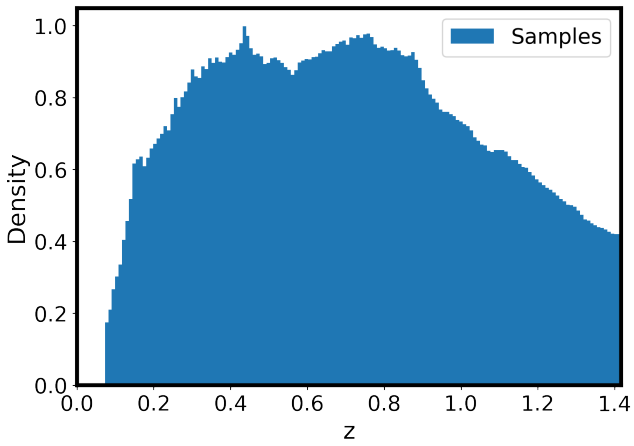

Fig. 2: True redshift distribution of galaxies in the simulated catalog MICECATv2.

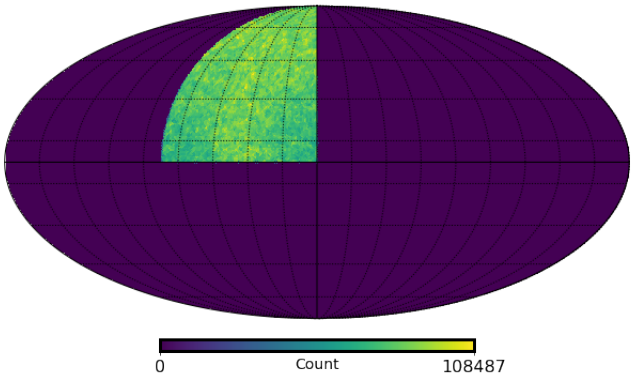

Fig. 3: Molleview projection of bin count per pixel in MICECATv2.

3.2. Simulated GWTC using galaxy catalog

The primary difference between using or not using a galaxy catalog is the underlying spatial distribution of galaxies that GW sources are going to be assigned to. The galaxies in the simulated galaxy catalog contain the spatial clustering, whereas when the galaxies that are chosen without a galaxy catalog, they do not feature spatial clustering. The former can be used to explore the cross-correlation technique whereas, for the second one, one cannot apply cross-correlation.

In the case of using a galaxy catalog the code loads from the catalog the galaxies that are assigned with luminosities in a certain band (chosen by a user). The underlying distribution of the redshifts of galaxies, in this case, is the one from the galaxy catalog, before we apply any merger rate model. The distribution of galaxies in the sky depends on the galaxy catalog. Most of the galaxy catalogs are only covering a small fraction of the sky. In this case, GW events are going to originate from that part only. Whole sky galaxy catalogs, like GLADE+ (Dálya et al. 2022), can be used to avoid this limitation. In the future, this module will be developed further to include different formation channels of GW sources and host galaxy properties. The code currently can not take into account galaxy gatalog incompleteness. This will be added in the future where GWSim will be able to account

for galaxies that host GW events but are not present in the galaxy catalog.

An example of a mock galaxy catalog that we have used within GWSim is the *MICE-Grand Challenge Galaxy and Halo Light-cone Catalog* (MICECATv2) (Fosalba et al. 2014; Crocce et al. 2015; Fosalba et al. 2015). This is an N-body simulation containing about 70 billion dark matter particles in a 3 Gpc/h^3 comoving volume. The dark matter halos and galaxies generated from this span a sky area of 5000 deg^2 and up to a redshift of $z = 1.4$, with no simulation box repetition. The galaxy catalog is made with the following values of the cosmological parameters for a flat-ΛCDM model $\Omega_m = 0.25$; $\sigma_8 = 0.8$; $n_s = 0.95$; $\Omega_b = 0.044$; $\Omega_\Lambda = 0.75$ and $h = 0.7$ [1].

The redshift distribution in MICECATv2 is shown in Fig. 2. For the observed luminosities of the galaxies, there are many different magnitudes, ranging in both color filter and cosmological surveys used to simulate them: DES (Abbott et al. (2021h)), Euclid (Scaramella et al. (2022)), COSMOS (Scoville et al. (2007)), SDSS (Margony (1999)), VISTA (Jarvis et al. (2012)). We apply an evolutionary correction to the apparent magnitudes as suggested in the MICECATv2[2].

$$m_{\rm evol} = m_{\rm true} - 0.8\Big(\arctan\left(1.5\,z\right) - 0.1489\Big) \tag{11}$$

where $m_{\rm true}$ is the observed magnitude in the catalog, z is the true redshift and $m_{\rm evol}$ is the evolved luminosity in redshift. We will use $m_{\rm evol}$ as the observed apparent magnitude, which goes in the luminosity weighting calculation: the user can ask for the probability for a galaxy to host a merger to be proportional to its absolute luminosity. The galaxy catalog also follows the observed luminosity functions, color distributions, and clustering as a function of luminosity and color at low-z.

In the following discussion, we explain how MICECATv2 can be used in GWSim to assign galaxies from the catalog to GWTC events. We start by reading the catalog file and selecting the data columns that we are interested in. We select the sky location in equatorial coordinates (RA and Dec), the true redshift of the galaxies, and the observed magnitude, $m_{\rm evol}$, for a specific band.

Since MICECATv2 is a simulated galaxy catalog, no redshift uncertainty is present. However, GWSim can consider redshift uncertainties by taking a certain percentage (user given) of the true redshift and using this as the 1σ standard deviation. It can also apply masking to the catalog, where it counts the number of galaxies per pixel over the entire sky, and when the number count is less than a certain threshold (user given), it assumes the pixel to be empty. The masking of the catalog over the sky is shown in Fig. 3, as it can be seen, it covers one-eighth of the sky. This masking function will be more useful in future work when real galaxy catalogs will be used, such as GLADE+, as it will introduce a way to measure the angular completeness in the galaxy catalog.

The true redshift of the galaxy is taken to be the mean of the galaxies redshift from MICECATv2 with no redshift errors[3]. Luminosity weighting is applied by using the corrected observed magnitude, $m_{\rm evol}$, and with the true redshift and cosmological parameters used, we retrieve the absolute magnitude, from which we then convert it into luminosity. We also use the fraction of the sky ($f_{\rm sky} = 0.128$) to correct the merger rate of the GW sources as the catalog is covering only a fraction of the sky.

4. Code validation

The validation of the code is done in two steps, (i) we check that the distributions are correctly sampled. This implies we check whether the mock catalog of GW sources matches the expected population of the GW sources. (ii) we produce a mock GW catalog and, by using codes that can infer cosmological parameters from GW data, we try to retrieve the injected cosmological parameters.

(i) *Comparison of the GW mock samples with the expected population:* Starting from the redshift independent mass models, various distributions and corresponding samples can be seen in Fig. A.1. The analytical expressions of the distributions are plotted as dashed orange curves whereas the drawn samples are histogrammed in blue. The comoving volume and the merger rate distributions as functions of redshift can be seen in Fig. B.1, as well as samples drawn from them. The code can also assign spins to events sampling from the models described in Tab. 2. For the results presented from the simulation here, we kept all the spins fixed to zero but an example of sampling of spins for the mass correlated model can be seen in Fig. C.1.

(ii) *Inference of the injected parameters from the mock GW samples:* GWSim computes the number of BBH events happening for certain values of the merger rate model parameters and an observational time selected by the user. Using a uniform in comoving volume distribution as the galaxies redshift distribution, we drew hosts of GW events following the merger rate model with $\gamma = 2.7$, $\kappa = 2.9$, $z_p = 1.9$, $R_0 = 20$ Gpc^{-3} yr^{-1}. We simulated 3 years of O1-, O2-, O3- and O4-like observation run for a realistic duty cycle with the Power Law+Gaussian distribution for m_1 that is shown in Fig. A.1. The duty factors for O1, O2, and O3 were similar to the real ones for the actual LVK observational runs, while for O4 we selected values similar to the ones from O3 which can be seen in Tab. 4.

The total number of mock BBH events is more than 200 thousand (calculated using $z_{max} = 5$). Of those, using a network SNR threshold $\rm SNR_{thr} = 12$, only 164 are detected (remind that the network SNR is the SNR obtained when combining data from all interferometers). The distributions of the detected events masses ($m_{1,s}$, $m_{2,s}$) and distribution of the detected events ($P(z)$) with redshift z can be seen in Fig. 4. In the same plots the distributions of parameters of all the simulated events, both detected (with $\rm SNR > SNR_{thr}$), and all the events are shown.

The detected events that pass the $\rm SNR_{thr}$ undergo parameter estimation using Bilby, to obtain the posteriors of their parameters. In Fig. 5 we present the distribution of the posteriors of six GW source parameters namely individual masses (m_1 and m_2) in detector frame, luminosity distance (D_L), source position (RA and Dec), and the inclination angle (θ_{JN}) for a single event. The true values are indicated by the dashed lines. The distribution of the detector frame masses and distance obtained from the posteriors of all the detected events can be seen in Fig. 6. On the same plots, we see the distribution of the true values of the events. The scatter plots of the masses and distances of the posterior samples of the events in the detector frame can be seen in Fig. 7. The true injected values of the events are also shown as orange stars in the same plots. This plot is not to illustrate that PE works event by event individually, but to show how the samples look and that the posterior samples do cluster around the true injected values.

[1] The catalog is available for download from the *CosmoHub* website (Tallada et al. 2020)

[2] The readme document on MICECATv2 (https://www.dropbox.com/s/0ffa8e7463n8h1q/README_MICECAT_v2.0_for_new_CosmoHub.pdf?dl=0)

[3] However, in the code there is an option to include Gaussian error on the true redshift.

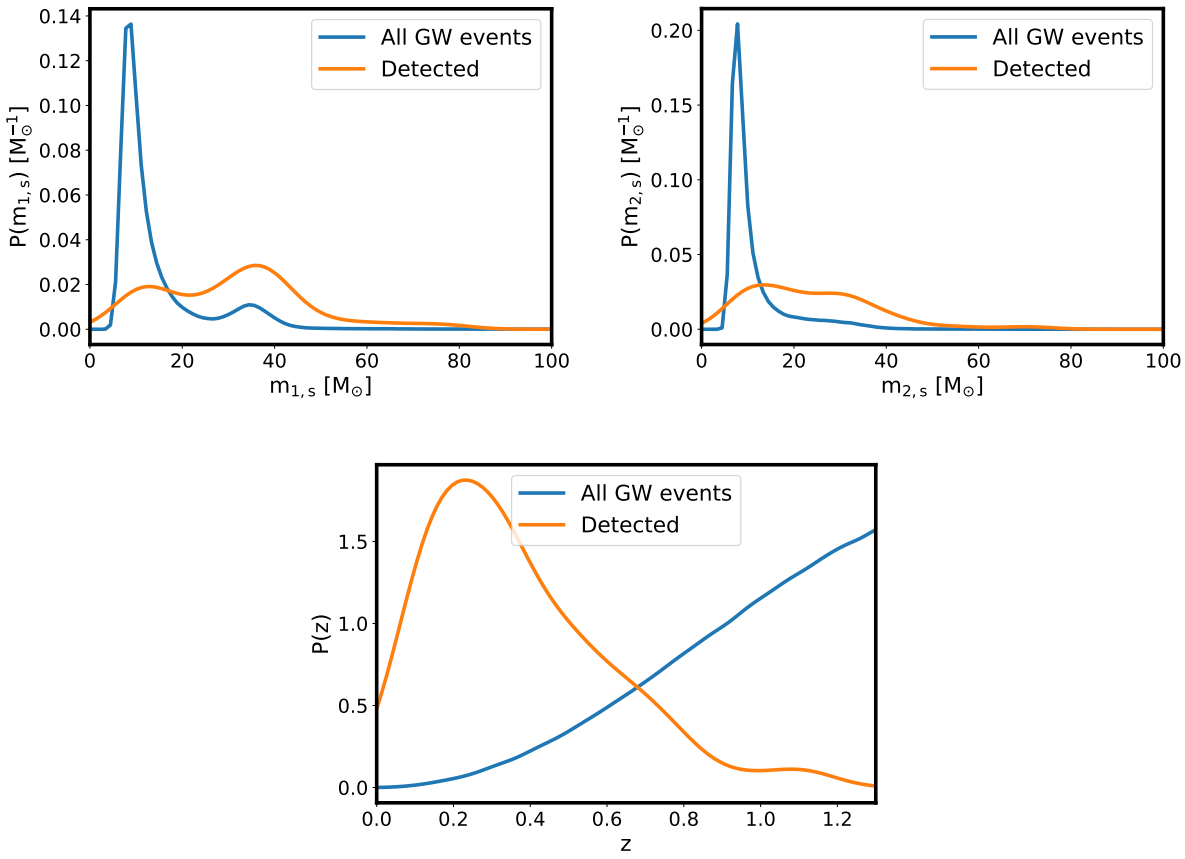

Fig. 4: True values of the primary component masses (top left) and lighter component masses (top right) distributions (analytic expression) in the source frame of the detected events for a simulation of O1-, O2-, O3-, and O4-like 3 observational years. The redshifts (bottom) distribution of the events is also shown. The network SNR threshold was set to $\mathrm{SNR_{thr}} = 12$. In the same plots, the distribution (KDE estimation) of all the simulated (detected and not detected) events can be seen. The population, cosmological, merger rate, and detector parameters used can be seen in Tab. 4.

Having at our disposal a set of posterior samples allows us to use a population and/or cosmology estimation code and retrieve the injected values. We select to run using the python package called `IcaroGW` which can estimate conjointly the population parameters of BBHs and the cosmological parameters (Mastrogiovanni et al. 2021a). For the mass model, we use a Powerlaw+Gaussian peak whereas for the cosmological model a flat-ΛCDM . In `IcaroGW` we fix Ω_m=0.3 and allow only H_0 to vary. The posteriors obtained can be seen in Fig. 8 along with the injected values (black dashed lines). The posterior distribution of the population parameters agrees within 95% C.L. with the injected values for most of the parameters. For a few parameters describing the mass distribution, the injected value falls in the tail of the posterior distribution. This is expected due to the uncertainty in measuring the GW source parameters (such as masses, and distance) which in turn is related to the detector noise of the O1+O2+O3+O4 configuration for the network of three detectors. With more detected events and better detector noise, the difference between the injected value and the maximum likelihood value is likely to reduce. This can be also seen in Fig. 9 where we show the distributions of the true values for m_1 (top) and m_2 (bottom) in the source frame for all detected events with $\mathrm{SNR_{thr}} = 12$. To the leftmost plots we can see the distributions for different injection datasets and for 3 years of total observational time with the same setup as the results from the simulation that was presented previously. As it can be seen the distributions cluster around a common curve, but still significant variation amongst the different curves can be seen, especially for m_2. By raising the observational time to 20 years, we detect more events and the distribution is better defined. This can be in the rightmost plots in the same figure.

5. Conclusions

We present a new package `GWSim` to simulate GW sources for different cosmological and GW population models. This code can incorporate five scenarios of GW mass models, four scenarios of GW spin distributions, two scenarios of the merger rate, and three models of the cosmological parameters, as described in Fig. 1. The code will be useful to simulate realistic mock samples of the GW sources for a different network of GW detectors. Currently, only second-generation detectors are being used for simulations but in the future, next-generation detectors like Cosmic Explorer (Reitze et al. 2019; Hall et al. 2020), and Einstein Telescope (Punturo et al. 2010) will also be included. For a chosen model of the GW source population parameters, and cosmological parameters the code generates samples of binary compact objects for a fixed amount of observation time, duty cycle, and configurations of the detector network. The code can also make use of a simulated galaxy catalog `MICECATv2` for as-

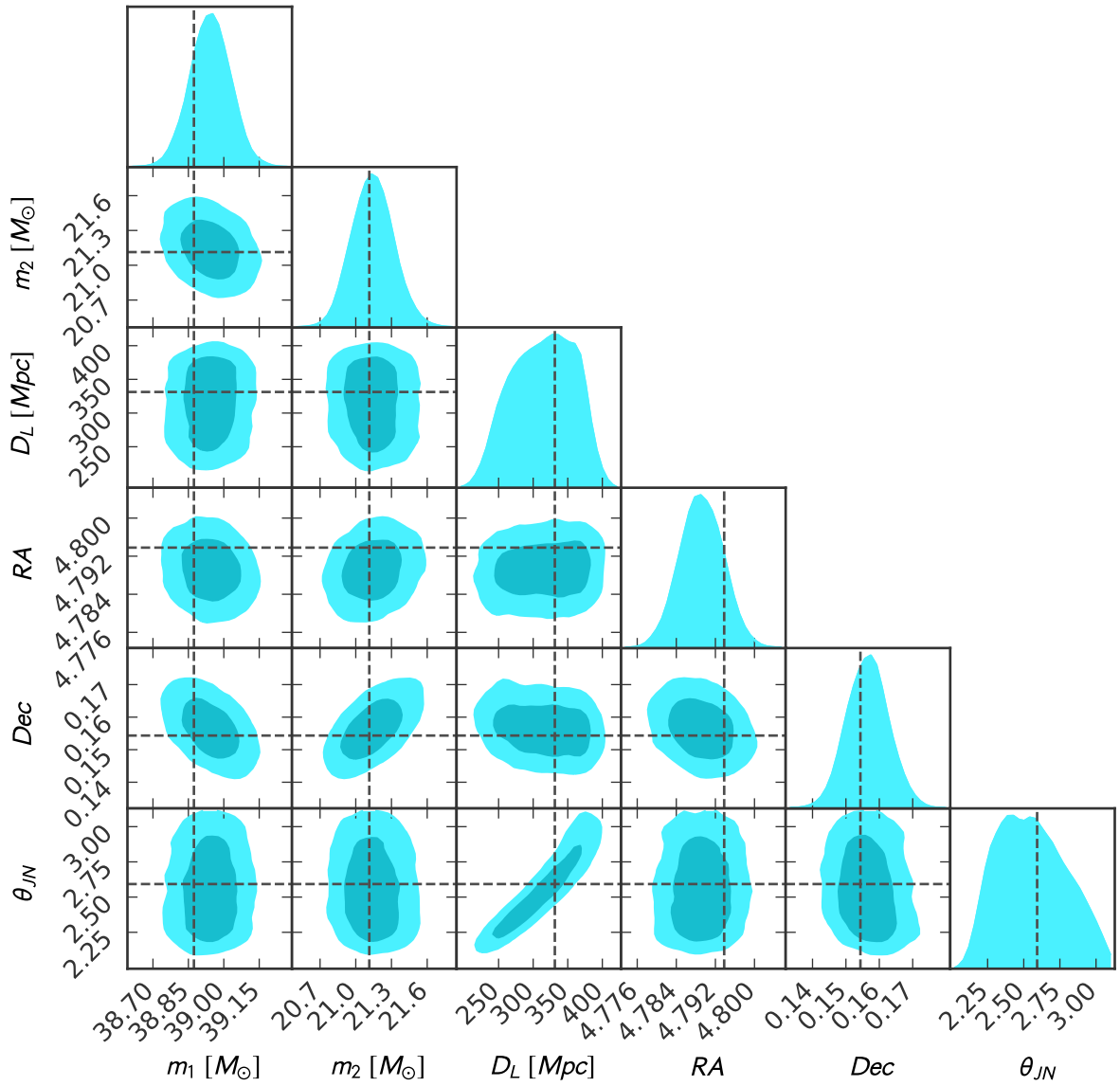

Fig. 5: Parameter estimation posteriors of one randomly selected event using `Bilby`. The parameters from top to bottom are: $m_1, m_2, D_L, Dec, RA, \theta_{JN}$. The latter is the inclination angle between the binary angular momentum and the line of sight from the source to the observer. The injected values for the parameters are presented with dashed black lines.

signing the mock GW samples to the redshift and sky position of the galaxies and making choices about different choices of apparent magnitude to populate the GW sources. In future a revised version will be made to include real galaxy catalog such as GLADE+ (Dálya et al. 2022) and Duncan (2022). The `GWSim` code is also wrapped up with the code `Bilby` for parameter estimation of the GW source parameters. This code will be useful to make simulations of GW sources and understand the interplay between GW source population and cosmology. This code can be used for the currently ongoing network of detectors and also for the next-generation GW detectors.

Acknowledgements

Authors thank Simone Mastrogiovanni for helpful review of the manuscript as a part of LVK P&P. C. Karathanasis is partially supported by the Spanish MINECO under the grants SEV-2016-0588 and PGC2018-101858-B-I00, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. The work of S. Mukherjee is a part of the ⟨`data|theory`⟩ `Universe-Lab` which is supported by the TIFR and the Department of Atomic Energy, Government of India. This analysis is carried out at the computing facility of the LSC cluster. We acknowledge the use of data from the MICE simulations, publicly available at `http://www.ice.cat/mice`. This work has made use of CosmoHub. CosmoHub has been developed by the Port d'Informació Científica (PIC), maintained through a collaboration of the Institut de Física d'Altes Energies (IFAE) and the Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT) and the Institute of Space Sciences (CSIC and IEEC), and was partially funded by the "Plan Estatal de Investigación Científica y Técnica y de Innovación" program of the Spanish government. We acknowledge the use of following packages in this work: Astropy (Astropy Collaboration et al. 2013, 2018), BILBY (Ashton et al. 2019), Giant-

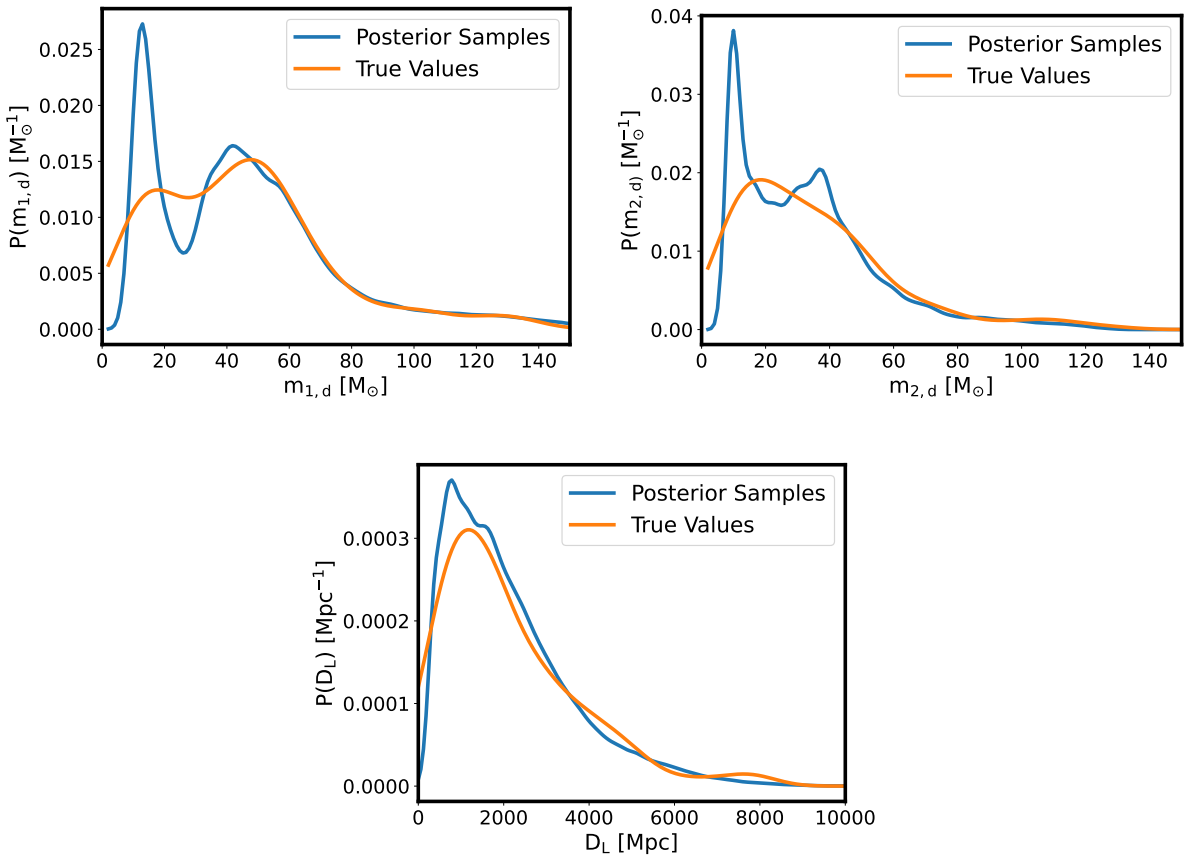

Fig. 6: Primary component masses (top left) and lighter component masses (top right) posterior samples distributions (KDE estimation) in the detector frame for a simulation of O1-, O2-, O3- and O4-like 3 observational years. The luminosity distance (bottom) distribution of the events is also shown. The network SNR threshold was set to $SNR_{thr} = 12$. In the same plots, the distribution (KDE estimation) of all the true values of the events can be seen. The population, cosmological, merger rate, and detector parameters used can be seen in Tab. 4.

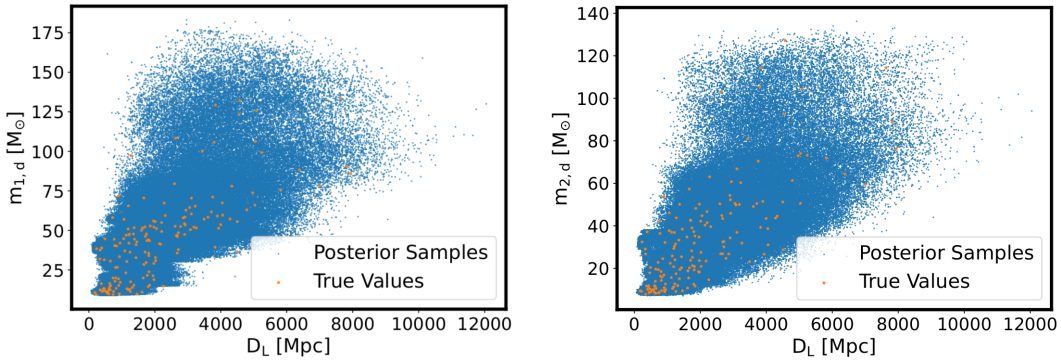

Fig. 7: Primary component masses (left) and lighter component masses (right) posterior samples distributions in the detector frame for a simulation of O1-, O2-, O3- and O4-like 3 observational years, as a function of luminosity distance. The network SNR threshold was set to $SNR_{thr} = 12$. In the same plots, the distribution of all the true values of the events can be seen. Obtained using parameters values of Tab. 4.

Triangle-Confusogram (Bocquet & Carter 2016), IPython (Pérez & Granger 2007), Matplotlib (Hunter 2007), NumPy (van der Walt et al. 2011), and SciPy (Jones et al. 2001–). The authors are grateful for computational resources provided by the LIGO Laboratory and supported by National Science Foundation Grants PHY-0757058 and PHY-0823459. This research has made use of data or software obtained from the Gravitational Wave Open Science Center (gw-openscience.org), a service of LIGO Laboratory, the LIGO Scientific Collaboration, the Virgo Collaboration, and KAGRA. LIGO Laboratory and Advanced LIGO are funded by the United States National Science Foundation (NSF) as well as the Science and Technology Facilities Council (STFC) of the

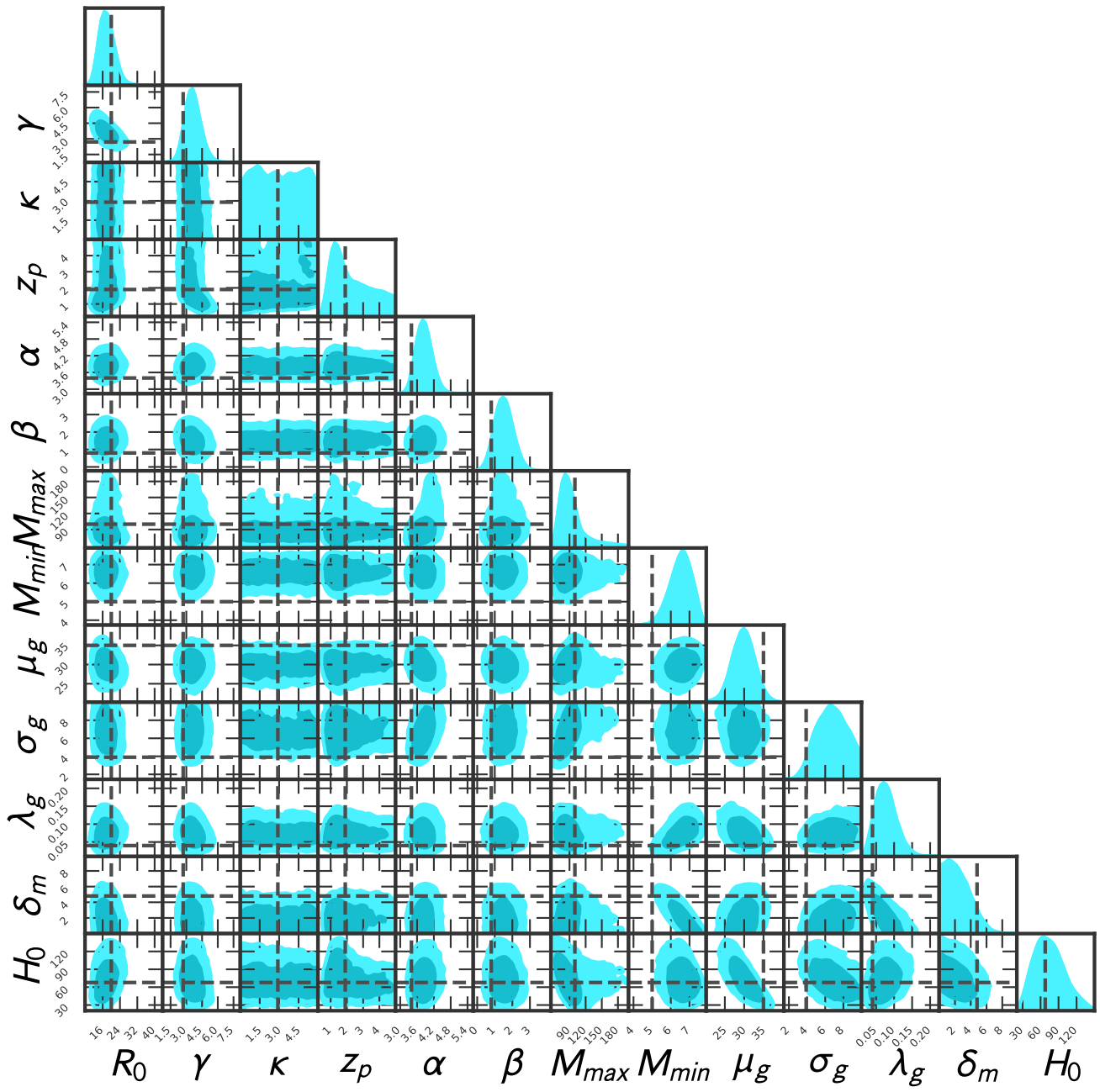

Fig. 8: Population and cosmological parameters posteriors obtained from an O1-, O2-, O3-and O4-like simulation using the cosmological inference python package `IcaroGW`. The dashed black lines indicate the injected values of parameters.

United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, Spain. The construction and operation of KAGRA are funded by Ministry of Education, Culture, Sports, Science and Technology (MEXT), and Japan Society for the Promotion of Science (JSPS), National Research Foundation (NRF) and Ministry of Science and ICT (MSIT) in Korea, Academia Sinica (AS) and the Ministry of Science and Technology (MoST) in Taiwan. This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation.

References

Aasi, J. et al. 2015, Class. Quant. Grav., 32, 074001
Abbott, B., Abbott, R., Abbott, T., et al. 2019a, Physical Review X, 9
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016a, Phys. Rev. Lett., 116, 131103
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2021a, The Astrophysical Journal, 909, 218

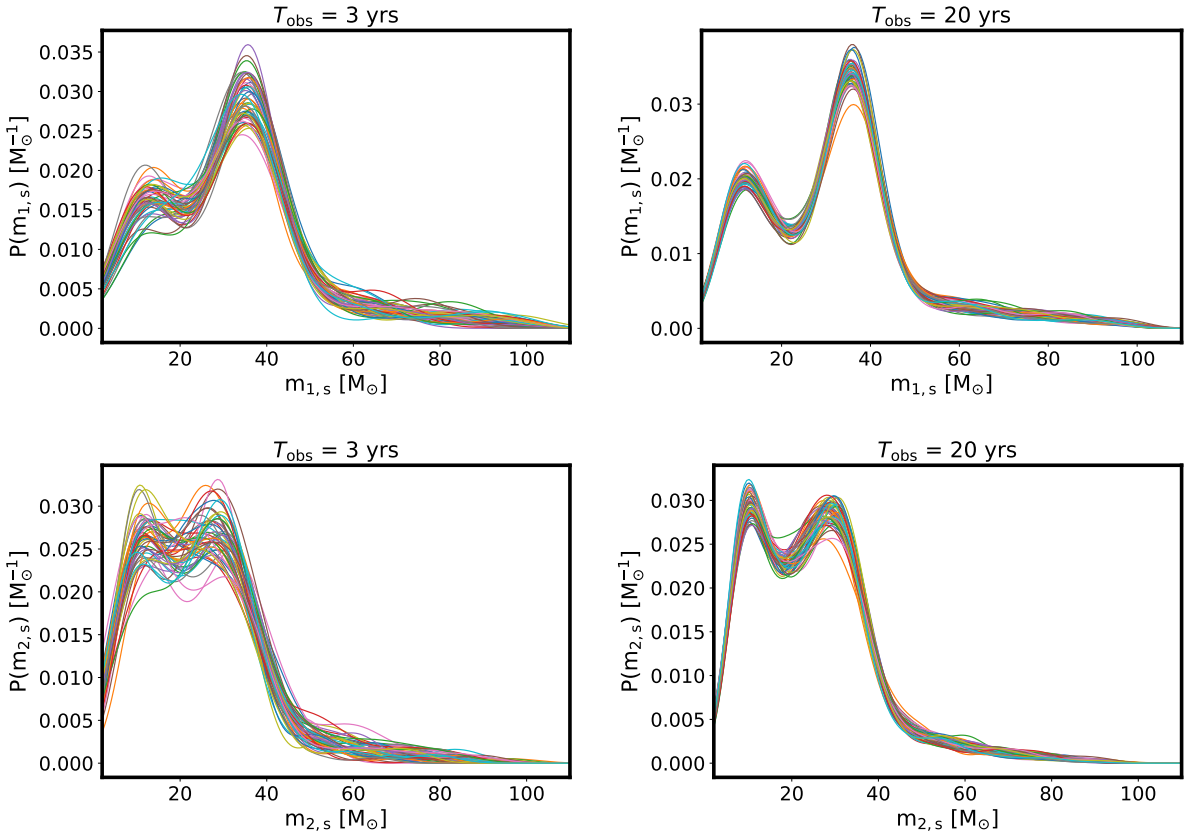

Fig. 9: Various m_1 (top) and m_2 (bottom) distributions of the detected events with SNR > 12 for T_{obs} = 3 years (left) and T_{obs} = 20 years (right) produced for different random seeds. Raising the total observational time leads to more events being detected and the distributions being better defined. The population, cosmological, merger rate, and detector parameters used can be seen in Tab. 4.

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2019b, The Astrophysical Journal, 882, L24
Abbott, B. P. et al. 2016b, Phys. Rev. D, 93, 112004, [Addendum: Phys.Rev.D 97, 059901 (2018)]
Abbott, B. P. et al. 2018, Living Rev. Rel., 21, 3
Abbott, R., Abbott, T., Abraham, S., et al. 2021b, Physical Review X, 11
Abbott, R., Abbott, T., Abraham, S., et al. 2020a, arXiv preprint arXiv:2010.14533
Abbott, R., Abbott, T. D., Acernese, F., et al. 2021c, GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo During the Second Part of the Third Observing Run
Abbott, R., Abbott, T. D., Acernese, F., et al. 2021d, The population of merging compact binaries inferred using gravitational waves through GWTC-3
Abbott, R., Abbott, T. D., Acernese, F., et al. 2021e, arXiv: 2108.01045
Abbott, R., Abe, H., Acernese, F., et al. 2021f, Constraints on the cosmic expansion history from GWTC-3
Abbott, R. et al. 2020b, Astrophys. J., 896, L44
Abbott, R. et al. 2021g, arXiv: 2111.03634 [arXiv:2111.03634]
Abbott, T. M. C., Adamó w, M., Aguena, M., et al. 2021h, The Astrophysical Journal Supplement Series, 255, 20
Acernese, F., Agathos, M., Agatsuma, K., et al. 2014, Classical and Quantum Gravity, 32, 024001
Akutsu, T. et al. 2020, arXiv:2005.05574 [arXiv:2005.05574]
Arabsalmani, M., Sahni, V., & Saini, T. D. 2013, Phys. Rev. D, 87, 083001
Ashton, G., Hübner, M., Lasky, P. D., et al. 2019, The Astrophysical Journal Supplement Series, 241, 27
Aso, Y., Michimura, Y., Somiya, K., et al. 2013, Phys. Rev. D, 88, 043007
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Banerjee, S., Baumgardt, H., & Kroupa, P. 2010, MNRAS, 402, 371
Bavera, S. S., Fragos, T., Qin, Y., et al. 2020, Astron. Astrophys., 635, A97
Bera, S., Rana, D., More, S., & Bose, S. 2020, Astrophys. J., 902, 79
Biscoveanu, S., Callister, T. A., Haster, C.-J., et al. 2022, Astrophys. J. Lett., 932, L19
Biscoveanu, S., Isi, M., Varma, V., & Vitale, S. 2021a, Phys. Rev. D, 104, 103018
Biscoveanu, S., Isi, M., Vitale, S., & Varma, V. 2021b, Phys. Rev. Lett., 126, 171103
Bocquet, S. & Carter, F. W. 2016, The Journal of Open Source Software, 1
Cañas-Herrera, G., Contigiani, O., & Vardanyan, V. 2021, ApJ, 918, 20
Callister, T. A., Haster, C.-J., Ng, K. K. Y., Vitale, S., & Farr, W. M. 2021, Astrophys. J. Lett., 922, L5
Calore, F., Cuoco, A., Regimbau, T., Sachdev, S., & Serpico, P. D. 2020, Physical Review Research, 2
Cao, L., Lu, Y., & Zhao, Y. 2018, MNRAS, 474, 4997
Chevallier, M. & Polarski, D. 2001, Int. J. Mod. Phys. D, 10, 213
Cigarrán Díaz, C. & Mukherjee, S. 2022, MNRAS, 511, 2782
Crocce, M., Castander, F. J., Gaztañ aga, E., Fosalba, P., & Carretero, J. 2015, Monthly Notices of the Royal Astronomical Society, 453, 1513
Del Pozzo, W. 2012, Phys. Rev. D, 86, 043011
Dominik, M., Belczynski, K., Fryer, C., et al. 2012, ApJ, 759, 52
Dominik, M., Berti, E., O'Shaughnessy, R., et al. 2015, Astrophys. J., 806, 263
du Buisson, L., Marchant, P., Podsiadlowski, P., et al. 2020, Mon. Not. Roy. Astron. Soc., 499, 5941
Duncan, K. J. 2022, Mon. Not. Roy. Astron. Soc., 512, 3662
Dálya, G., Díaz, R., Bouchet, F. R., et al. 2022, Monthly Notices of the Royal Astronomical Society, 514, 1403
Elbert, O. D., Bullock, J. S., & Kaplinghat, M. 2018, Mon. Not. Roy. Astron. Soc., 473, 1186
Eldridge, J. J., Stanway, E. R., & Tang, P. N. 2019, Mon. Not. Roy. Astron. Soc., 482, 870
Ezquiaga, J. M. & Holz, D. E. 2022, arXiv:2202.08240 [arXiv:2202.08240]
Farmer, R., Renzo, M., de Mink, S. E., Marchant, P., & Justham, S. 2019, ApJ, 887, 53
Farr, W. M., Fishbach, M., Ye, J., & Holz, D. 2019, Astrophys. J. Lett., 883, L42
Farr, W. M., Stevenson, S., Coleman Miller, M., et al. 2017, Nature, 548, 426
Finke, A., Foffa, S., Iacovelli, F., Maggiore, M., & Mancarella, M. 2021, arXiv:2101.12660 [arXiv:2101.12660]
Fishbach, M. & Kalogera, V. 2021, Astrophys. J. Lett., 914, L30
Fishbach, M. et al. 2019, Astrophys. J. Lett., 871, L13

Fosalba, P., Crocce, M., Gaztañaga, E., & Castander, F. J. 2015, MNRAS, 448, 2987
Fosalba, P., Gaztañ aga, E., Castander, F. J., & Crocce, M. 2014, Monthly Notices of the Royal Astronomical Society, 447, 1319
Fuller, J. & Ma, L. 2019, Astrophys. J. Lett., 881, L1
Gerosa, D., Berti, E., O'Shaughnessy, R., et al. 2018, Phys. Rev. D, 98, 084036
Gray, R., Hernandez, I. M. n., Qi, H., et al. 2020, Phys. Rev. D, 101, 122001
Gregory, M. H. e. a. 2010, Classical and Quantum Gravity, 27, 084006
Hall, A., Gow, A. D., & Byrnes, C. T. 2020, Phys. Rev. D, 102, 123524
Hunter, J. D. 2007, Computing In Science & Engineering, 9, 90
Jarvis, M. J., Bonfield, D. G., Bruce, V. A., et al. 2012, Monthly Notices of the Royal Astronomical Society, 428, 1281
Jones, E., Oliphant, T., Peterson, P., et al. 2001–, SciPy: Open source scientific tools for Python, [Online; accessed <today>]
Karathanasis, C., Mukherjee, S., & Mastrogiovanni, S. 2022, Binary black holes population and cosmology in new lights: Signature of PISN mass and formation channel in GWTC-3
Kushnir, D., Zaldarriaga, M., Kollmeier, J. A., & Waldman, R. 2016, Mon. Not. Roy. Astron. Soc., 462, 844
Lamberts, A., Garrison-Kimmel, S., Clausen, D. R., & Hopkins, P. F. 2016, Mon. Not. Roy. Astron. Soc., 463, L31
Leyde, K., Mastrogiovanni, S., Steer, D. A., Chassande-Mottin, E., & Karathanasis, C. 2022, arXiv:2202.00025 [arXiv:2202.00025]
LIGO, Virgo, & Kagra. 2021, arXiv: 2111.03604 [arXiv:2111.03604]
Linder, E. V. 2003, Phys. Rev. Lett., 90, 091301
MacLeod, C. L. & Hogan, C. J. 2008, Phys. Rev. D, 77, 043512
Madau, P. & Dickinson, M. 2014, Annual Review of Astronomy and Astrophysics, 52, 415–486
Mancarella, M., Genoud-Prachex, E., & Maggiore, M. 2021, arXiv: 2112.05728 [arXiv:2112.05728]
Mandel, I. & de Mink, S. E. 2016, MNRAS, 458, 2634
Mandel, I. & O'Shaughnessy, R. 2010, Class. Quant. Grav., 27, 114007
Margony, B. 1999, Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 357, 93
Mastrogiovanni, S., Leyde, K., Karathanasis, C., et al. 2021a, Physical Review D, 104
Mastrogiovanni, S., Leyde, K., Karathanasis, C., et al. 2021b, arXiv:2103.14663 [arXiv:2103.14663]
Mukherjee, S. 2022, MNRAS[arXiv:2112.10256]
Mukherjee, S. & Dizgah, A. M. 2021, arXiv:2111.13166 [arXiv:2111.13166]
Mukherjee, S., Krolewski, A., Wandelt, B. D., & Silk, J. 2022, arXiv:2203.03643 [arXiv:2203.03643]
Mukherjee, S. & Silk, J. 2021, Mon. Not. Roy. Astron. Soc., 506, 3977
Mukherjee, S. & Wandelt, B. D. 2018, arXiv:1808.06615 [arXiv:1808.06615]
Mukherjee, S., Wandelt, B. D., Nissanke, S. M., & Silvestri, A. 2021a, Phys. Rev. D, 103, 043520
Mukherjee, S., Wandelt, B. D., & Silk, J. 2020, Mon. Not. Roy. Astron. Soc., 494, 1956
Mukherjee, S., Wandelt, B. D., & Silk, J. 2021b, Mon. Not. Roy. Astron. Soc., 502, 1136
Oguri, M. 2016, Phys. Rev. D, 93, 083511
O'Shaughnessy, R., Gerosa, D., & Wysocki, D. 2017, Phys. Rev. Lett., 119, 011101
O'Shaughnessy, R., Kalogera, V., & Belczynski, K. 2010, ApJ, 716, 615
Palmese, A., Bom, C. R., Mucesh, S., & Hartley, W. G. 2021, arXiv: 2111.06445 [arXiv:2111.06445]
Pérez, F. & Granger, B. E. 2007, Computing in Science and Engineering, 9, 21
Punturo, M. et al. 2010, Class. Quant. Grav., 27, 194002
Reitze, D. et al. 2019, Bull. Am. Astron. Soc., 51, 035
Renzo, M., Farmer, R. J., Justham, S., et al. 2020, Mon. Not. Roy. Astron. Soc., 493, 4333
Rodriguez, C. L., Zevin, M., Pankow, C., Kalogera, V., & Rasio, F. A. 2016, Astrophys. J. Lett., 832, L2
Santoliquido, F., Mapelli, M., Giacobbo, N., Bouffanais, Y., & Artale, M. C. 2021, Mon. Not. Roy. Astron. Soc., 502, 4877
Scaramella, R., Amiaux, J., Mellier, Y., et al. 2022, Astronomy & Astrophysics, 662, A112
Scelfo, G., Boco, L., Lapi, A., & Viel, M. 2020, JCAP, 10, 045
Scelfo, G., Spinelli, M., Raccanelli, A., et al. 2022, JCAP, 01, 004
Schutz, B. F. 1986, Nature, 323, 310
Scoville, N., Aussel, H., Brusa, M., et al. 2007, The Astrophysical Journal Supplement Series, 172, 1
Soares-Santos, M. et al. 2019, Astrophys. J. Lett., 876, L7
Spera, M. & Mapelli, M. 2017, Mon. Not. Roy. Astron. Soc., 470, 4739
Stevenson, S., Berry, C. P. L., & Mandel, I. 2017, Mon. Not. Roy. Astron. Soc., 471, 2801
Talbot, C. & Thrane, E. 2018, The Astrophysical Journal, 856, 173
Tallada, P., Carretero, J., Casals, J., et al. 2020, CosmoHub: Interactive exploration and distribution of astronomical data on Hadoop
Taylor, S. R., Gair, J. R., & Mandel, I. 2012, Phys. Rev. D, 85, 023535
Tse, M., Yu, H., Kijbunchoo, N., et al. 2019, Phys. Rev. Lett., 123, 231107
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science and Engineering, 13, 22
Vitale, S., Farr, W. M., Ng, K., & Rodriguez, C. L. 2019, Astrophys. J. Lett., 886, L1
You, Z.-Q., Zhu, X.-J., Ashton, G., Thrane, E., & Zhu, Z.-H. 2021, Astrophys. J., 908, 215

Appendix A: Mass distributions

Here we show some of the available mass distributions and showcase the ability of the code to properly sample those. In Fig. A.1 the histogram of the samples drawn from the input distributions is shown in blue and the analytical curves are shown as orange dashed curves.

Appendix B: Merger rate distributions

Here we show some of the available merger rate distributions and showcase the ability of the code to properly sample those distributions. In Fig. B.1 the samples drawn from the distributions are histogrammed in blue and the analytical curves are shown as orange dashed curves.

Appendix C: Spin distributions

Here we show a spin sampling of the mass correlated model. In Fig. C.1 one can see the samples that the code drew in the $\chi_{\rm eff} - q$ space, $\chi_1 - q$ space and $\chi_2 - q$ space. Fig. C.2 shows the individual spin amplitudes χ_1 and χ_2 as a function of z in the case of the spin-redshift model where both expectation value and standard deviation depend linearly of z (see Sec. 2.4.2).

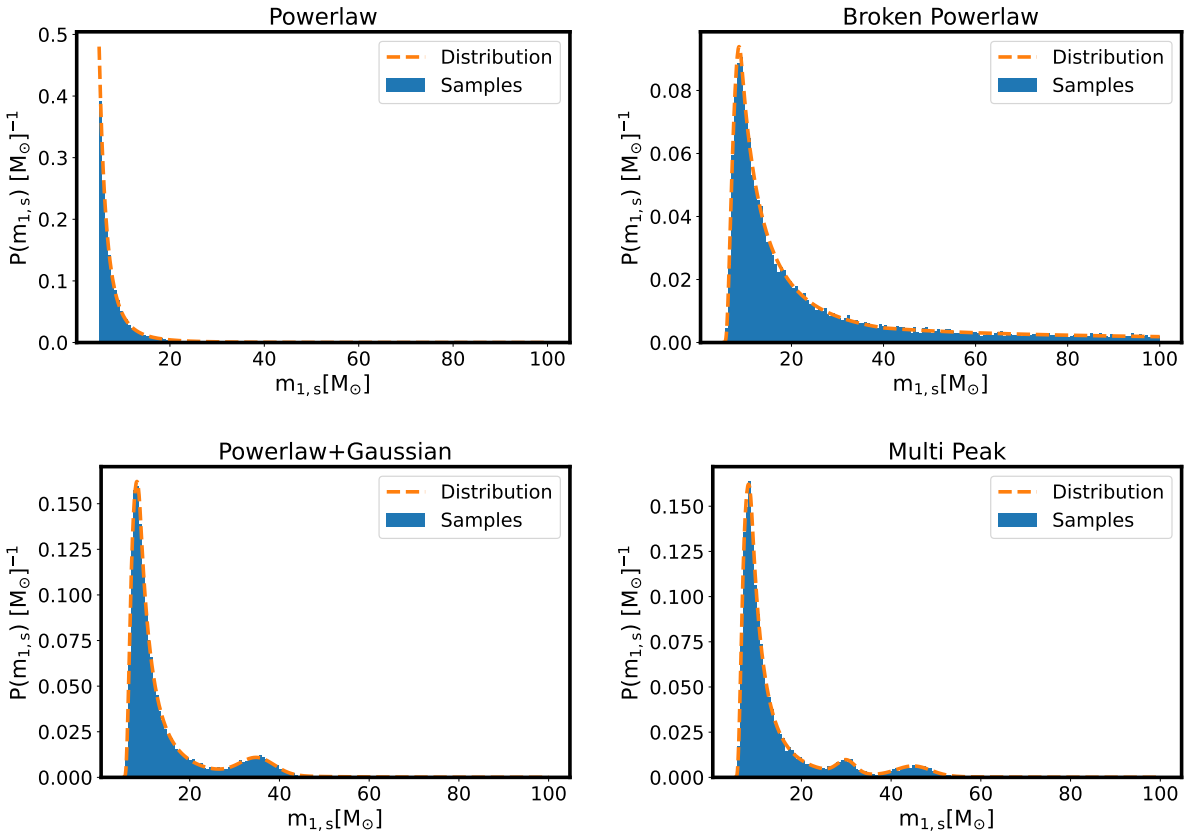

Fig. A.1: Empirical distributions (in blue) of the source-frame mass samples drawn from various mass distributions. The analytical expressions of the distributions are plotted in orange dashed curves. Top left: a Truncated Power Law model. Top right: a Broken Power Law model with smoothing at the lower end of the distribution. Bottom left: a Power Law+Gaussian peak model with smoothing at the lower end of the distribution. Bottom right: a Multi-Peak model with smoothing at the lower end of the distribution.

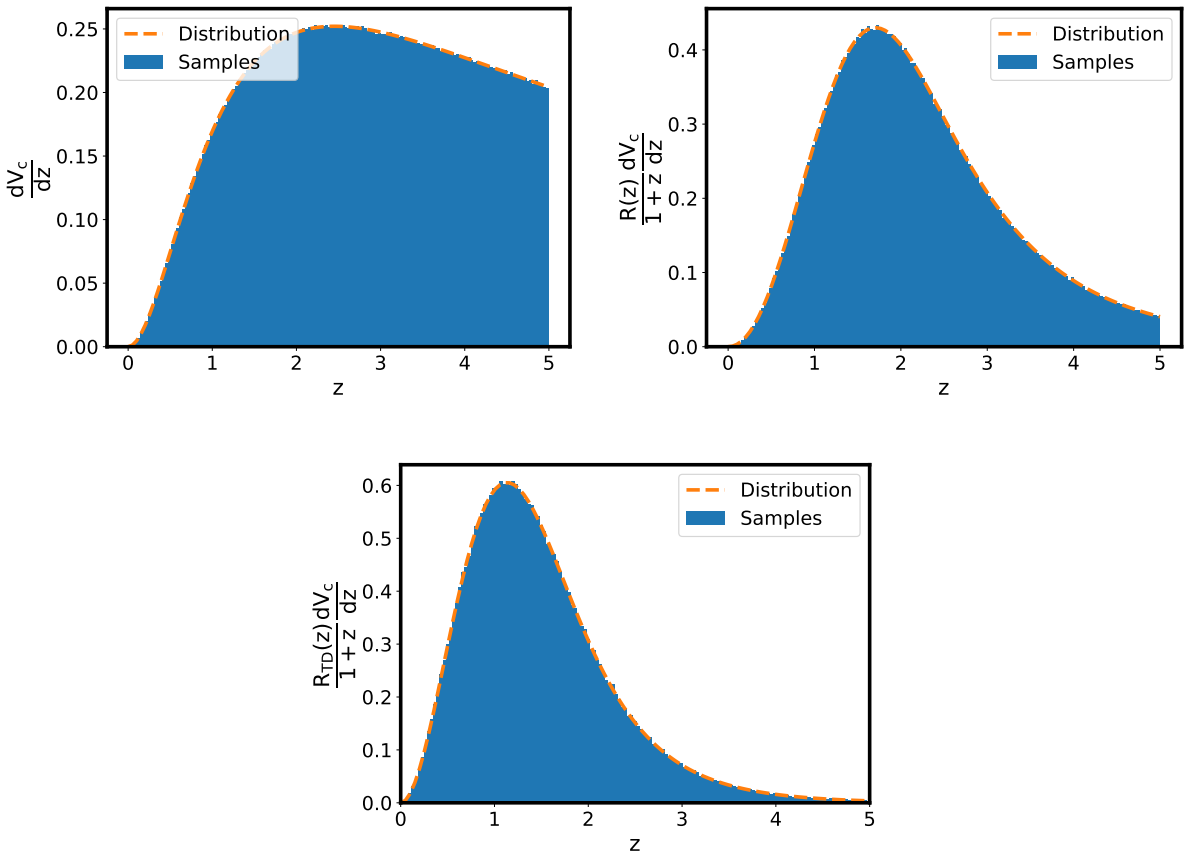

Fig. B.1: Empirical distribution of the redshift samples draws (in blue) from the uniform in comoving volume distribution (top left), the merger rate model of Eq. 2 (top right) for $z_p = 1.9, \gamma = 2.7, \kappa = 2.9$ and the merger rate model in the case of time delay model described by Eq. 3 (bottom) for $d = -1, t_d^{min} = 0.5$ Gyrs. The analytical expressions of the distributions are plotted in orange dashed curves. Obtained using parameters values of Tab. 4.

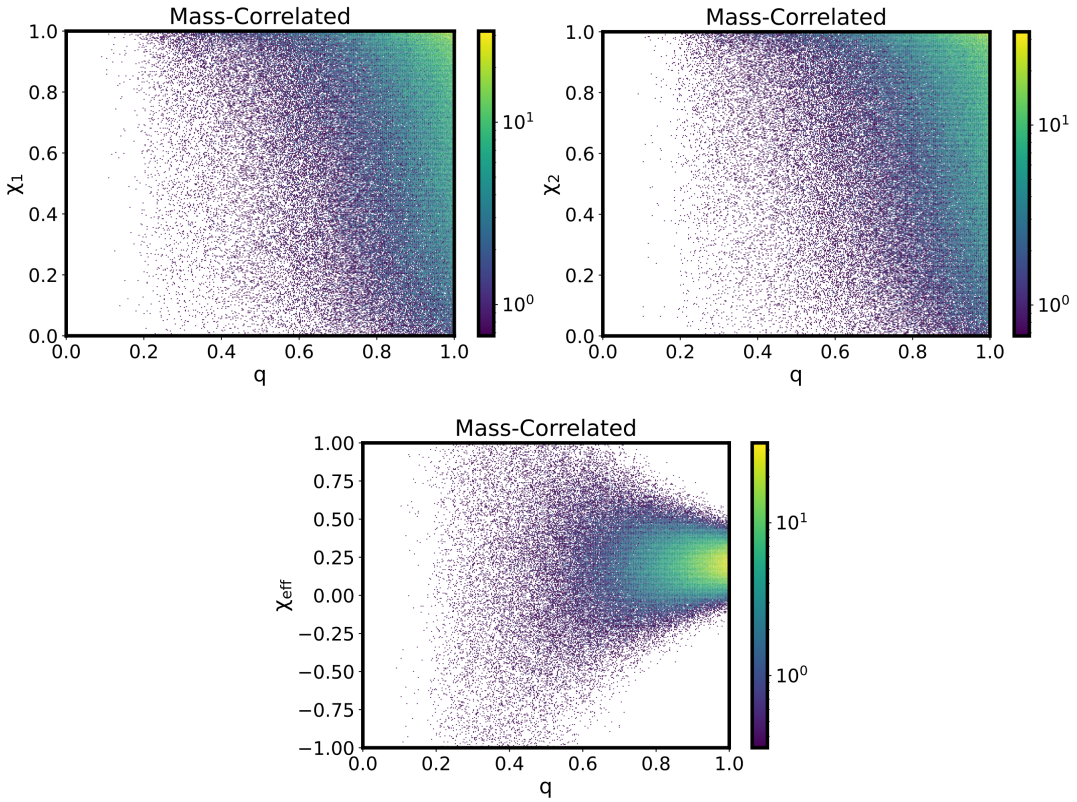

Fig. C.1: The 2D distribution of the spin sampling for the case of the mass correlated model. Top figures: individual spin amplitudes χ_1 (left) and χ_2 (right) as a function of the mass ratio q. Bottom: effective spin amplitude $\chi_{\rm eff}$ as a function of the mass ratio q. The distributions were calculated with $\alpha_s = 0.013, \beta_s = -1.53, \chi_0 = 0.2, \sigma_0 = 0.52$. The mass distribution was the Powerlaw+Gaussian distribution seen in A.1.

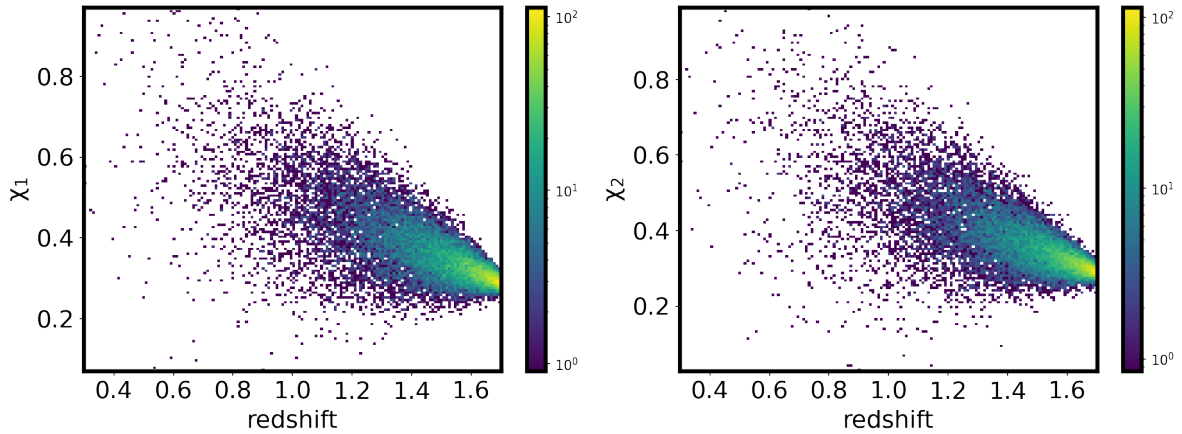

Fig. C.2: The 2D distribution of the sampled spin amplitudes χ_1 (left) and χ_2 (right) for the case of the spin model evolving with the redshift. The parameters used are $\chi_0 = 0.8$, $\delta\chi = -0.3$, $\sigma_0 = 0.3$, $\delta\sigma = -0.17$, see section 2.4.2 for the spin-redshift dependence.