Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **507,** 3362–3380 (2021) https://doi.org/10.1093/mnras/stab2315
Advance Access publication 2021 August 13

Signatures of hierarchical mergers in black hole spin and mass distribution

Hiromichi Tagawa,[1]★ Zoltán Haiman,[2] Imre Bartos,[3] Bence Kocsis [4] and Kazuyuki Omukai[1]

[1]*Astronomical Institute, Graduate School of Science, Tohoku University, Aoba, Sendai 980-8578, Japan*
[2]*Department of Astronomy, Columbia University, 550 W. 120th St., New York NY-10027, USA*
[3]*Department of Physics, University of Florida, PO Box 118440, Gainesville FL-32611, USA*
[4]*Rudolf Peierls Centre for Theoretical Physics, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK*

Accepted 2021 August 5. Received 2021 August 5; in original form 2021 April 21

ABSTRACT

Recent gravitational wave (GW) observations by LIGO/Virgo show evidence for hierarchical mergers, where the merging BHs are the remnants of previous BH merger events. These events may carry important clues about the astrophysical host environments of the GW sources. In this paper, we present the distributions of the effective spin parameter ($\chi_{\rm eff}$), the precession spin parameter ($\chi_{\rm p}$), and the chirp mass ($m_{\rm chirp}$) expected in hierarchical mergers. Under a wide range of assumptions, hierarchical mergers produce (i) a monotonic increase of the average of the typical total spin for merging binaries, which we characterize with $\bar{\chi}_{\rm typ}\equiv\overline{(\chi_{\rm eff}^2+\chi_{\rm p}^2)^{1/2}}$, up to roughly the maximum $m_{\rm chirp}$ among first-generation (1g) BHs, and (ii) a plateau at $\bar{\chi}_{\rm typ}\sim 0.6$ at higher $m_{\rm chirp}$. We suggest that the maximum mass and typical spin magnitudes for 1g BHs can be estimated from $\bar{\chi}_{\rm typ}$ as a function of $m_{\rm chirp}$. The GW data observed in LIGO/Virgo O1–O3a prefers an increase in $\bar{\chi}_{\rm typ}$ at low $m_{\rm chirp}$, which is consistent with the growth of the BH spin magnitude by hierarchical mergers at $\sim 2\sigma$ confidence. A Bayesian analysis using the $\chi_{\rm eff}$, $\chi_{\rm p}$, and $m_{\rm chirp}$ distributions suggests that 1g BHs have the maximum mass of $\sim$15–30 $\rm M_\odot$ if the majority of mergers are of high-generation BHs (not among 1g–1g BHs), which is consistent with mergers in active galactic nucleus discs and/or nuclear star clusters, while if mergers mainly originate from globular clusters, 1g BHs are favoured to have non-zero spin magnitudes of $\sim$0.3. We also forecast that signatures for hierarchical mergers in the $\bar{\chi}_{\rm typ}$ distribution can be confidently recovered once the number of GW events increases to $\gtrsim O(100)$.

Key words: black hole physics – gravitational waves – methods: data analysis.

1 INTRODUCTION

Recent detections of gravitational waves (GWs) by LIGO (Aasi et al. 2015) and Virgo (Acernese et al. 2015) have shown evidence for a high rate of black hole (BH)–BH and neutron star (NS)–NS mergers in the Universe (Abbott et al. 2019, 2021; Venumadhav et al. 2020). However, proposed astrophysical pathways to mergers remain highly debated. Indeed there are currently a large number of such possible pathways, with widely different environments and physical processes. A possible list of these currently includes isolated binary evolution (e.g. Dominik et al. 2012; Kinugawa et al. 2014; Belczynski et al. 2016; Spera et al. 2019) accompanied by mass transfer (Inayoshi et al. 2017; Pavlovskii et al. 2017; van den Heuvel, Portegies Zwart & de Mink 2017), common envelope ejection (e.g. Paczynski 1976; Ivanova et al. 2013), envelope expansion (Tagawa, Kocsis & Saitoh 2018), chemically homogeneous evolution in a tidally distorted binary (de Mink & Mandel 2016; Mandel & de Mink 2016; Marchant et al. 2016), evolution of triple or quadruple systems (e.g. Antonini, Toonen & Hamers 2017; Silsbee & Tremaine 2017; Fragione & Kocsis 2019; Michaely & Perets 2019), gravitational capture (e.g. O'Leary, Kocsis & Loeb 2009; Gondán et al. 2018; Rasskazov & Kocsis 2019), dynamical evolution in open clusters (e.g. Banerjee 2017; Kumamoto, Fujii & Tanikawa 2018; Rastello et al. 2019), young stellar clusters (e.g. Ziosi et al. 2014; Di Carlo et al. 2019; Rastello et al. 2020), and dense star clusters (e.g. Portegies Zwart & McMillan 2000; Samsing, MacLeod & Ramirez-Ruiz 2014; O'Leary, Meiron & Kocsis 2016; Rodriguez, Chatterjee & Rasio 2016; Fragione & Kocsis 2018; Fragione et al. 2019), and interaction in active phases of galactic nucleus discs (e.g. Bartos et al. 2017; Stone, Metzger & Haiman 2017; McKernan et al. 2018; Tagawa, Haiman & Kocsis 2020).

Recently several GW events were reported by LIGO and Virgo whose measured physical properties pose interesting constraints on their astrophysical origin. These include nine candidates for mergers in the upper-mass gap ($\sim$50–130 $\rm M_\odot$), such as GW190521 (Abbott et al. 2019; Zackay et al. 2019; Abbott et al. 2020a, 2021; The LIGO Scientific Collaboration 2020a). Additionally, mergers with very unequal masses have been reported – GW190412 ($q=0.28^{+0.13}_{-0.07}$, The LIGO Scientific Collaboration 2020b) and GW190814 ($q=0.112^{+0.008}_{-0.009}$, Abbott et al. 2020b) – which are also atypical in stellar evolutionary models of isolated binaries (Gerosa, Vitale & Berti 2020; Olejak et al. 2020; Zevin et al. 2020b). The object in the lower mass gap in GW190814 and a non-zero spin for the primary BH ($a_1=0.43^{+0.16}_{-0.26}$) in GW190412 are consistent with a scenario in which the merging compact objects (COs) had experienced previous episode(s) of mergers or significant accretion. These events suggest that growth by gas accretion or hierarchical mergers may be common

★ E-mail: htagawa@astr.tohoku.ac.jp

among COs (see e.g. O'Leary et al. 2016; Abbott et al. 2020a; Fragione, Loeb & Rasio 2020; Gerosa et al. 2020; Hamers & Safarzadeh 2020; Liu & Lai 2021; Rodriguez et al. 2020; Safarzadeh & Haiman 2020; Safarzadeh et al. 2020; Samsing et al. 2020; Yang et al. 2020b; Tagawa et al. 2021a,b).

Hierarchical mergers may occur in dynamical environments, such as globular clusters (GCs), nuclear star clusters (NSCs), and active galactic nucleus (AGN) accretion discs. In GCs, up to $\sim$ 10–20 per cent of detected mergers may be caused by high-generation (high-g) BHs depending on spin magnitudes of 1g BHs (O'Leary et al. 2016; Rodriguez et al. 2019). Repeated mergers of BHs and stars may produce intermediate-mass BHs (BHs with masses of $\sim$100–$10^4\,M_\odot$) in NSCs without supermassive BHs (SMBHs; Antonini, Gieles & Gualandris 2019; Askar, Davies & Church 2021; Mapelli et al. 2020). In NSCs with SMBHs, it is uncertain how often hierarchical mergers occur (e.g. Arca Sedda 2020).

In AGN discs, hierarchical mergers are predicted to be frequent due to the high escape velocity and efficient binary formation and evolution facilitated by gaseous (Yang et al. 2019; McKernan et al. 2020) and stellar interactions (Tagawa et al. 2020). Yang et al. (2019) and McKernan et al. (2020), McKernan, Ford & O'Shaughnessy (2020) identified the expected mass ratio and spin distribution of hierarchical mergers in hypothetical migration traps (MTs) of AGN discs, defined to be regions where objects accumulate rapidly as they interact with the accretion discs analogously to planetary migration.[1] Tagawa et al. (2020, 2021b) showed that hierarchical mergers take place in AGN discs without MTs and derived the corresponding mass and spin distributions self-consistently. In the latter models (e.g. Tagawa et al. 2020), the mass and spin distributions of merging BHs are significantly different compared to those in the former models. This is mainly due to binary–single interactions which take place frequently at large orbital radii where the gas density is very low and gas effects drive the binaries toward merger more slowly and allow ample time for such binary–single interactions.

Several authors have investigated the properties of GWs associated with hierarchical mergers (Gerosa & Berti 2017; Yang et al. 2019; Doctor et al. 2020; Kimball et al. 2020). Gerosa & Berti (2017) estimated the fraction of future detected sources contributed by hierarchical mergers under the assumption that first-generation (1g) BHs have a flat spin distribution and binary components are drawn independently. Fishbach, Holz & Farr (2017) estimated the required number of events to detect hierarchical mergers using the distribution of the BH spin magnitudes. Doctor et al. (2020) constructed a toy model to obtain the properties of hierarchical mergers from the distribution of sub-populations for BHs under various assumptions for coagulation and depletion in the population and constrained parameters using LIGO/Virgo O1–O2 data. Kimball et al. (2020) examined whether the observed events in the same catalogue are compatible with hierarchical mergers particularly in GCs. These models found no evidence for a high rate of hierarchical mergers in this early catalogue. More recently, by analysing the ensemble of events detected during LIGO/Virgo's O1-O3a observing runs, Kimball et al. (2021) and Tiwari & Fairhurst (2021) found preference for at least one, but probably multiple hierarchical mergers in the detected sample. The conclusion of Kimball et al. (2021) strongly depends on the assumed escape velocity in the host environment, with higher escape velocities favouring a larger number of hierarchical mergers.

In this paper, we focus on distributions of the effective and precession spin parameters ($\chi_{\rm eff}$ and $\chi_{\rm p}$) and the chirp mass ($m_{\rm chirp}$), and predict characteristic features in them expected from hierarchical mergers. We use $m_{\rm chirp}$ as this variable is most precisely determined by GW observations, and $\chi_{\rm eff}$ and $\chi_{\rm p}$ as these characterize the BH spin magnitudes in a binary. Here, $\chi_{\rm eff}$ and $\chi_{\rm p}$ are defined as

$$\chi_{\rm eff} = \frac{m_1 a_1 \cos\theta_1 + m_2 a_2 \cos\theta_2}{m_1 + m_2}, \tag{1}$$

and

$$\chi_{\rm p} = \max\left(a_1 \sin\theta_1, q\,\frac{4q+3}{4+3q} a_2 \sin\theta_2\right), \tag{2}$$

(Hannam et al. 2014; Schmidt, Ohme & Hannam 2015), where m_1 and m_2 are the masses, a_1 and a_2 are the spin magnitudes, θ_1 and θ_2 are the angles between the orbital angular momentum directions and the BH spins of the binary components, $q \equiv m_2/m_1 \leq 1$ is the mass ratio, and $m_{\rm chirp} \equiv (m_1 m_2)^{3/5}(m_1 + m_2)^{-1/5}$. We identify and characterize features expected in hierarchical mergers using mock GW data, and find that intrinsic properties (maximum mass and typical spin magnitude) of 1g BHs can be constrained by recovering the features, which enables us to distinguish astrophysical models. By analysing the GW data obtained in LIGO/Virgo O1–O3a, we investigate whether such features are consistent with observed GW data, and identify the astrophysical population models most consistent with the data. Finally, using mock GW data, we estimate how well parameters characterizing the spin distribution can be recovered in future catalogues depending on the number of events.

The paper is organized as follows. In Section 2, we describe our method to construct mock GW data and detect signatures for hierarchical mergers. We present our main results in Section 3, and give our conclusions in Section 4.

2 METHOD

2.1 Overview

We introduce a mock data set generated by a simple *N*-body toy model (Section 2.2), which allows us to explore hierarchical mergers more generically (Section 3.1). To identify features in the distributions representative of hierarchical mergers, we use a simple analytic model characterizing the spin distribution profile (Section 2.3.1), and apply it to the observed GW data (Section 3.2.1) and the *N*-body toy model (Section 3.3). Furthermore, to assess how well model predictions match the observed GW data, we also use a Bayes factor to assess relative likelihoods of models (including the *N*-body toy model and a physical model for mergers in AGN discs adopted from our simulations in Tagawa et al. 2021b; Section 3.2.2).

In the analyses, we mostly use $\chi_{\rm typ} \equiv (\chi_{\rm eff}^2 + \chi_{\rm p}^2)^{1/2}$ as it characterizes the spin magnitudes of BHs in binaries, and it is easily calculated from the quantities $\chi_{\rm eff}$ and $\chi_{\rm p}$ taken from LIGO Scientific Collaboration & Virgo Collaboration (2020, 2021). However, one should be aware of the following properties of $\chi_{\rm typ}$. First, unlike $\chi_{\rm eff}$, $\chi_{\rm p}$ is not conserved up to 2PN (e.g. Gerosa et al. 2020), suffering additional uncertainties due to its modulation. Secondly, due to the geometry, the contribution of $\chi_{\rm p}$ is on average larger than $\chi_{\rm eff}$ by a factor of $\sim$3 in cases of isotropic BH spins (equation 26 in the

[1]Note that the orbital radii where this takes place were derived by assuming Type-I migration (Bellovary et al. 2016), but these assumptions may be inconsistent for BHs embedded in AGN discs as gaps may be opened in the accretion discs (see e.g. equations 45–46, Kocsis, Yunes & Loeb 2011). Also, Pan & Yang (2021) found that the traps can disappear if radiation pressure is correctly accounted for.

Appendix). Thirdly, $\chi_{\rm p}$ is often unconstrained in the LIGO/Virgo events (e.g. Fig. B1).

2.2 Constructing mock GW data

To understand and analyse the distributions of $\chi_{\rm eff}$, $\chi_{\rm p}$, and $m_{\rm chirp}$ typically expected in hierarchical mergers, we employ mock GW data.

2.2.1 Overall procedure

We construct mock data by following the methodology of Doctor et al. (2020):

(i) Sample $N_{\rm 1g}$ BHs from 1g population as described in Section 2.2.2. We set $N_{\rm 1g} = 10^6$ to ensure a sufficient number for detectable mergers. We call this sample S.

(ii) Choose $\omega N_{\rm ng}$ pairs from S by weighing the pairing probability Γ (Section 2.2.2), where ω is the fraction of BHs that merge at each step, and $N_{\rm ng}$ is the number of BHs in the sample S ($N_{\rm ng} = N_{\rm 1g}$ in the first iteration).

(iii) Compute the remnant mass and spin, and the kick velocity for merging pairs assuming random directions for BH spins, where we use the method described in Tagawa et al. (2020). Update the sample S by removing BHs that have merged, and adding merger remnants if the kick velocity is smaller than the escape velocity ($v_{\rm esc}$).

(iv) Repeat steps 2–3 for $N_{\rm s}$ steps.

(v) Determine the fraction of detectable mergers by assessing whether signal-to-noise ratio (SNR) of mergers exceeds the detection criteria (Section 2.2.3). Randomly choose $N_{\rm obs}$ observed mergers from the detectable merging pairs. Add observational errors following Section 2.2.3, and construct a mock GW data set.

By changing the underlying parameters of the merging binaries in mock GW data ($\boldsymbol{\lambda}_0$; presented in the next section), we can construct various $\chi_{\rm eff}$, $\chi_{\rm p}$, and $m_{\rm chirp}$ distributions expected in hierarchical mergers. For example, $N_{\rm s}$ and ω influence the fraction of hierarchical mergers ($\propto\sim \omega^{N_{\rm s}}$), while $N_{\rm s}$ specifies the maximum generation and mass of BHs.

2.2.2 First generation BHs and pairing

We assume that the masses of 1g BHs are drawn from the power-law distribution as

$$p_{m_{\rm 1g}} \propto \begin{cases} m_{\rm 1g}^{-\alpha} & \text{for } m_{\rm min} < m_{\rm 1g} < m_{\rm max}, \\ 0 & \text{otherwise}, \end{cases} \tag{3}$$

where α is the power-law slope, $m_{\rm min}$ and $m_{\rm max}$ are the minimum and maximum masses, respectively.

We set the dimensionless spin magnitude for 1g BHs to

$$a_{\rm ini} = |a_{\rm ave} + a_{\rm uni} U[-1:1]|, \tag{4}$$

where $U[-1{:}\ 1]$ represent uniform distribution randomly chosen from -1 to 1, and $a_{\rm ave}$ and $a_{\rm uni}$ are parameters characterizing initial spins of 1g BHs. We assume that the spin magnitude for 1g BHs does not depend on the masses of 1g BHs. This assumption may be justified for single BHs, for which slow rotation is motivated by theoretical considerations (Fuller & Ma 2019). Here we assume $a_{\rm ave} = a_{\rm uni} = 0$ in the fiducial model. On the other hand, for mergers of field binaries (FBs), a large fraction of BHs may experience tidal synchronization, and the dispersion of the BH spin magnitudes decreases with BH masses (e.g. Hotokezaka & Piran 2017; Bavera et al. 2020; Safarzadeh, Farr & Ramirez-Ruiz 2020). The spin distribution expected in this pathway is considered in Section 3.1.2.

We assume the redshift distribution of merging BHs as

$$p_z \propto \frac{{\rm d}V_c}{{\rm d}z}\frac{1}{1+z}, \tag{5}$$

so that a merger rate density is uniform in comoving volume and source-frame time. Here, ${\rm d}V_c/{\rm d}z$ is calculated assuming ΛCDM cosmology with the Hubble constant $H_0 \simeq 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$, the matter density today $\Omega_{\rm m0} = 0.24$, and the cosmological constant today $\Omega_{\Lambda 0} = 0.76$ (Planck Collaboration XIII 2016). We also investigate different choices in Section A (see also Fishbach, Holz & Farr 2018; Yang et al. 2020a). We set the maximum redshift to be 1.5 considering LIGO/Virgo sensitivities (The LIGO Scientific Collaboration 2019).

To draw merging pairs, we simply assume that the interaction rate depends on the binary masses with a form

$$\Gamma \propto (m_1 + m_2)^{\gamma_{\rm t}}\, q^{\gamma_{\rm q}}, \tag{6}$$

as employed in Doctor et al. (2020). This parametrization enables us to mimic the effects that massive and equal-mass binaries are easy to merge in plausible models due to exchanges at binary–single interactions, mass segregation in clusters, interaction with ambient gas, mass transfer, or common-envelope evolution (e.g. O'Leary et al. 2016; Rodriguez et al. 2019; Tagawa et al. 2021b; Olejak et al. 2020).

Using the model described above and adding observational errors (Section 2.2.3), we can construct a mock observational data set. The parameter set characterizing a mock data set is $\boldsymbol{\lambda}_0 = \{\alpha, m_{\rm min}, m_{\rm max}, a_{\rm ave}, a_{\rm uni}, \gamma_{\rm t}, \gamma_{\rm q}, \omega, N_{\rm s}, v_{\rm esc}, N_{\rm obs}\}$. The fiducial choice of $\boldsymbol{\lambda}_0$ is described in Section 2.2.4 and Table 1.

2.2.3 Mock observational errors

To construct mock GW data, we need to put observational errors on observables. The true values of observables $\boldsymbol{\theta}$ are produced through the procedures in Sections 2.2.1 and 2.2.2 assuming a set of the population parameters $\boldsymbol{\lambda}_0$. To incorporate observational errors to the mock data, we refer to the prescription in Fishbach & Holz (2020). We assume that the binary is detected if the SNR of the signal in a single detector exceeds eight. We set the typical SNR, ρ_0, of a binary with parameters $m_{\rm chirp}$, $\chi_{\rm eff}$, and the luminosity distance $d_{\rm L}$ to

$$\rho_0 = 8\left[\frac{m_{\rm chirp}(1+z)}{m_{\rm chirp,8}}\right]^{5/6}\frac{d_{\rm L,8}}{d_{\rm L}}\left(1+\frac{3}{8}\chi_{\rm eff}\right), \tag{7}$$

where we fix $m_{\rm chirp,8} = 10\,{\rm M}_\odot$ and $d_{\rm L,8} = 1\,{\rm Gpc}$ (see equation 26 in Fishbach et al. 2018). This scaling approximates the amplitude of a GW signal, $m_{\rm chirp,8}$ and $d_{\rm L,8}$ are chosen to roughly match the typical values detected by LIGO at design sensitivity (Chen et al. 2017), and the dependence on $\chi_{\rm eff}$ roughly reproduces results in The LIGO Scientific Collaboration (2019). We calculate $d_{\rm L}$ from z assuming ΛCDM cosmology as stated above. The true SNR depends on the angular factor Θ, and is given by

$$\rho = \rho_0 \Theta. \tag{8}$$

Θ plays the combined role of the sky location, inclination, and polarization on the measured GW amplitude. We tune the width of the distribution to control the uncertainty of the measured signal strength, which in turn controls the uncertainty on the measured luminosity distance. We simply set Θ to a lognormal distribution with

$$\log\Theta \sim N\left(0, \frac{0.3}{1+\frac{\rho_0}{8}}\right), \tag{9}$$

following Fishbach et al. (2018).

Table 1. Fiducial values of our model parameters.

Parameter	Fiducial value
The number of observed events	$N_{\rm obs} = 1000$
Frequency of mergers for high-mass binaries	$\gamma_{\rm t} = 2$
Frequency of mergers for equal-mass binaries	$\gamma_{\rm q} = 2$
The spin magnitudes for 1g BHs	$a_{\rm ave} = 0$, $a_{\rm uni} = 0$
Maximum and minimum masses for 1g BHs	$m_{\rm max} = 20\,{\rm M}_\odot$, $m_{\rm min} = 5\,{\rm M}_\odot$
Power-law exponent in the mass function for 1g BHs	$\alpha = 1$
Fraction of BHs that merges at each step	$\omega = 0.1$
Number of merger steps	$N_{\rm s} = 4$
Escape velocity of systems hosting BHs	$v_{\rm esc} = 1000\,{\rm km\,s^{-1}}$
The parameter for correlation between the steps and the redshift	$w_z = \infty$ (no correlation)

From the true parameters ρ, $m_{\rm chirp}(1+z)$, z, $\chi_{\rm eff}$ and Θ, we assume that the four parameters, the SNR ($\rho_{\rm obs}$), the chirp mass ($m_{\rm chirp,\,obs}$), $\chi_{\rm eff,\,obs}$, and $\chi_{\rm p,\,obs}$, are given with errors as below. We assume that the fractional uncertainty on the detector-frame chirp mass is

$$\sigma_{m_{\rm chirp}} = \frac{8}{\rho_{\rm obs}} 0.04\, m_{\rm chirp}(1+z), \tag{10}$$

that on the SNR is

$$\sigma_\rho = 1 \tag{11}$$

following Fishbach & Holz (2020), and that on $\chi_{\rm eff}$ and $\chi_{\rm p}$ is, respectively,

$$\sigma_{\chi_{\rm eff}} = 0.1\frac{8}{\rho_0}, \tag{12}$$

and

$$\sigma_{\chi_{\rm p}} = 0.2\frac{8}{\rho_0}, \tag{13}$$

which roughly match typical observational error magnitudes in Abbott et al. (2019, 2020a). We assume that the observed median values $\tilde{m}_{\rm chirp,obs}$, $\tilde{\rho}_{\rm obs}$, $\tilde{\chi}_{\rm eff,obs}$, and $\tilde{\chi}_{\rm p,obs}$, respectively, from a normal distribution centred on the true values $m_{\rm chirp}(1+z)$, ρ, $\chi_{\rm eff}$, and $\chi_{\rm p}$ with the standard deviation $\sigma_{m_{\rm chirp}}$, σ_ρ, $\sigma_{\chi_{\rm eff}}$, and $\sigma_{\chi_{\rm p}}$. We further assume that the posterior distributions of $m_{\rm chirp}$, ρ, $\chi_{\rm eff}$, and $\chi_{\rm p}$ including errors for GW data in the $i^{\rm th}$ event are, respectively, calculated by drawing from a normal distribution centred on $\tilde{m}_{\rm chirp,obs}$, $\tilde{\rho}_{\rm obs}$, $\tilde{\chi}_{\rm eff,obs}$, and $\tilde{\chi}_{\rm p,obs}$ with the standard deviation $\sigma_{m_{\rm chirp}}$, σ_ρ, $\sigma_{\chi_{\rm eff}}$, and $\sigma_{\chi_{\rm p}}$. An observed value of z is calculated from $d_{\rm L}$ derived by incorporating the observed values to equation (7) and the relation between z and $d_{\rm L}$ so that equation (7) is valid for derived z.

2.2.4 Numerical choices

Table 1 lists the parameter values adopted in the fiducial model. Referring to Fuller & Ma (2019), we set small BH spin magnitudes for 1g BHs as $a_{\rm ave} = a_{\rm uni} = 0$. The power-law slope in the mass function for 1g BHs is given as $\alpha = 1$. Assuming mergers in (active phase of) NSCs, where hierarchical mergers are probably most frequent, we set $m_{\rm max} = 20\,{\rm M}_\odot$ as NSCs are mainly metal-rich (e.g. Do et al. 2018; Schödel et al. 2020), $v_{\rm esc} = 1000\,{\rm km\,s^{-1}}$ typically expected for merging sites of binaries (Tagawa et al. 2020), $\gamma_{\rm t} = 2$ and $\gamma_{\rm q} = 2$ as high and equal-mass BHs are easier to merge in dynamical environments, and $\omega = 0.1$ and $N_{\rm s} = 4$ to reproduce frequent hierarchical mergers (Table 2 415.0pt; Tagawa et al. 2021b).

2.3 Reconstruction of the spin distribution

Here, we present a way to detect features for hierarchical mergers that possibly appear in the distribution of spins and masses.

2.3.1 Model characterizing the spin distribution

Given the universal trends of hierarchical mergers in the averaged spin magnitude as a function of masses for merging binaries (Section 3.1.1), we investigate how well such trends can be reconstructed using a finite number of events. To do this, we replace the procedure above with a simple parametric analytic toy model, directly describing the distribution of the three variables ($\boldsymbol{\theta} = \{\chi_{\rm eff}, \chi_{\rm p}, m_{\rm chirp}\}$) in terms of a set of the parameters ($\boldsymbol{\lambda}$) as

$$p(\boldsymbol{\theta}|\boldsymbol{\lambda}) = N(\chi_{\rm typ}|\mu_\chi, \sigma_\chi)T[-1, 1], \tag{14}$$

where $N(x_0|x_1, x_2)$ represents the probability to return x_0 for the normal distribution with the mean x_1 and the standard deviation x_2, $T[-1, 1]$ means to truncate the normal distribution to the range $[-1, 1]$ and normalize N so that the integral of N in this range is 1, $\chi_{\rm typ} \equiv (\chi_{\rm p}^2 + \chi_{\rm eff}^2)^{1/2}$,

$$\mu_\chi = \begin{cases} a_\mu m_{\rm chirp} + b_\mu & \text{for } m_{\rm chirp} < m_{\rm crit} \\ b_\mu & \text{for } m_{\rm chirp} \geq m_{\rm crit}, \end{cases} \tag{15}$$

and

$$\sigma_\chi = \begin{cases} a_\sigma m_{\rm chirp} + b_\sigma & \text{for } m_{\rm chirp} < m_{\rm crit} \\ b_\sigma & \text{for } m_{\rm chirp} \geq m_{\rm crit}. \end{cases} \tag{16}$$

We use $\chi_{\rm typ}$ since it roughly represents the spin magnitudes of BHs in a binary. Hence, this model has five parameters $\boldsymbol{\lambda} = \{a_\mu, b_\mu, a_\sigma, b_\sigma, m_{\rm crit}\}$ characterizing the $\chi_{\rm typ}$ profile as a function of $m_{\rm chirp}$. The functional form of the model (equation 14) is motivated by the prediction that hierarchical mergers favour a plateau in the distribution of $\bar{\chi}_{\rm typ}$ *versus* $m_{\rm chirp}$ at high $m_{\rm chirp}$ as the BH spin magnitudes roughly converge to a constant value of $\sim$0.7 as a result of mergers with isotropic spin directions, while $\bar{\chi}_{\rm typ}$ roughly linearly approaches the value at the plateau from lower $m_{\rm chirp}$ according to Figs 1 and A1.

We simply adopt the same functional form for σ_χ with ν_χ. Since the BHs formed from mergers typically have spins dominated by the orbital angular momentum of their progenitor binary (i.e. $\sim$0.7), the dispersion in the $\chi_{\rm typ}$ distribution is expected to converge to a constant beyond $m_{\rm crit}$, producing a plateau. This motivates the functional form of equation (16) to describe the relation between the spins and mass for hierarchical mergers.

The model parameters, $\boldsymbol{\lambda}$, are estimated from GW data through a Bayesian analysis, whose details are described in the next section.

Table 2. Properties of hierarchical mergers in our models. The first and second columns indicate the model number and its variation from the fiducial model (Table 1). The third and fourth columns show the fraction of high-g mergers among all and detectable mergers, respectively. The fifth column shows the maximum chirp mass ($m_{\rm chirp,\,max}$) among $N_{\rm obs} = 10^3$ detectable mergers. The sixth and seventh columns show the average and the standard deviation of $\chi_{\rm p}$ among all merging pairs.

model	Parameter	High-g fraction	High-g detection fraction	$m_{\rm chirp,\,max}[{\rm M}_\odot]$	$\bar{\chi}_{\rm p}$	$\sigma(\chi_{\rm p})$
M1	Fiducial	0.33	0.68	56	0.17	0.26
M2	Globular cluster (GC)	0.063	0.17	44	0.030	0.13
M3	Field binary (FB)	0	0	33	0	0
M4	Migration trap (MT)	0.31	0.80	42	0	0
M5	$a_{\rm uni} = 1$	0.32	0.73	52	0.50	0.21
M6	$a_{\rm ave} = 0.99$	0.31	0.70	51	0.75	0.20
M7	$a_{\rm ave} = 0.66$, $a_{\rm uni} = 0.1$	0.33	0.72	55	0.55	0.13
M8	$a_{\rm ave} = 0.5$	0.33	0.74	65	0.46	0.12
M9	$m_{\rm max} = 30\,{\rm M}_\odot$	0.35	0.73	70	0.18	0.26
M10	$N_{\rm obs} = 50$, $N_{\rm s} = 3$	0.25	0.62	28	0.13	0.24
M11	$N_{\rm obs} = 50$, $N_{\rm s} = 2$	0.15	0.28	24	0.077	0.19
M12	$N_{\rm obs} = 50$, $N_{\rm s} = 2$, $\omega = 0.05$	0.077	0.18	19	0.040	0.14
M13	$N_{\rm obs} = 50$, $N_{\rm s} = 2$, $\omega = 0.03$	0.046	0.14	19	0.023	0.11

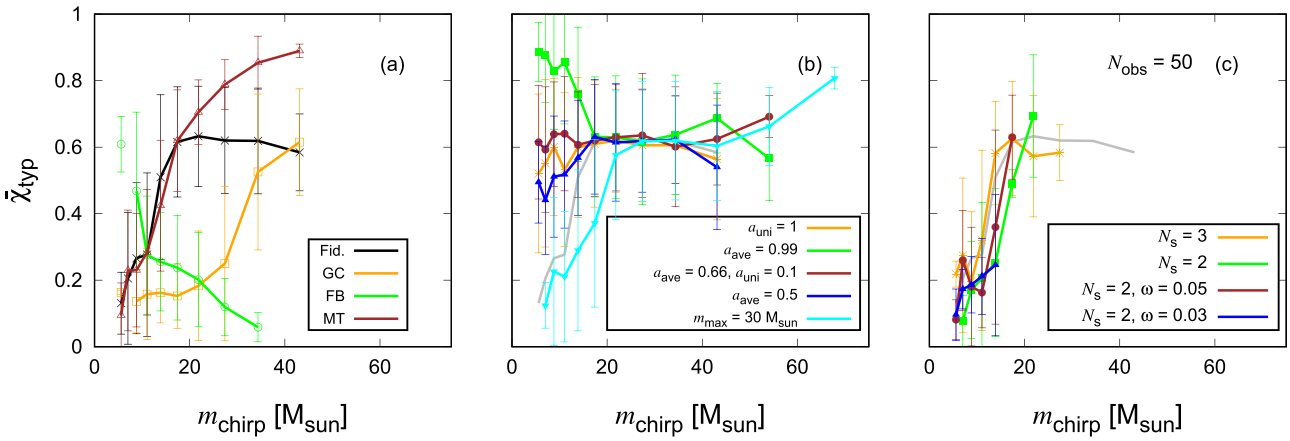

Figure 1. The mean dimensionless total spin $\bar{\chi}_{\rm typ}$ as a function of $m_{\rm chirp}$ for models M1–M13 (Table 2). We use $N_{\rm obs} = 10^3$ detectable mergers. In panels (b)–(c), the profiles for model M1 are presented by grey lines. Bars correspond to 1σ credible intervals.

2.3.2 *Bayesian analysis*

To derive the posterior distribution of $\boldsymbol{\lambda}$ from a data set $\{\boldsymbol{d}_i\}$, $p(\boldsymbol{\lambda}|\{\boldsymbol{d}_i\})$, we use the Bayesian formalism as follows. Here, $\boldsymbol{d}_i$ encodes the measurable parameters ($\boldsymbol{\theta}$) and also includes their random noise in the ith event. Bayes' rule gives

$$p(\boldsymbol{\lambda}|\{\boldsymbol{d}_i\}) = \frac{p(\{\boldsymbol{d}_i\}|\boldsymbol{\lambda})\pi(\boldsymbol{\lambda})}{p(\{\boldsymbol{d}_i\})}, \tag{17}$$

where $p(\{\boldsymbol{d}_i\}|\boldsymbol{\lambda})$ is the likelihood to obtain $\{\boldsymbol{d}_i\}$ for $\boldsymbol{\lambda}$, $\pi(\boldsymbol{\lambda})$ is the prior probability for the model parameters $\boldsymbol{\lambda}$, and the evidence $p(\{\boldsymbol{d}_i\})$ is the integral of the numerator over all $\boldsymbol{\lambda}$.

We assume that each GW detection is independent so that

$$p(\{\boldsymbol{d}_i\}|\boldsymbol{\lambda}) = \prod_{i=1}^{N_{\rm obs}} p(\boldsymbol{d}_i|\boldsymbol{\lambda}). \tag{18}$$

The probability of making observation i is

$$p(\boldsymbol{d}_i|\boldsymbol{\lambda}) = \frac{\int {\rm d}\boldsymbol{\theta}\, p(\boldsymbol{d}_i|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{\lambda})}{A(\boldsymbol{\lambda})}, \tag{19}$$

where the normalization factor $A(\boldsymbol{\lambda})$ is given by

$$\begin{aligned} A(\boldsymbol{\lambda}) &= \int_{\boldsymbol{d}>{\rm threshold}} {\rm d}\boldsymbol{d} \int {\rm d}\boldsymbol{\theta}\, p(\boldsymbol{d}|\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{\lambda}) \\ &= \int {\rm d}\boldsymbol{\theta}\, p_{\rm det}(\boldsymbol{\theta})p(\boldsymbol{\theta}|\boldsymbol{\lambda}), \end{aligned} \tag{20}$$

$$p_{\rm det}(\boldsymbol{\theta}) = \int_{\boldsymbol{d}>{\rm threshold}} p(\boldsymbol{d}|\boldsymbol{\theta}){\rm d}\boldsymbol{d}, \tag{21}$$

is the detection probability for a given set of parameters, and 'threshold' denotes that the event $\boldsymbol{d}$ is detectable when $\boldsymbol{d}$ is above the threshold. To reduce computational costs, we assume that $A(\boldsymbol{\lambda})$ is constant. This assumption does not affect our results as $A(\boldsymbol{\lambda})$ varies by less than a factor of 1.1 if the spin directions of BHs are assumed to be isotropic, meaning that the variation of $A(\boldsymbol{\lambda})$ per each steps in the Monte Carlo method (Section 2.3.3) is negligible. This is because the detection probability is influenced only by $\chi_{\rm eff}$ by changing $\boldsymbol{\lambda}$ (see equations 7 and 14), and the reduction and enhancement of the detectable volume for mergers with negative and positive $\chi_{\rm eff}$ are mostly cancelled out.

The likelihood $p(\boldsymbol{d}_i|\boldsymbol{\theta})$ can be rewritten in terms of the posterior probability density function (PDF) $p(\boldsymbol{\theta}|\boldsymbol{d}_i)$ that is estimated in the analysis assuming prior $\pi(\boldsymbol{\theta})$ as

$$p(\boldsymbol{d}_i|\boldsymbol{\theta}) = \frac{p(\boldsymbol{\theta}|\boldsymbol{d}_i)p(\boldsymbol{d}_i)}{\pi(\boldsymbol{\theta})}. \tag{22}$$

The posterior PDF $p(\boldsymbol{\theta}|\boldsymbol{d}_i)$ has information on errors, and it is often discretely sampled with S_i samples from the posterior, $\{^j\boldsymbol{\theta}^{(i)}\}$, for $j \in [1, S_i]$. Because the samples are drawn according to the posterior, the parameter space volume associated with each sample is inversely proportional to the local PDF, $d^j\theta^{(i)} \propto [p(^j\theta^{(i)}|d^{(i)})]^{-1}$, which allows us to replace the integral with a discrete sum (e.g. Mandel, Farr & Gair 2019; Vitale et al. 2020). Overall, the posterior distribution of $\boldsymbol{\lambda}$ is given as

$$\begin{aligned} &p(\boldsymbol{\lambda}|\{\boldsymbol{d}_i\}) \\ &= \frac{\pi(\boldsymbol{\lambda})}{p(\{\boldsymbol{d}_i\})} \prod_{i=1}^{N_{\rm obs}} \frac{\frac{1}{S_i}\sum_{j=1}^{S_i} p(^j\boldsymbol{\theta}^{(i)}|\boldsymbol{\lambda}) \frac{p(\boldsymbol{d}_i)}{\pi(\boldsymbol{\theta})}}{A(\boldsymbol{\lambda})} \\ &\propto \pi(\boldsymbol{\lambda}) \prod_{i=1}^{N_{\rm obs}} \frac{\frac{1}{S_i}\sum_{j=1}^{S_i} p(^j\boldsymbol{\theta}^{(i)}|\boldsymbol{\lambda}) \frac{1}{\pi(\boldsymbol{\theta})}}{A(\boldsymbol{\lambda})}, \end{aligned} \tag{23}$$

where we factor out the evidence factors $p(\{\boldsymbol{d}_i\})$ and $\prod_{i=1}^{N_{\rm obs}} p(\boldsymbol{d}_i)$ since it is independent of $\boldsymbol{\lambda}$ and does not affect the relative values of the posterior $p(\boldsymbol{\lambda}|\{\boldsymbol{d}_i\})$. We use a flat prior distribution for $\pi(\boldsymbol{\lambda})$. We set $\pi(\boldsymbol{\theta}) \propto d_L^2(z)$ following the standard priors used in the LIGO/Virgo analysis of individual events (Veitch et al. 2015). We assume flat priors on $\chi_{\rm p}$ and $\chi_{\rm eff}$. Note that this is different from the LIGO/Virgo analysis which used uniform priors for the component spin magnitudes and they are appropriately transformed to priors for $\chi_{\rm p}$ and $\chi_{\rm eff}$. We set $S_i = 3N_{\rm obs}$ so that we can take into account uncertainties whose probability is in the order of $\sim 1/N_{\rm obs}$.

2.3.3 *Markov Chain Monte Carlo methods*

We calculate the posterior distribution (equation 23) using Markov chain Monte Carlo (MCMC) methods. We track one chain for 10^7 steps, set the first half to a burn-in period, check convergence by verifying that values for parameters after the burn-in period are oscillating around a constant average and dispersion. We adopt Metropolis–Hastings algorithm (e.g. Hastings 1970), and set a proposal distribution to the normal distribution with the values at each step as the means and the standard deviations for a_μ, b_μ, a_σ, b_σ, and $m_{\rm crit}$ to be 0.0001 $\rm M_\odot^{-1}$, 0.01, 0.0001 $\rm M_\odot^{-1}$, 0.01, and 1.0 $\rm M_\odot$, respectively. The standard deviations of the proposal distribution are roughly given by the typical standard deviations of the posterior distribution divided by $\sim$4 as this setting works well for convergence. We do not pose thinning to a posterior distribution as the autocorrelation for each variable between adjacent steps is already as small as $\lesssim 10^{-5}$. We restrict $m_{\rm crit}$ in the ranges from $m_{\rm min}$ to the maximum $m_{\rm chirp}$ among observed events.

3 RESULTS

In Section 3.1, we investigate characteristic features in hierarchical mergers, using our flexible tool (Section 2.2) to generate mock GW data sets for a large range of input parameter combinations. In Section 3.2, we analyse GW data observed in LIGO/Virgo O1–O3a. We first derive signatures and properties of hierarchical mergers (Section 3.2.1), using the simple fitting formula for spin *versus* chirp mass (Section 2.3.1). We then assess (Section 3.2.2) how well the predictions in our mock GW catalogues and in our physical AGN disc models (Tagawa et al. 2021b), in fact, match these observed GW data. Finally, in Section 3.3, we analyse mock GW data, and investigate how well the signatures of hierarchical models, described by the simple fitting formulae ($\bar{\chi}_{\rm typ}$ *versus* $m_{\rm chirp}$), can be recovered from future larger GW catalogues.

Table 3. Adopted parameter values for several populations. The differences with respect to the fiducial model (Table 1) are listed.

	Globular cluster (GC)
1	$m_{\rm max} = 45\,\rm M_\odot$
2	$N_{\rm s} = 2$
3	$\omega = 0.03$
4	$v_{\rm esc} = 30\,\rm km\,s^{-1}$
	Field binary (FB)
1	$m_{\rm max} = 45\,\rm M_\odot$
2	$N_{\rm s} = 1$
3	$\theta_1 = \theta_2 = 0$
4	$a_{\rm uni}$ follows equation (27)
	Migration trap (MT)
1	$\theta_1 = \theta_2 = 0$

3.1 Profiles for average spin parameters

3.1.1 *Dependence on population parameters*

We first show the parameter dependence of the $\bar{\chi}_{\rm typ}$ profile as a function of $m_{\rm chirp}$ using mock GW events, in which hierarchical mergers are assumed to be frequent. In Table 2, we list the model varieties we have investigated. These include the fiducial model (M1), and 12 different varieties (models M2–M13). We examine different choices of the initial spin magnitudes (models M5–M8) and the maximum mass of 1g BHs (model M9), the fraction of hierarchical mergers (models M10–M13), and the several parameter sets mimicking different populations (models M2–M4, Table 3). We also investigate a variety of additional models in the appendix (models M14–M28, Table A1).

Fig. 1 shows the profiles for models M1–M13 (Table 2). For models in which hierarchical mergers are frequent (panels b and c of Fig. 1 and Fig. A1), there are universal trends for hierarchical mergers in the $\bar{\chi}_{\rm typ}$ profiles: (i) increase (or decrease) of $\bar{\chi}_{\rm typ}$ to $\sim$0.6 at low $m_{\rm chirp}$. (ii) plateau of $\bar{\chi}_{\rm typ}$ with $\sim$0.6 at high $m_{\rm chirp}$. Thus, the profile is roughly characterized by two lines if hierarchical mergers are frequent, mergers originate mostly from one population, and the typical spin magnitude for 1g BHs does not depend on their masses. The profile of $\bar{\chi}_{\rm typ}$ strongly depends on $a_{\rm ave}$, $a_{\rm uni}$, and $m_{\rm max}$ (Fig. 1 b), while it is less affected by the other parameters (see Fig. A1).

The typical value of $\bar{\chi}_{\rm typ} \sim 0.6$ at the plateau can be understood as follows. When masses and spin magnitudes between the primary and secondary BHs are similar ($m_1 \sim m_2$ and $a_1 \sim a_2 \sim a_0$) and the directions of BH spins are isotropic, the typical magnitude of mass-weighted BH spins is

$$|\boldsymbol{a}_{\rm w}| = \left\langle \left| \frac{m_1\boldsymbol{a}_1 + m_2\boldsymbol{a}_2}{m_1 + m_2} \right| \right\rangle \sim \frac{\sqrt{7}}{3} a_0, \tag{24}$$

where $\langle \dots \rangle$ represents an average over the number of samples. If we approximate

$$\begin{aligned} \bar{\chi}_{\rm p} &\simeq \langle |\boldsymbol{a}_0||\cos\theta| \rangle \sim \tfrac{\pi}{4} a_0, \\ \overline{|\chi_{\rm eff}|} &\simeq \langle |\boldsymbol{a}_{\rm w}||\sin\theta| \rangle \sim \tfrac{1}{2} |\boldsymbol{a}_{\rm w}|, \end{aligned} \tag{25}$$

then

$$\bar{\chi}_{\rm typ} = \left(\overline{|\chi_{\rm eff}|}^2 + \bar{\chi}_{\rm p}^2\right)^{1/2} \sim \left[\left(\frac{\sqrt{7}}{3}\frac{1}{2}\right)^2 + \left(\frac{\pi}{4}\right)^2\right]^{1/2} a_0 \sim 0.90 a_0. \quad (26)$$

Since merger remnants typically have spin magnitudes of $a_0 \sim 0.7$ (Buonanno, Kidder & Lehner 2008), $\bar{\chi}_{\rm typ} \sim 0.6$ for mergers among high-g BHs, which is roughly consistent with the value at the plateau (Figs 1 and A1). Note that when $q \ll 1$, $|\boldsymbol{a}_{\rm w}| \sim a_0$ and so the average value is slightly enhanced to $\bar{\chi}_{\rm typ} \sim 0.93 a_0$.

As $m_{\rm max}$ increases, the bending point between the two lines increases (grey and cyan lines in Fig. 1 b). This is because $m_{\rm max}$ determines the critical mass above which all merging BHs are of high generations with high spins of $\sim$0.7. As the bending point is not influenced by the other parameters, the maximum mass of 1g BHs can be estimated from the bending point of the $\bar{\chi}_{\rm typ}$ profile. Note that since the bending points of the $\chi_{\rm p}$ and $\chi_{\rm eff}$ profiles are similar in shape to that of the $\chi_{\rm typ}$ profile for mergers with isotropic BH spins (Fig. 3 a), either $\chi_{\rm typ}$, $\chi_{\rm p}$, or $\chi_{\rm eff}$ can constrain the maximum mass of 1g BHs if the profiles are reconstructed well.

Additionally, $a_{\rm ave}$ and $a_{\rm uni}$ influence $\bar{\chi}_{\rm typ}$ at the smallest values of $m_{\rm chirp}$ (Fig. 1 b). This suggests that typical spin magnitudes of 1g BHs can be presumed by spins at small $m_{\rm chirp}$. However, note that $\bar{\chi}_{\rm typ}$ at small $m_{\rm chirp}$ is also influenced by the observational errors on $\chi_{\rm p}$ and $\chi_{\rm eff}$. Due to the smaller errors on $|\chi_{\rm eff}|$ compared to $\chi_{\rm p}$, $\overline{|\chi_{\rm eff}|}$ may constrain the typical spin values of 1g BHs more precisely using a number of events (green and orange lines in Fig. 3a). Note that $\bar{\chi}_{\rm p} > \overline{|\chi_{\rm eff}|}$ when the BH spins are isotropic due to their definition. In model M7, the average and the dispersion of the spin magnitude for 1g BHs are set to be roughly the same as for the merger remnants. In such cases, the signatures of hierarchical mergers cannot be identified from the spin distributions (brown line in Fig. 1 b). Also, for models in which the typical spin magnitude for 1g BHs are close to $\sim$0.7 (e.g. models M5 and M8), a large number of events are needed to detect the hierarchical merger signatures.

In Fig. 1 (c), we can see how the features for hierarchical mergers in the $\chi_{\rm typ}$ profile are influenced by the fraction of hierarchical mergers for $N_{\rm obs} = 50$. The plateau at high $m_{\rm chirp}$ is seen for $N_{\rm s} = 3$ (orange), while the rise of $\chi_{\rm typ}$ to $\sim$0.6 at low $m_{\rm chirp}$ is seen for $N_{\rm s} = 2$ with $\omega \geq 0.05$ (green and brown). These suggest that with $N_{\rm obs} = 50$ the plateau and the rise of $\chi_{\rm typ}$ to $\sim$0.6 can be confirmed when the detection fraction of mergers of high-g BHs roughly exceeds $\sim$0.5 and $\sim$0.15, respectively (models M10, M12; Table 2).

To summarize, the profile of $\bar{\chi}_{\rm typ}$ is mostly affected only by $a_{\rm ave}$, $a_{\rm uni}$, and $m_{\rm max}$, while the other parameters may affect the maximum $m_{\rm chirp}$ or the frequency of high-g mergers (Tables 2 and A1).

3.1.2 *Contribution from multiple populations*

In the fiducial model (M1), the parameter values (Table 1) are roughly adjusted to reproduce properties of mergers in AGN discs outside of MTs (Tagawa et al. 2021b) or NSCs. The $\chi_{\rm typ}$ profile is similar, but the $m_{\rm chirp}$ distribution is different between the fiducial model and physically motivated models derived in Tagawa et al. (2021b). The former is because the profile is characterized by the few parameters ($m_{\rm max}$, $a_{\rm uni}$, $a_{\rm ave}$) as found in Section 3.1, while the latter is because the $m_{\rm chirp}$ distributions are affected by how BHs pair with other BHs and merge in AGN discs.

In this section, we additionally consider the spin distributions for mergers typically expected in several environments, including GCs, FBs, and MTs of AGN discs. Values of the parameters adopted to mimic these populations are listed in Table 3. Figs 2 and 3, and panel (a) in Fig. 1 present the distributions and the profiles of the spin parameters ($\chi_{\rm typ}$, $\chi_{\rm p}$, and $\chi_{\rm eff}$) as functions of $m_{\rm chirp}$ for these populations. Fig. 4 is the same as Fig. 3, but mergers are contributed by a mixture of two populations. Some contribution from multiple populations to the observed events is also favoured by the analysis in Zevin et al. (2020a).

For mergers in GCs, we set lower escape velocity $v_{\rm esc} = 30\,{\rm km\,s^{-1}}$, $N_{\rm s} = 2$ and $\omega = 0.03$ to reproduce the detection fraction of hierarchical mergers of $\sim$10–20 per cent, which is predicted by theoretical studies (e.g. O'Leary et al. 2016; Rodriguez et al. 2019; Table 2). We chose higher $m_{\rm max} = 45\,{\rm M}_\odot$ as GCs are composed of metal-poor stars (e.g. Peng et al. 2006; Leaman, VandenBerg & Mendel 2013; Brodie et al. 2014); other parameters are the same as those for AGN discs. Note that $m_{\rm max}$ in metal-poor environments is uncertain due to uncertainties on the reaction rate of carbon burning (Farmer et al. 2019) and the enhancement of the helium core mass by rotational mixing (Chatzopoulos & Wheeler 2012; Yoon, Dierks & Langer 2012; Vink et al. 2021).

Due to higher $m_{\rm max}$, $\bar{\chi}_{\rm typ}$ continues to increase until higher $m_{\rm chirp}$ (panel b in Fig. 3, see also Rodriguez et al. 2018) compared to the fiducial model (panel a). Also, 90 percentile regions are distributed around $\chi_{\rm eff} \sim 0$ and $\chi_{\rm p} \sim 0$ (Fig. 2 b) as a large fraction of mergers are among 1g BHs. Thus, the distribution of $\bar{\chi}_{\rm typ}$ at low $m_{\rm chirp}$ is clearly different between mergers in AGN discs and GCs, mainly due to the difference of $m_{\rm max}$ and the fraction of mergers among high-g BHs. If mergers are comparably contributed both by GCs and AGN discs, steep increase of $\bar{\chi}_{\rm typ}$ against $m_{\rm chirp}$ appears twice (panel a in Fig. 4). Thus, mixture of these populations can be discriminated by analysing the spin distribution. Note that the intermediate line between the two increases in the $\bar{\chi}_{\rm typ}$ profile is roughly characterized by the ratio of mergers from AGN discs and GCs. Hence, the contribution from multiple populations would be distinguishable by analysing the profile by using a number of GW events.

For mergers among FBs, we set $N_{\rm s} = 1$ and $m_{\rm max} = 45\,{\rm M}_\odot$. Although BH spin distributions are highly uncertain, we refer to Bavera et al. (2020), who proposed that $\bar{\chi}_{\rm eff}$ is high at low $m_{\rm chirp}$ of $\lesssim$ 10–20 $\rm M_\odot$ as low-mass progenitors have enough time to be tidally spun up. We assume that $a_{\rm uni}$ follows

$$a_{\rm uni} = \begin{cases} 1 \ \text{for} \ m_{\rm 1g} \leq 15\,{\rm M}_\odot \\ (30\,{\rm M}_\odot - m_{\rm 1g})/15\,{\rm M}_\odot \\ \qquad \text{for} \ 15\,{\rm M}_\odot \leq m_{\rm 1g} \leq 30\,{\rm M}_\odot \\ 0 \ \text{for} \ 30\,{\rm M}_\odot \leq m_{\rm 1g}. \end{cases} \quad (27)$$

BH spins are assumed to be always aligned with the orbital angular momentum of binaries, although we do not always expect spins to be aligned (e.g. Kalogera 2000; Rodriguez et al. 2016). In such a setting, $\overline{|\chi_{\rm eff}|}$ decreases as $m_{\rm chirp}$ increases (panels c of Figs 2 and 3). Also, non-zero $\chi_{\rm p}$ is due to assumed observational errors (orange line in Fig. 3 c). The profile expected for the binary evolution channel is significantly different from those expected for the other channels. If mergers arise comparably from FBs and GCs, $\overline{|\chi_{\rm eff}|}$ exceeds $\bar{\chi}_{\rm p}$ at low $m_{\rm chirp}$ (panel c of Fig. 4). As contribution from mergers in FBs enhances $\overline{|\chi_{\rm eff}|}$ relative to $\bar{\chi}_{\rm p}$ at low $m_{\rm chirp}$, we could constrain the contribution from FBs using the ratio of $\overline{|\chi_{\rm eff}|}$ to $\bar{\chi}_{\rm p}$. Observed events so far suggest that $|\chi_{\rm eff}|$ is typically lower than $\chi_{\rm p}$ at low $m_{\rm chirp}$ (panel e of Fig. 3), implying that the contribution to the observed mergers from FBs is minor, unless adopted spins for 1g BHs need significant revisions.

For mergers in MTs, we assume that parameters are the same as in the fiducial model (Table 1), while BH spins are always aligned

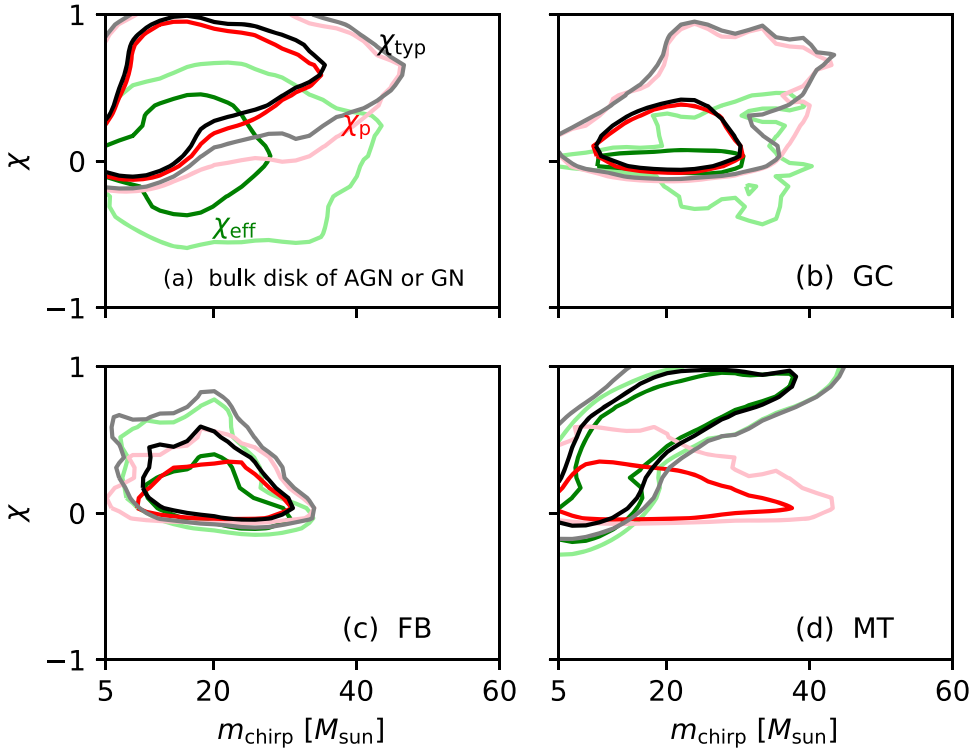

Figure 2. The 90 (dark lines) and 99 (light lines) percentile distributions in the spin *versus* $m_{\rm chirp}$ plane. Black, red, and green lines represent the distributions of $\chi_{\rm typ}$, $\chi_{\rm p}$, and $\chi_{\rm eff}$, respectively. Panels (a), (b), (c), and (d) show distributions for mergers in AGNs (bulk discs), GCs, FBs, and MTs, respectively.

with the orbital angular momentum of the binaries. Such alignment is expected for binaries in MTs where randomization of the binary orbital angular momentum directions by binary–single interactions is inefficient due to rapid hardening and merger caused by gas dynamical friction (unlike in gaps formed further out in the disc where these interactions were found to be very important by Tagawa et al. 2020), and so the BH spins are aligned with circumbinary discs due to the Bardeeen–Petterson effect (Bardeen & Petterson 1975), and circumbinary discs are aligned with the binaries due to viscous torque (e.g. Moody, Shi & Stone 2019). Here, we assume that the orbital angular momentum directions of binaries are the same as that of the AGN disc referring to Lubow, Seibert & Artymowicz (1999), which is different from the assumption (anti-alignment with 50 per cent) adopted in Yang et al. (2019). In this model, the $\chi_{\rm p}$ and $|\chi_{\rm eff}|$ distributions are significantly different from those in the other models (panels d of Figs 2 and 3). The value of $\chi_{\rm eff}$ at high $m_{\rm chirp}$ is typically high, while $\chi_{\rm p}$ is low. When mergers originate comparably in MTs and GCs, $\overline{|\chi_{\rm eff}|}$ significantly exceeds $\bar{\chi}_{\rm p}$ in a wide range of $m_{\rm chirp}$ (Fig. 4 d). As $\overline{|\chi_{\rm eff}|}$ is typically lower than $\bar{\chi}_{\rm p}$ in the observed events in all $m_{\rm chirp}$ bins (Fig. 4 e), the contribution from MTs to the detected mergers is probably minor.

3.2 Application to LIGO/Virgo O1–O3a data

3.2.1 Reconstruction of spin profiles

We analyse the GW data observed in LIGO/Virgo O1–O3a reported by Abbott et al. (2019, 2021). Although $\chi_{\rm p}$ and $\chi_{\rm typ}$ suffer large uncertainties (e.g. Fig. B1), their median values indicate a positive correlation with $m_{\rm chirp}$. Such positive correlation is, if confirmed, consistent with the growth of BH spin magnitudes by hierarchical mergers as presented in Figs 1, 3, and A1.

To confirm the features in the χ-profiles due to hierarchical mergers, we reconstruct the $\bar{\chi}_{\rm typ}$ profile from the observed GW data in the way described in Section 2.3. We discretize the posteriors for $m_{\rm chirp}$, $\chi_{\rm eff}$, and $\chi_{\rm p}$ with 20, 40, and 20 bins in the ranges from the minimum to the maximum of posteriors for $m_{{\rm chirp},\,i}$, from -1 to 1, and from 0 to 1, respectively. Note that the prior and posterior distributions for some events are similar to each other, which means that $\chi_{\rm p}$ is less constrained by the waveforms. To exclude events in which $\chi_{\rm p}$ are not well estimated, we only use events in which the Kullback–Leibler (KL) divergence between prior and posterior samples evaluated using heuristic estimates of $\chi_{\rm p}$ ($D_{\rm KL}$) exceeds a critical value of $D_{\rm KL,\,cri} = 0$, 0.05, 0.1, 0.15, or 0.2. We consider that $\chi_{\rm p}$ for events with non-zero $D_{\rm KL}$ is statistically useful to understand the spin distribution. We use the events with $m_2 \geq 5\,{\rm M}_\odot$ provided in LIGO Scientific Collaboration & Virgo Collaboration (2020) and LIGO Scientific Collaboration & Virgo Collaboration (2021) as we do not model mergers of neutron stars. Then, the number of events with $D_{\rm KL} \geq 0$, 0.05, 0.1, 0.15, and 0.2 are 44, 28, 20, 12, and 7, respectively. We present 1σ errors on the estimated parameters below unless stated otherwise.

The reconstructed $\bar{\chi}_{\rm typ}$ profiles for $D_{\rm KL,\,cri} = 0$, 0.05, 0.1, and 0.15 are, respectively, presented by orange lines in panels (a)–(d) of Fig. 5, and the posterior distributions and correlations of the reconstructed parameters for $D_{\rm KL,\,cri} = 0$ are presented in Fig. C1 in the Appendix. For $D_{\rm KL,\,cri} = 0$, 0.05, 0.1, 0.15, and 0.20, respectively, $\bar{\chi}_{\rm typ}$ at the plateau is $b_\mu = 0.51^{+0.14}_{-0.07}$, $0.55^{+0.14}_{-0.08}$, $0.55^{+0.19}_{-0.11}$, $0.62^{+0.22}_{-0.23}$, and $0.66^{+0.20}_{-0.18}$, the critical chirp mass at the bending point of the $\bar{\chi}_{\rm typ}$ profile is $m_{\rm crit} = 31^{+23}_{-15}\,{\rm M}_\odot$, $29^{+21}_{-12}\,{\rm M}_\odot$, $30^{+21}_{-14}\,{\rm M}_\odot$, $31^{+18}_{-15}\,{\rm M}_\odot$, and $36^{+17}_{-16}\,{\rm M}_\odot$, and the slope of $\bar{\chi}_{\rm typ}$ at $m_{\rm chirp} < m_{\rm crit}$ is $a_\mu = 10^{+7}_{-4} \times 10^{-3}\,{\rm M}_\odot^{-1}$, $12^{+8}_{-5} \times 10^{-3}\,{\rm M}_\odot^{-1}$, $12^{+7}_{-6} \times 10^{-3}\,{\rm M}_\odot^{-1}$, $15^{+12}_{-7} \times 10^{-3}\,{\rm M}_\odot^{-1}$, and $15^{+8}_{-6} \times 10^{-3}\,{\rm M}_\odot^{-1}$. To understand the influence of GW190521, which seems to have a large impact on spin distributions due to its

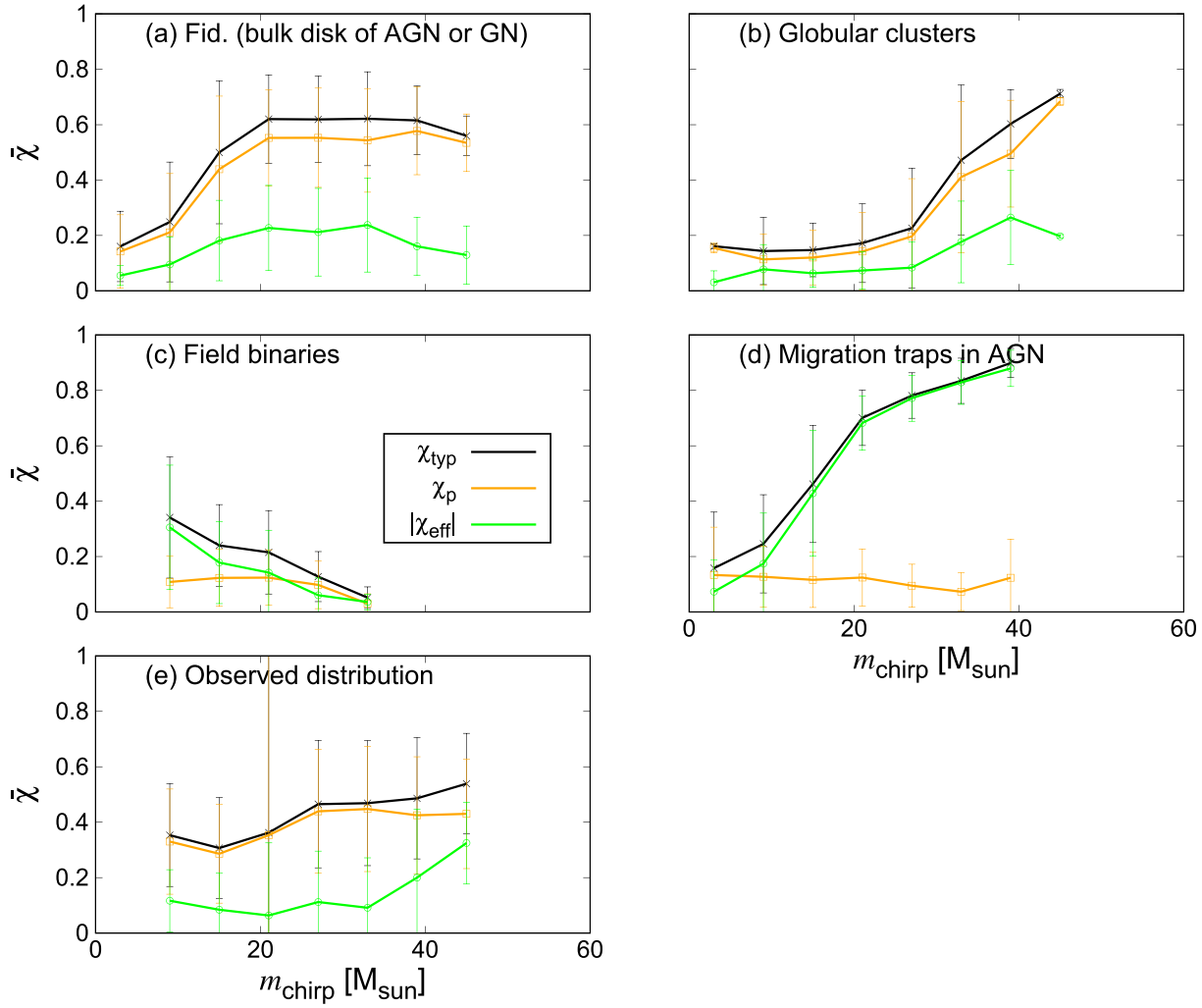

Figure 3. The profiles for the average spin parameters as a function of $m_{\rm chirp}$ for 10^3 detectable mergers. Black, orange, and green lines represent the average of $\chi_{\rm typ}$, $\chi_{\rm p}$, and $|\chi_{\rm eff}|$, respectively. Panels (a)–(e), respectively, present the distributions for mergers in AGN discs (M1), GCs (M2), FBs (M3), MTs (M4), and those observed by LIGO/Virgo O1–O3a. The averages for observed distributions (e) are calculated by averaging the medians of the parameters estimated in observed events. Bars correspond to 1σ credible intervals.

large mass and $\chi_{\rm p}$, we repeated our analysis excluding this event. In this case, for $D_{\rm KL,\,cri} = 0$, 0.05, and 0.1, respectively, $b_\mu = 0.50^{+0.15}_{-0.07}$, $0.52^{+0.14}_{-0.08}$, and $0.59^{+0.24}_{-0.14}$, $m_{\rm crit} = 36^{+19}_{-14}\,{\rm M}_\odot$, $25^{+23}_{-11}\,{\rm M}_\odot$, and $37^{+19}_{-17}\,{\rm M}_\odot$, and $a_\mu = 8^{+4}_{-4} \times 10^{-3}\,{\rm M}_\odot^{-1}$, $12^{+10}_{-6} \times 10^{-3}\,{\rm M}_\odot^{-1}$, and $12^{+4}_{-5} \times 10^{-3}\,{\rm M}_\odot^{-1}$, while for $D_{\rm KL,\,cri} = 0.15$ and 0.2, the parameters are not well determined due to the small number of events. For $D_{\rm LK,\,crit} \leq 0.1$, the evaluated values of the parameters are similar with and without GW190521.

The positive value of the slope (a_μ), i.e. the increase of $\bar{\chi}_{\rm typ}$ at low $m_{\rm chirp}$ is confirmed with $\gtrsim 2\sigma$ confidence, which is a tell-tale sign of frequent hierarchical mergers. Also, according to the analysis in Section 3.1, the detection of the rise of $\bar{\chi}_{\rm typ}$ at low $m_{\rm chirp}$ with $N_{\rm obs} = 50$ roughly requires that the detection fraction of mergers of high-g BHs exceeds $\sim$0.15. As the number of events is smaller than 50 (e.g. $N_{\rm obs} = 28$ for $D_{\rm KL,\,cri} = 0.05$), the high-g detection fraction would be even higher than $\sim$0.15. Thus, hierarchical mergers are preferred from the analysis. Note that accretion can also produce a positive correlation, but $|\chi_{\rm eff}| > \chi_{\rm p}$ is predicted in such cases similarly to mergers in MTs (panel d of Fig. 3). As $|\chi_{\rm eff}| < \chi_{\rm p}$ is predicted by GW observations (panel e of Fig. 3), accretion is disfavoured as a process enhancing the BH spin magnitudes.

For $D_{\rm KL,\,cri} = 0.05$, 0.1, 0.15, and 0.20 (panels b, c, and d of Fig. 5), the value of $\bar{\chi}_{\rm typ}$ at the plateau ($b_\mu \sim 0.6$) is consistent with that expected from hierarchical mergers ($\sim$0.6), which possibly supports frequent hierarchical mergers with the high-g detection fraction to be $\gtrsim 0.5$ (Section 3.1.1). On the other hand, for $D_{\rm KL,\,cri} = 0$, $b_\mu \sim 0.5$, which is somewhat lower than the expected value of 0.6. This is presumably because $\chi_{\rm p}$ values for events with $D_{\rm KL} \leq 0.05$ are not well constrained and just reflect assumed priors. Also, note that events with high $\chi_{\rm p}$ might tend to be missed as the waveform for large $\chi_{\rm p}$ (Apostolatos et al. 1994; Kidder 1995; Pratten et al. 2020) or spin (Kesden, Sperhake & Berti 2010; Gerosa et al. 2019) mergers often accompany strong amplitude modulation, reducing SNRs.

Here, $\bar{\chi}_{\rm typ}$ at $m_{\rm chirp} = m_{\rm min}$ is closely related to the typical spin magnitude for 1g BHs (Fig. 1 b). If we assume the median values for $\bar{\chi}_{\rm typ}$ and $m_{\rm crit}$, $\bar{\chi}_{\rm typ}$ at $m_{\rm chirp} = 5\,{\rm M}_\odot$ is $0.26^{+0.10}_{-0.18}$, $0.27^{+0.12}_{-0.20}$, $0.26^{+0.14}_{-0.19}$, $0.23^{+0.18}_{-0.32}$, and $0.29^{+0.20}_{-0.25}$ for $D_{\rm KL,\,cri} = 0$, 0.05, 0.1, 0.15, and 0.20, respectively. These suggest that 1g BHs typically have $a \lesssim 0.4$. Since this value is effectively enhanced by the observational errors on $\bar{\chi}_{\rm p}$, the estimated typical spin magnitude of 1g BHs is still consistent with $\sim$0 as predicted by stellar evolution models (Fuller & Ma 2019), which is also verified later (Section 3.3).

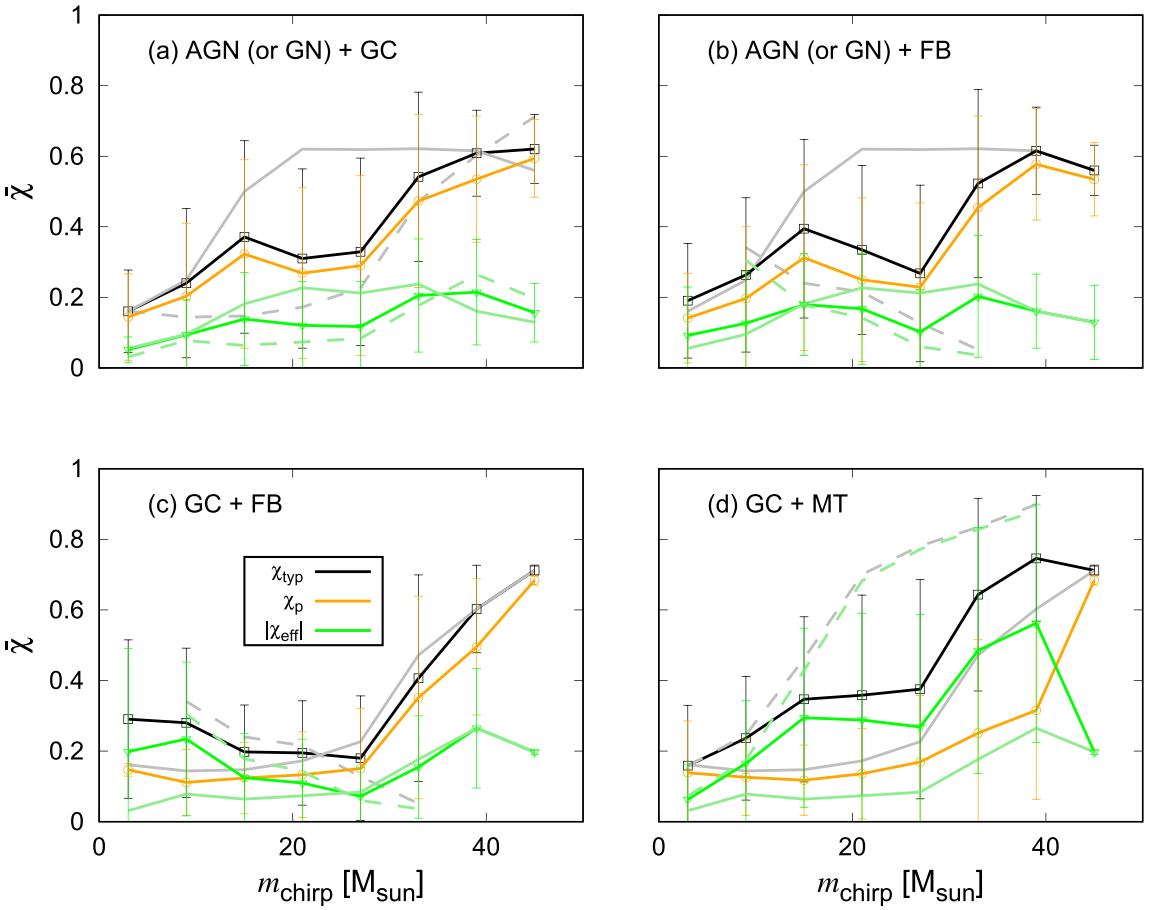

Figure 4. Same as Fig. 3, but for mixture of two models. Both channels contribute to 1000 mergers, and contributing two models are specified in the upper left of each panel. 'AGN', 'GC', 'FB', and 'MT' represent mergers in the AGN discs, GCs, FBs, and MTs, respectively. Thin and dashed lines present the $\chi_{\rm eff}$ and $\chi_{\rm typ}$ profiles only for the former and the latter populations in the legend, respectively.

The critical chirp mass at the bending point of the $\bar{\chi}_{\rm typ}$ profile ($m_{\rm crit}$) is related to the maximum mass of 1g BHs (Fig. 1 f). The analysis loosely constrains the parameter to $m_{\rm crit} \sim 15$–$50\,{\rm M}_\odot$, from which we discuss in Section 3.3 that the maximum mass of 1g BHs is estimated to be $\sim$20–$60\,{\rm M}_\odot$. However, it needs a caution that $m_{\rm crit}$ is restricted from $5\,{\rm M}_\odot$ to the maximum chirp mass among the event ($\sim 67\,{\rm M}_\odot$) in this analysis, which may artificially produce the bending point and the plateau. To confidently confirm the plateau, $m_{\rm crit}$ needs to be precisely constrained compared to the allowed range for $m_{\rm crit}$ of 5–$67\,{\rm M}_\odot$, which would require further events (see also Section 3.3).

3.2.2 *Bayes factors on spins and mass distributions*

In the previous section we focus on the $\bar{\chi}_{\rm typ}$ profile, while here we use the distributions of $\chi_{\rm eff}$, $\chi_{\rm p}$, and $m_{\rm chirp}$ and discuss the preferred values for underlying parameters $\boldsymbol{\lambda}_0$.

To assess the relative likelihood to produce each event in different models, we calculate the Bayes factors between pairs of models,

$$K_{\rm A,B} = \frac{P(\boldsymbol{d}|A)}{P(\boldsymbol{d}|B)}, \tag{28}$$

where

$$P(\boldsymbol{d}|A) = \prod_i P(\boldsymbol{d}_i|A), \tag{29}$$

$P(\boldsymbol{d}_i|A)$ is the likelihood of obtaining data $\boldsymbol{d}_i$ observed in the GW event i from model A,

$$P(\boldsymbol{d}_i|A) = \int P(\boldsymbol{d}_i|m_{\rm chirp}, \chi_{\rm eff}, \chi_{\rm p})\ P(m_{\rm chirp}, \chi_{\rm eff}, \chi_{\rm p}|A)\ {\rm d}m_{\rm chirp}{\rm d}\chi_{\rm eff}{\rm d}\chi_{\rm p} \tag{30}$$

and $P(m_{\rm chirp}, \chi_{\rm eff}, \chi_{\rm p}|A)$ is the probability distribution of $m_{\rm chirp}$, $\chi_{\rm eff}$, and $\chi_{\rm p}$ in model A. We calculate the three dimensional likelihood $P(\boldsymbol{d}_i|m_{\rm chirp}, \chi_{\rm eff}, \chi_{\rm p})$ for the events.

We calculate the Bayes factors for events with $D_{\rm KL} \geq D_{\rm KL,\,crit} = 0$, 0.05, 0.1, 0.15, and 0.2. We consider $D_{\rm KL,\,crit} = 0.05$ as the fiducial value, and mostly discuss the Bayes factors for $D_{\rm KL,\,crit} = 0.05$ below. Note that the events with positive Bayes factors for $D_{\rm KL,\,crit} = 0.1$, 0.15, or 0.2 always have positive Bayes factors also for $D_{\rm KL,\,crit} = 0.05$ somewhat incidentally.

To calculate $P(m_{\rm chirp}, \chi_{\rm eff}, \chi_{\rm p}|A)$, we first count mergers in $30 \times 30 \times 30$ uniform bins in $\chi_{\rm eff}$, $m_{\rm chirp}$, and $\chi_{\rm p}$ for model A. The maximum and minimum values of $m_{\rm chirp}$ for the bins are set to 100 and $5\,{\rm M}_\odot$, respectively. In this section, we generate 1000 mergers for each model. To include error distributions for the variables ($m_{\rm chirp}$, $\chi_{\rm eff}$, $\chi_{\rm p}$) to $P(m_{\rm chirp}, \chi_{\rm eff}, \chi_{\rm p}|A)$, we sample 10 different realizations for each merger event predicted by the model. To reduce the statistical fluctuation in the distribution of $\chi_{\rm eff}$, $m_{\rm chirp}$, and $\chi_{\rm p}$ due to the finite number of mergers in our models, we perform a kernel-density

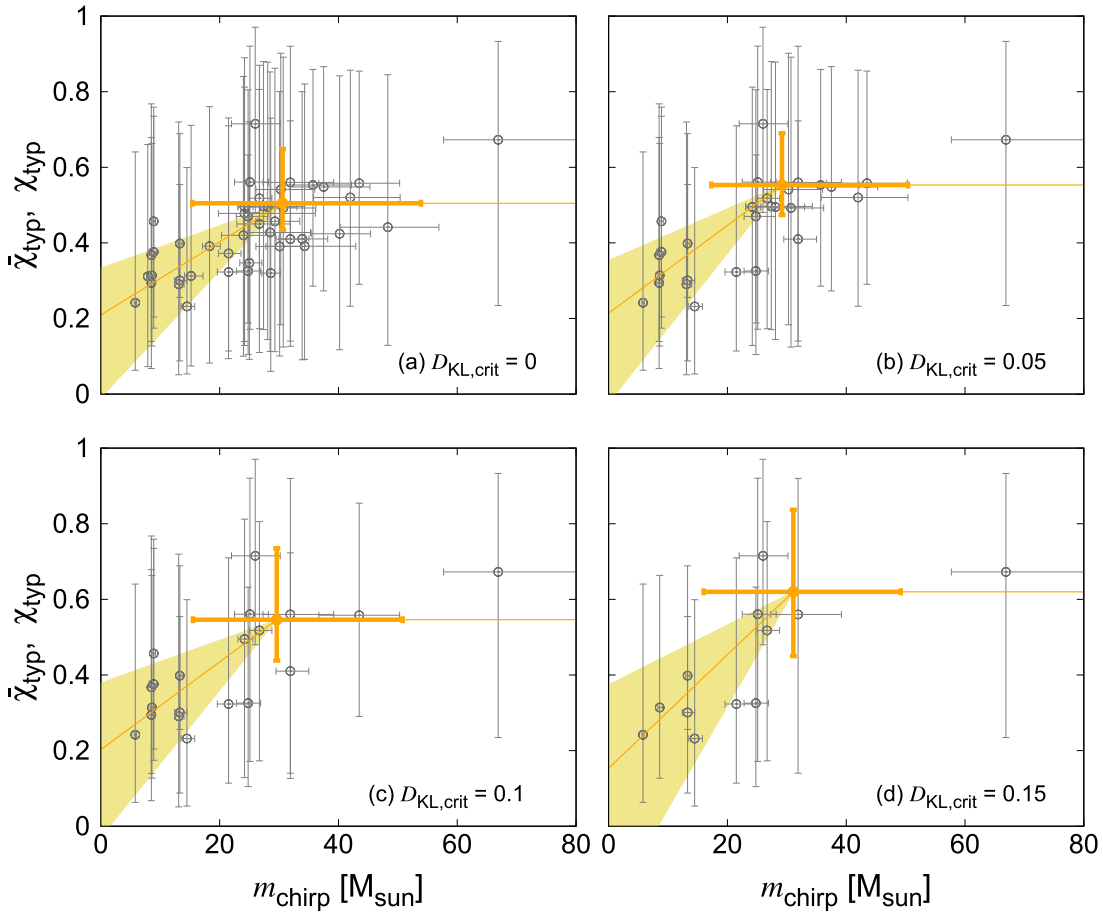

Figure 5. The $\bar{\chi}_{\rm typ}$ profile as a function of $m_{\rm chirp}$ constructed using the GW data observed in LIGO/Virgo O1–O3a. Orange line presents the recovered profile using the model described in Section 2.3 (equation 15). Bars and shaded regions represent 1σ credible intervals for $\bar{\chi}_{\rm typ}$ at the plateau b_μ, the critical chirp mass at the bending point of the profile $m_{\rm crit}$, and the slope of the $\bar{\chi}_{\rm typ}$ at a lower chirp mass a_μ. Panels (a), (b), (c), and (d) presents results for events with $D_{\rm KL} \geq 0$, 0.05, 0.1, and 0.15, respectively. Grey plots are the same as those in Fig. B1 (c).

estimate for the distribution using Gaussian kernels whose bandwidth is chosen to satisfy the Scott's Rule (Scott 1992). We calculate $P(\boldsymbol{d}_i|m_{\rm chirp}, \chi_{\rm eff}, \chi_{\rm p})$ by means of 300 samples generated according to the observed posterior distributions as used in the previous section.

For reference, we also calculate the Bayes factors for the two parameters, $m_{\rm chirp}$ and $\chi_{\rm eff}$, using the 44 events used in the analysis with $D_{\rm KL} \geq 0$ in the previous section.

Table 4 lists the Bayes factors for some models relative to the fiducial model (=*B*, Table 1). The Bayes factors suggest that, compared to the $m_{\rm chirp}$, $\chi_{\rm eff}$, and $\chi_{\rm p}$ distributions typically expected for mergers in FBs and MTs (Table 3), the observed distribution is much more consistent with those in AGN discs. This is because high $|\chi_{\rm eff}|$ and low $\chi_{\rm p}$ expected for mergers either in FBs or MTs (panels c and d in Figs 2 and 3) are incompatible with the observed distribution of $|\chi_{\rm eff}| < \chi_{\rm p}$ (Fig. 3 e).

For mergers in GCs, the models with small spin magnitudes for 1g BHs are less favoured. This is presumably because infrequent hierarchical mergers ($\lesssim$ 20 per cent) in GCs are difficult to explain typically high values of $\chi_{\rm p}$ if 1g BHs have low spin magnitudes. On the other hand, for $a_{\rm ave} = 0.3$ and $\alpha = 2$, the Bayes factor for $D_{\rm KL,\,crit} = 0.05$ is as high as $\sim 10^{0.4}$. Thus, if mergers originate from GCs, 1g BHs are favoured to have high spin magnitudes and follow a bottom heavy initial mass function.

For mergers in AGN discs or NSCs, the models with non-zero values for initial BH spins ($a_{\rm ave} = 0.3$) as well as a high value for $\sigma_{\chi_{\rm p}}$($\sim$ 0.3–0.4) have high Bayes factors of $10^{0.5}$ and $10^{0.05}$–$10^{0.4}$ for $D_{\rm KL,\,crit} = 0.05$, respectively. This is because non-zero $\bar{\chi}_{\rm typ}$ at low $m_{\rm chirp}$ in the observed distribution (Fig. B1) can be explained by adjusting these variables (Fig. 1). Also, large values for α, which effectively shift the $\chi_{\rm typ}$ and $m_{\rm chirp}$ distribution towards lower $m_{\rm chirp}$, and accordingly raises $\bar{\chi}_{\rm typ}$ at low $m_{\rm chirp}$ (e.g. Fig. A1 e). This is presumably the reason why the model with $\alpha(=2)$ has a high Bayes factor of $10^{0.8}$ at $m_{\rm max} \sim 25\,{\rm M}_\odot$ compared to the models with $\alpha = 1$ ($K_{A,\,B} \lesssim 1$).

Preferred values for $m_{\rm max}$ are probably as low as $\sim$15–30 $\rm M_\odot$ if the typical spin magnitude for 1g BHs is low. For $\alpha = 1$, in the models with $N_{\rm s} = 3$, 4, and 5, respectively, $m_{\rm max} = 25$–$30\,{\rm M}_\odot$, $m_{\rm max} = 20$–$25\,{\rm M}_\odot$, and $m_{\rm max} = 15$–$20\,{\rm M}_\odot$ is preferred. The difference in preference of $m_{\rm max}$ for different $N_{\rm s}$ is because both variables are constrained by the maximum $m_{\rm chirp}$ among the GW events. In any case, the preferred values of $m_{\rm max} = 15$–$30\,{\rm M}_\odot$ are roughly consistent with the values estimated in the previous section.

We also compare the properties inferred from GW observations with those predicted for mergers in AGN discs, which are calculated from one-dimensional *N*-body simulations, combined with a semi-

Table 4. The parameters of model A which are different from each population model (shown in Table 3) and the logarithm of their Bayes factor $K_{A,B}$ relative to the fiducial model ('B'). The Bayes factors for the three parameters with $D_{KL,cri} = 0$, 0.05, 0.1, 0.15, and 0.2, and that for the two parameters are presented from the second to seventh columns. We highlight the models with positive Bayes factors in the five rightmost columns in boldface.

Parameters	$\log_{10}K_{A,B}$ in 3D					$\log_{10}K_{A,B}$ in 2D
	$D_{KL,cri} = 0$	$D_{KL,cri} = 0.05$	$D_{KL,cri} = 0.1$	$D_{KL,cri} = 0.15$	$D_{KL,cri} = 0.2$	–
			AGN disc or NSC			
$m_{max} = 15\,M_\odot$	−6.9	−4.0	−1.8	−3.0	−1.5	−7.2
$m_{max} = 25\,M_\odot$	1.3	−0.05	−1.1	−0.31	−0.27	**1.3**
$m_{max} = 30\,M_\odot$	−0.9	−2.0	−2.7	−1.4	−0.90	**0.085**
$m_{max} = 35\,M_\odot$	−3.2	−3.7	−4.6	−2.4	−1.6	−1.8
$m_{max} = 45\,M_\odot$	−8.2	−7.6	−7.4	−4.2	−2.6	−5.7
$\alpha = 2$	−0.065	**0.35**	**0.65**	**0.065**	**0.29**	−0.56
$\alpha = 2$, $m_{max} = 15\,M_\odot$	−12	−8.3	−4.5	−4.1	−3.3	−15
$\alpha = 2$, $m_{max} = 25\,M_\odot$	1.5	**0.83**	**0.14**	−0.060	**0.13**	**1.4**
$\alpha = 2$, $m_{max} = 30\,M_\odot$	1.4	**0.12**	−1.0	−0.58	−0.33	**1.9**
$\alpha = 2$, $m_{max} = 35\,M_\odot$	0.11	−1.1	−2.0	−1.2	−0.57	**0.72**
$\alpha = 2$, $m_{max} = 45\,M_\odot$	−2.7	−3.1	−3.4	−2.0	−0.87	−1.3
$a_{uni} = 1$	−2.2	−1.7	−2.1	−1.7	−1.0	−1.3
$a_{ave} = 0.3$	0.51	**0.53**	**0.042**	−0.21	**0.062**	−0.035
$a_{ave} = 0.5$	−1.5	−0.85	−1.2	−1.2	−0.67	−1.1
$a_{ave} = 0.7$	−5.5	−4.1	−4.3	−3.0	−2.0	−2.4
$\sigma_p = 0.3$	−0.021	**0.053**	−0.059	−0.15	−0.16	−0.24
$\sigma_p = 0.4$	0.37	**0.42**	−0.075	−0.20	−0.072	**0.40**
$v_{esc} = 30\,km\,s^{-1}$	−14	−8.7	−6.0	−5.7	−5.7	−13
$w = 0.03$	−26	−23	−12	−13	−8.3	−30
$w = 0.2$	−3.2	−2.4	−3.1	−1.8	−0.81	−2.1
$\gamma_t = 0$	−21	−16	−11	−11	−8.2	−25
$\gamma_t = 4$	−3.7	−2.8	−3.1	−1.7	−0.64	−3.2
$\gamma_q = 0$	−6.5	−4.5	−2.3	−2.8	−2.0	−6.7
$\gamma_q = 4$	0.39	**0.12**	−0.26	−0.18	−0.058	**0.64**
$N_s = 2$	−49	−31	−25	−29	−25	−48
$N_s = 2$, $m_{max} = 30\,M_\odot$	−0.55	−3.3	−5.7	−2.9	−4.0	−0.84
$N_s = 2$, $m_{max} = 45\,M_\odot$	−3.6	−6.1	−7.9	−3.9	−3.6	−0.95
$N_s = 2$, $m_{max} = 60\,M_\odot$	−14	−13	−12	−6.9	−5.0	−9.5
$N_s = 3$	−4.1	−3.0	−1.5	−1.8	−2.0	−5.9
$N_s = 3$, $m_{max} = 25\,M_\odot$	1.0	−0.55	−1.2	−0.56	−0.77	**0.81**
$N_s = 3$, $m_{max} = 30\,M_\odot$	1.9	−0.11	−1.7	−0.61	−0.77	**2.1**
$N_s = 3$, $m_{max} = 35\,M_\odot$	0.24	−2.1	−3.8	−1.8	−1.5	**1.3**
$N_s = 3$, $m_{max} = 45\,M_\odot$	−6.7	−7.2	−7.7	−4.2	−2.9	−3.7
$N_s = 5$	0.084	**0.16**	−0.50	−0.48	**0.14**	**0.44**
$N_s = 5$, $m_{max} = 15\,M_\odot$	−0.96	−0.17	−0.091	−0.76	−0.065	−0.68
$N_s = 5$, $m_{max} = 30\,M_\odot$	−3.6	−3.4	−3.8	−2.1	−1.1	−2.6
$N_s = 5$, $m_{max} = 45\,M_\odot$	−11	−8.5	−7.8	−4.5	−2.4	−7.9
$N_s = 6$	−2.8	−2.0	−2.6	−1.5	−0.42	−2.0
			Globular cluster			
Fiducial	−2.3	−6.8	−8.7	−5.3	−4.8	−0.83
$\alpha = 2$	4.2	−0.50	−1.5	−1.4	−2.1	**5.8**
$m_{max} = 30\,M_\odot$	−12	−10	−8.7	−9.1	−7.0	−15
$a_{uni} = 1$	−12	−13	−16	−8.9	−9.6	−10
$a_{ave} = 0.3$	−0.84	−5.1	−7.8	−4.8	−5.2	−3.3
$a_{ave} = 0.5$	−1.8	−5.0	−8.6	−4.6	−4.6	−4.2
$a_{ave} = 0.7$	−12	−12	−14	−8.7	−7.9	−8.7
$v_{esc} = 100\,km\,s^{-1}$	−3.1	−7.1	−8.5	−4.8	−4.9	−1.4
$w = 0.05$	−2.1	−5.4	−7.1	−3.8	−3.9	**0.54**
$w = 0.1$	−3.3	−5.5	−7.3	−4.1	−3.5	−0.72
$\alpha = 2$, $a_{uni} = 1$	−4.1	−4.8	−5.6	−5.6	−4.8	−2.6
$\alpha = 2$, $a_{ave} = 0.3$	4.9	**0.38**	−0.72	−1.1	−2.4	**2.6**
$\alpha = 2$, $a_{ave} = 0.5$	0.15	−2.9	−5.0	−4.4	−4.6	−2.3
			Field binary			
Fiducial	−18	−17	−19	−12	−10	−9.1
			Migration trap			
Fiducial	−61	−37	−23	−17	−14	−57

Table 4 – *continued*

Parameters	$\log_{10}K_{A,B}$ in 3D					$\log_{10}K_{A,B}$ in 2D
	$D_{KL,cri}=0$	$D_{KL,cri}=0.05$	$D_{KL,cri}=0.1$	$D_{KL,cri}=0.15$	$D_{KL,cri}=0.2$	–
			AGN disc (Tagawa et al. 2021b)			
$f_{m1g}=1$	−4.0	−2.5	−1.6	−1.5	−0.63	−3.9
$f_{m1g}=1.33$	−2.7	−1.9	−1.5	−0.53	−0.19	−2.5
$f_{m1g}=1.66$	−2.0	−1.8	−1.8	−0.45	−0.16	−2.0
$f_{m1g}=2$	1.1	**0.032**	−0.79	−0.15	**0.045**	**1.2**
$f_{m1g}=3$	2.1	−0.19	−1.7	−1.7	−1.1	**4.3**

analytical model used in Tagawa et al. (2021b). We adopt the fiducial model in Tagawa et al. (2021b), while we investigate several variations in which the initial BH masses are multiplied by f_{m1g} =1, 1.33, 1.66, 2, and 3 so that m_{max} = 15, 20, 25, 30, and 45 $M_\odot$, respectively. Since 1g BH masses are 5–15 $M_\odot$ in the fiducial model, the minimum BH mass is given by $5f_{m1g}\,M_\odot$, in which the minimum chirp mass is $\sim 8.7 f_{m1g}\,M_\odot$. To eliminate a reduction of the likelihood due to the lack of 1g BHs in the low mass ranges, we here calculate Bayes factors only using events with $m_{chirp} \geq 8.7 f_{m1g}\,M_\odot$. The errors on m_{chirp}, χ_{eff}, and χ_p are simply given by the normal distribution with the standard deviation of $0.08\,m_{chirp}$, 0.12, and 0.2, respectively. The Bayes factors are listed in the bottom five rows in Table 4, which indicate that $m_{max} \sim 30\,M_\odot$ (f_{m1g} = 2) is preferred. Thus, the properties predicted for AGN disc-assisted mergers are likely to be consistent with the observed properties of the GW events.

Here, events with high Bayes factors for $D_{KL,crit} = 0.05$ tend to have high Bayes factors for the two dimensional likelihood (bold number in the third and rightmost columns of Table 4). We consider that this fact would further support the preferred models discussed above.

Overall, our analyses suggest that m_{max} = 15–30 $M_\odot$ with a high fraction of hierarchical mergers, or high spin magnitudes of ∼0.3 for 1g BHs is favoured. The former may support mergers in NSCs including AGN discs, while the latter may be consistent with those in GCs. Further events would be required to assess these possibilities in more detail.

We also discuss the spin distribution suggested in The LIGO Scientific Collaboration (2021). First, we compare the average and the standard deviation of χ_p predicted by models and those estimated from LIGO/Virgo O1–O3a data. By analysing the observed GW data, The LIGO Scientific Collaboration (2021) estimated that the average and the standard deviation of χ_p are $0.21^{+0.15}_{-0.14}$ and $0.09^{+0.21}_{-0.07}$, respectively, assuming a truncated mass model. These values are consistent with models in which hierarchical mergers are frequent such as models M1, M9–M11 (Table 2), M18–M21, M25–M28 (Table A1). Also, the average and the standard deviation χ_p for the model of GC with $a_{ave} = 0.3$ is 0.25 and 0.062, respectively, which are also consistent with the values estimated from the observed data. This fact further supports our claim that frequent hierarchical mergers or high spin magnitudes of ∼0.3 for 1g BHs is favoured. Here, note that the dependence of the spins on masses expected from hierarchical mergers is taken into account in our analysis, which would be a critical difference from that in The LIGO Scientific Collaboration (2020a).

Next, we discuss the fraction of mergers with positive and negative χ_{eff}. The LIGO Scientific Collaboration (2021) analysed the GW data observed in LIGO/Virgo O1–O3a, and estimated that $0.67^{+0.16}_{-0.16}$ (the 90 per cent credible intervals) and $0.27^{+0.17}_{-0.15}$ of mergers have $\chi_{eff} > 0.01$ and $\chi_{eff} < -0.01$, respectively. In the fiducial model (Table 1), the fraction of mergers with $\chi_{eff} > 0.01$ is 0.54 and that for $\chi_{eff} < -0.01$ is 0.41. The larger fraction for positive χ_{eff} compared to that for negative one in the model is due to the assumed dependence of ρ_0 on χ_{eff} in equation (7). The fraction of mergers with negative χ_{eff} in the model is somewhat higher than that estimated in The LIGO Scientific Collaboration (2021). Such difference may be due to large uncertainties for the estimated fraction, while it may suggest that the dependence of ρ_0 on χ_{eff} is stronger than that adopted in equation (7), or BH spins are moderately aligned toward the binary angular momentum directions due to interactions with gas, tidal synchronization, or alignment of spins for progenitor stars.

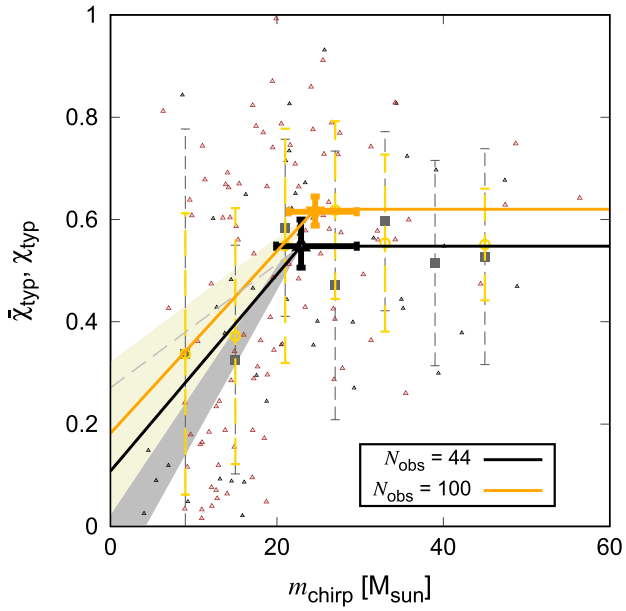

Figure 6. The $\bar{\chi}_{typ}$ profile as a function of m_{chirp} constructed using the mock GW data for 100 (orange), and 44 (black) observed events. Thick lines present the recovered profiles as in Fig. 5. The triangles corresponds to the median values of χ_{typ} for all mock events. The dashed lines show 1σ credible intervals for observed values of $\bar{\chi}_{typ}$ for mock events.

3.3 Reconstruction of the spin profile from mock GW data

We investigate how well the $\bar{\chi}_{typ}$ profile can be reconstructed from mock GW data (Section 2.2.3) for different values of N_{obs} by performing the MCMC method as described in Section 2.3. Fig. 6

shows $\bar{\chi}_{\rm typ}$ as a function of $m_{\rm chirp}$ for $N_{\rm obs} = 44$ (black) and $N_{\rm obs} = 100$ (orange) for the model with the fiducial setting (Table 1) but $m_{\rm max} = 30\,{\rm M}_\odot$ and $\alpha = 2$, which is preferred from observed GW events (Sections 3.2.1 and 3.2.2).

As the parameter estimate tends to be biased in small $N_{\rm obs}$, we additionally perform 10 models for $N_{\rm obs} = 44$ with same settings with independent realizations of the initial condition. By averaging the estimated parameters for 11 models, $\bar{\chi}_{\rm typ}$ at the plateau is $b_\mu = 0.63^{+0.07}_{-0.05}$ with the standard deviation $\sigma(b_\mu) = 0.04$, the critical chirp mass is $m_{\rm crit} = 24^{+6}_{-4}\,{\rm M}_\odot$ with $\sigma(m_{\rm crit}) = 5\,{\rm M}_\odot$, and the slope of $\bar{\chi}_{\rm typ}$ in $m_{\rm chirp} < m_{\rm crit}$ is $a_\mu = 25^{+9}_{-8} \times 10^{-3}\,{\rm M}_\odot^{-1}$ with $\sigma(b_\mu) = 8 \times 10^{-3}\,{\rm M}_\odot^{-1}$. As these uncertainties on the reconstructed parameters from the GW mock data are similar to those derived from the observed GW data in Section 3.2.1, we conclude that the GW mock data are a useful tool to understand how well the spin profile can be reconstructed.

The critical chirp mass is estimated to be $m_{\rm crit} = 25^{+6}_{-3}\,{\rm M}_\odot$ for $N_{\rm obs} = 100$, and $m_{\rm crit} = 24.9^{+0.9}_{-0.8}\,{\rm M}_\odot$ for $N_{\rm obs} = 1000$. Here, the estimated value of $m_{\rm crit}$ is lower than $m_{\rm max}$ by ~ 20 per cent mostly because $m_{\rm chirp} = (m_1 + m_2)[q(1+q)^{-2}]^{3/5} \lesssim 0.87 m_{\rm max}$. As the analysis on the observed GW events in Section 3.2.1 derives $m_{\rm crit} \sim 15$–$50\,{\rm M}_\odot$, $m_{\rm max} \sim 20$–$60\,{\rm M}_\odot$ is roughly inferred according to the relation of $m_{\rm max} \sim 1.2\,m_{\rm crit}$.

The average spin parameter $\bar{\chi}_{\rm typ}$ at $m_{\rm chirp} = m_{\rm min}$ is related to the typical spin magnitude of 1g BHs (e.g. Fig. 1). $\bar{\chi}_{\rm typ}$ at $m_{\rm chirp} = 5\,{\rm M}_\odot$ is $0.20^{+0.14}_{-0.17}$ for $N_{\rm obs} = 44$, $0.27^{+0.11}_{-0.13}$ for $N_{\rm obs} = 100$, and $0.23^{+0.03}_{-0.02}$ for $N_{\rm obs} = 1000$. These values derived from the model with $a_{\rm ave} = a_{\rm uni} = 0$ are similar to the value ($\sim 0.3 \pm \sim 0.1$) derived from the observed GW data (Section 3.2.1), suggesting that the typical spin magnitude of 1g BHs inferred from the observed GW events is still consistent with ~ 0.

For $N_{\rm obs} = 100$, $\bar{\chi}_{\rm typ}$ at the plateau is $b_\mu = 0.62^{+0.04}_{-0.03}$, which is similar to the expected value for hierarchical mergers (~ 0.6, Section 3.1.1). Also, the mass at the bending point is well constrained with $N_{\rm obs} = 100$ as mentioned above. Thus, with $N_{\rm obs} \geq 100$, parameters characterizing properties of hierarchical mergers, e.g. a value of $\bar{\chi}_{\rm typ}$ and $m_{\rm crit}$ at the plateau, are more precisely constrained.

Finally, to investigate whether the bending point is robustly verified, we also fit the distribution by a straight line, i.e. assuming $m_{\rm crit} \to \infty$ in equation (14), and calculate the Bayes factor of the model with broken lines (equation 14) compared to the model with a single line ($m_{\rm crit} \to \infty$), where we set the likelihood function to equation (14) with the fitted parameters. For $N_{\rm obs} = 44$, 100, and 1000, the logarithm of the Bayes factor is 1.5, 2.1, and 24, respectively. If we adopt the Akaike information criterion (Akaike 1974), the model with the broken lines is preferred by a factor of $\sim 10^{1.6}$ for $N_{\rm obs} = 100$, and the preference increases as $N_{\rm obs}$ increases. In the analysis using the observed data in Section 3.2.1, although we assumed the existence of the plateau, the Bayes factors using the observed events (with $D_{\rm KL,\,cri} = 0$, 0.05, 0.1, and 0.15) are in the range of $10^{-0.2}$–$10^{0.2}$, suggesting that the existence of the plateau is uncertain. Our analysis suggests that as the number of GW events increases to $\gtrsim O(100)$, the existence of the plateau can be confirmed with high significance.

4 SUMMARY AND CONCLUSIONS

In this paper we have investigated characteristic distributions of $\chi_{\rm eff}$, $\chi_{\rm p}$, $\chi_{\rm typ} = (\chi_{\rm p}^2 + \chi_{\rm eff}^2)^{1/2}$, and $m_{\rm chirp}$ expected from hierarchical mergers among stellar-mass BHs. We then used a toy model to derive the profile of the average of $\chi_{\rm typ}$ as a function of $m_{\rm chirp}$ for the events observed by LIGO/Virgo O1–O3a. We also investigated how well predictions in different models match observed spin and mass distributions by using Bayes factors. Finally, we estimate how well the $\chi_{\rm typ}$ profile can be reconstructed using mock GW data expected in hierarchical mergers. Our main results are summarized as follows:

(i) If hierarchical mergers are frequent, and the spin distribution of first-generation (1g) BHs does not strongly depend on their mass, the $\bar{\chi}_{\rm typ}$ profile as a function of $m_{\rm chirp}$ is characterized by a monotonic increase of $\bar{\chi}_{\rm typ}$ with $m_{\rm chirp}$ up to the maximum chirp mass among 1g BHs, and reaches a plateau of $\bar{\chi}_{\rm typ}$ with ~ 0.6 at higher $m_{\rm chirp}$ (Fig. 1). With ~ 50 events, the plateau and the rise of $\chi_{\rm typ}$ to 0.6 can be confirmed if the detection fraction of mergers of high-g BHs roughly exceeds ~ 0.5 and ~ 0.15, respectively.

(ii) The maximum mass for 1g BHs can be estimated by constraining the transition point between the two regimes in the $\bar{\chi}_{\rm typ}$ profile. Also, the typical spin magnitude for 1g BHs is constrained from $\bar{\chi}_{\rm typ}$ at around minimum $m_{\rm chirp}$ among GW events.

(iii) The $\bar{\chi}_{\rm typ}$ profile reconstructed from the LIGO/Virgo O1–O3a data prefers an increase in $\bar{\chi}_{\rm typ}$ at $m_{\rm chirp} \lesssim 15$–$50\,{\rm M}_\odot$ with $\sim 2\sigma$ confidence (Fig. 5), consistent with the evolution of BH spin magnitudes by hierarchical mergers. The maximum mass and the typical spin magnitude of 1g BHs are loosely constrained to be ~ 20–$60\,{\rm M}_\odot$ and $\lesssim 0.4$ with $\sim 1\sigma$ credible intervals, respectively.

(iv) A Bayesian analysis using the $\chi_{\rm eff}$, $\chi_{\rm p}$, and $m_{\rm chirp}$ distributions suggests that 1g BHs are preferred to have the maximum mass of $m_{\rm max} \sim 15$–$30\,{\rm M}_\odot$ if hierarchical mergers are frequent, which is consistent with mergers in AGN discs and/or nuclear star clusters. On the other hand, if mergers mainly originate from GC (in which $m_{\rm max}$ is assumed to be $45\,{\rm M}_\odot$), 1g BHs are favoured to have spin magnitudes of ~ 0.3. These favoured models are also consistent with the average and the standard deviation of $\chi_{\rm p}$ estimated in The LIGO Scientific Collaboration (2021).

(v) By using observed data of more than ~ 100 events in the future, we will be able to recover parameters characterizing the $\bar{\chi}_{\rm typ}$ distribution (e.g. the existence of the plateau and the value of $\bar{\chi}_{\rm typ}$ at the plateau b_μ) more precisely.

ACKNOWLEDGEMENTS

The authors greatly thank V. Gayathri for providing us the posterior distributions, Davide Gerosa for providing us values for the Kullback–Leibler divergence, and Berry Christopher and Kengo Tomida for valuable comments. This work was financially supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number JP21J00794 and for Basic Research by the Ministry of Education, Science and Culture of Japan (HT:17H01102, 17H06360, KO:17H02869, 17H01102, 17H06360). ZH acknowledges support from NASA grant NNX15AB19G and NSF grants 1715661 and 2006176. This work received founding from the European Research Council (ERC) under the European Union's Horizon 2020 Programme for Research and Innovation ERC-2014-STG under grant agreement No. 638435 (GalNUC) (to BK). IB acknowledges support from the Alfred P. Sloan Foundation and from the University of Florida. Simulations and analyses were carried out on Cray XC50 at the Center for Computational Astrophysics, National Astronomical Observatory of Japan.

DATA AVAILABILITY

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

REFERENCES

Abbott B. P. et al., 2019, Phys. Rev. X, 9, 031040
Abbott R. et al., 2020a, ApJ, 900, L13
Abbott R. et al., 2020b, ApJ, 896, L44
Abbott R. et al., 2021, Phys. Rev. X, 11, 021053
Acernese F. et al., 2015, Class. Quantum Gravity, 32, 024001
Akaike H., 1974, Proc. IEEE, 19, 716
Antonini F., Toonen S., Hamers A. S., 2017, ApJ, 841, 77
Antonini F., Gieles M., Gualandris A., 2019, MNRAS, 486, 5008
Apostolatos T. A., Cutler C., Sussman G. J., Thorne K. S., 1994, Phys. Rev. D, 49, 6274
Arca Sedda M., 2020, ApJ, 891, 47
Askar A., Davies M. B., Church R. P., 2021, MNRAS, 502, 2682
Banerjee S., 2017, MNRAS, 467, 524
Bardeen J. M., Petterson J. A., 1975, ApJ, 195, L65
Bartos I., Kocsis B., Haiman Z., Márka S., 2017, ApJ, 835, 165
Bavera S. S. et al., 2020, A&A, 635, A97
Belczynski K., Daniel E. H., Bulik T., O'Shaughnessy R., 2016, Nature, 534, 512
Bellovary J. M., Mac Low M.-M., McKernan B., Ford K. E. S., 2016, ApJ, 819, L17
Brodie J. P. et al., 2014, ApJ, 796, 52
Buonanno A., Kidder L. E., Lehner L., 2008, Phys. Rev. D, 77, 026004
Chatzopoulos E., Wheeler J. C., 2012, ApJ, 748, 42
Chen H.-Y., Holz D. E., Miller J., Evans M., Vitale S., Creighton J., 2021, Class. Quantum Gravity, 38, 055010
de Mink S. E., Mandel I., 2016, MNRAS, 460, 3545
Di Carlo U. N., Giacobbo N., Mapelli M., Pasquato M., Spera M., Wang L., Haardt F., 2019, MNRAS, 487, 2947
Do T., Kerzendorf W., Konopacky Q., Marcinik J. M., Ghez A., Lu J. R., Morris M. R., 2018, ApJ, 855, L5
Doctor Z., Wysocki D., O'Shaughnessy R., Holz D. E., Farr B., 2020, ApJ, 893, 35
Dominik M., Belczynski K., Fryer C., Holz D. E., Berti E., Bulik T., Mandel I., O'Shughnessy R., 2012, ApJ, 759, 52
Farmer R., Renzo M., de Mink S. E., Marchant P., Justham S., 2019, ApJ, 887, 53
Fishbach M., Holz D. E., 2020, ApJ, 891, L27
Fishbach M., Holz D. E., Farr B., 2017, ApJ, 840, L24
Fishbach M., Holz D. E., Farr W. M., 2018, ApJ, 863, L41
Fragione G., Kocsis B., 2018, Phys. Rev. Lett., 121, 161103
Fragione G., Kocsis B., 2019, MNRAS, 486, 4781
Fragione G., Grishin E., Leigh N. W. C., Perets H. B., Perna R., 2019, MNRAS, 488, 47
Fragione G., Loeb A., Rasio F. A., 2020, ApJ, 902, L26
Fuller J., Ma L., 2019, ApJ, 881, L1
Gerosa D., Berti E., 2017, Phys. Rev. D, 95, 124046
Gerosa D., Lima A., Berti E., Sperhake U., Kesden M., O'Shaughnessy R., 2019, Class. Quantum Gravity, 36, 105003
Gerosa D., Mould M., Gangardt D., Schmidt P., Pratten G., Thomas L. M., 2020, Phys. Rev. D, 103, 064067
Gerosa D., Vitale S., Berti E., 2020, Phys. Rev. Lett., 125, 101103
Gondán L., Kocsis B., Raffai P., Frei Z., 2018, ApJ, 860, 5
Hamers A. S., Safarzadeh M., 2020, ApJ, 898, 99
Hannam M., Schmidt P., Bohé A., Haegel L., Husa S., Ohme F., Pratten G., Pürrer M., 2014, Phys. Rev. Lett., 113, 151101
Hastings W. K., 1970, Biometrika, 57, 97
Hotokezaka K., Piran T., 2017, ApJ, 842, 111
Inayoshi K., Hirai R., Kinugawa T., Hotokezaka K., 2017, MNRAS, 468, 5020
Ivanova N. et al., 2013, A&AR, 21, 59
Kalogera V., 2000, ApJ, 541, 319
Kesden M., Sperhake U., Berti E., 2010, Phys. Rev. D, 81, 084054
Kidder L. E., 1995, Phys. Rev. D, 52, 821
Kimball C., Talbot C., Berry C. P. L., Carney M., Zevin M., Thrane E., Kalogera V., 2020, ApJ, 900, 177
Kimball C. et al., 2021, ApJ, 915, L35
Kinugawa T., Inayoshi K., Hotokezaka K., Nakauchi D. T. N., 2014, MNRAS, 442, 2963
Kocsis B., Yunes N., Loeb A., 2011, Phys. Rev. D, 84, 024032
Kumamoto J., Fujii M. S., Tanikawa A., 2018, MNRAS, 486, 3942
Leaman R., VandenBerg D. A., Mendel J. T., 2013, MNRAS, 436, 122
Liu B., Lai D., 2021, MNRAS, 502, 2049
Lubow S. H., Seibert M., Artymowicz P., 1999, ApJ, 526, 1001
Mandel I., de Mink S. E., 2016, MNRAS, 458, 2634
Mandel I., Farr W. M., Gair J. R., 2019, MNRAS, 486, 1086
Mapelli M., Santoliquido F., Bouffanais Y., Arca Sedda M., Giacobbo N., Artale M. C., Ballone A., 2020, preprint (arXiv:2007.15022)
Marchant P., Langer N., Podsiadlowski P., Tauris T., Moriya T., 2016, A&A, 588, A50
McKernan B. et al., 2018, ApJ, 866, 66
McKernan B., Ford K. E. S., O'Shaugnessy R., Wysocki D., 2020, MNRAS, 494, 1203
McKernan B., Ford K. E. S., O'Shaughnessy R., 2020, MNRAS, 498, 4088
Michaely E., Perets H. B., 2019, ApJ, 887, L36
Moody M. S. L., Shi J.-M., Stone J. M., 2019, ApJ, 875, 66
O'Leary R. M., Kocsis B., Loeb A., 2009, MNRAS, 395, 2127
O'Leary R. M., Meiron Y., Kocsis B., 2016, ApJ, 824, L12
Olejak A., Fishbach M., Belczynski K., Holz D. E., Lasota J. P., Miller M. C., Bulik T., 2020, ApJ, 901, L39
Paczynski B., 1976, in Eggleton P., Mitton S., Whelan J., eds, Proc. IAU Symp. 73, Structure and Evolution of Close Binary Systems. Kluwer, Dordrecht, p.75
Pan Z., Yang H., 2021, Phys. Rev. D, 103, 103018
Pavlovskii K., Ivanova N., Belczynski K., Van K. X., 2017, MNRAS, 465, 2092
Peng E. W. et al., 2006, ApJ, 639, 95
Planck Collaboration XIII, 2016, A&A, 594, A13
Portegies Zwart S. F., McMillan S. L. W., 2000, ApJ, 528, L17
Pratten G., Schmidt P., Buscicchio R., Thomas L. M., 2020, Phys. Rev. Res., 2, 043096
Rasskazov A., Kocsis B., 2019, ApJ, 881, 20
Rastello S., Amaro-Seoane P., Arca-Sedda M., Capuzzo-Dolcetta R., Fragione G., Tosta e Melo I., 2019, MNRAS, 483, 1233
Rastello S., Mapelli M., Di Carlo U. N., Giacobbo N., Santoliquido F., Spera M., Ballone A., Iorio G., 2020, MNRAS, 497, 1563
Rodriguez C. L. et al., 2020, ApJ, 896, L10
Rodriguez C. L., Zevin M., Pankow C., Kalogera V., Rasio F. A., 2016, ApJ, 832, L2
Rodriguez C. L., Chatterjee S., Rasio F. A., 2016, Phys. Rev. D., 93, 084029
Rodriguez C. L., Amaro-Seoane P., Chatterjee S., Rasio F. A., 2018, Phys. Rev. Lett., 120, 151101
Rodriguez C. L., Zevin M., Amaro-Seoane P., Chatterjee S., Kremer K., Rasio F. A., Ye C. S., 2019, Phys. Rev. D, 100, 043027
Safarzadeh M., Haiman Z., 2020, ApJ, 903, L21
Safarzadeh M., Hamers A. S., Loeb A., Berger E., 2020, ApJ, 888, L3
Safarzadeh M., Farr W. M., Ramirez-Ruiz E., 2020, ApJ, 894, 129
Samsing J. et al., 2020, preprint (arXiv:2010.09765)
Samsing J., MacLeod M., Ramirez-Ruiz E., 2014, ApJ, 784, 71
Schmidt P., Ohme F., Hannam M., 2015, Phys. Rev. D, 91, 024043
Schödel R., Nogueras-Lara F., Gallego-Cano E., Shahzamanian B., Gallego-Calvente A. T., Gardini A., 2020, A&A, 641, A102
Scott D. W., 1992, Multivariate Density Estimation: Theory, Practice, and Visualization., 2nd edn. Wiley, New York
Silsbee K., Tremaine S., 2017, ApJ, 836, 39
Spera M., Mapelli M., Giacobbo N., Trani A. A., Bressan A., Costa G., 2019, MNRAS, 485, 889
Stone N. C., Metzger B. D., Haiman Z., 2017, MNRAS, 464, 946
Tagawa H., Kocsis B., Saitoh R. T., 2018, Phys. Rev. Lett., 120, 261101
Tagawa H., Haiman Z., Bartos I., Kocsis B., 2020, ApJ, 899, 26
Tagawa H., Haiman Z., Kocsis B., 2020, ApJ, 898, 25
Tagawa H., Kocsis B., Haiman Z., Bartos I., Omukai K., Samsing J., 2021a, ApJ, 907, L20
Tagawa H., Kocsis B., Haiman Z., Bartos I., Omukai K., Samsing J., 2021b, ApJ, 908, 194

The LIGO Scientific Collaboration, 2015, Class. Quantum Gravity, 32, 074001
The LIGO Scientific Collaboration, 2019, Phys. Rev. D, 100, 064064
The LIGO Scientific Collaboration, 2020a, Phys. Rev. Lett., 125, 101102
The LIGO Scientific Collaboration, 2020b, Phys. Rev. D, 102, 043015
The LIGO Scientific Collaboration, Virgo Collaboration, 2020, LIGO Document P1800370-v5, Available at: https://dcc.ligo.org/LIGO-P1800370/public
The LIGO Scientific Collaboration, 2021, ApJ, 913, L7
The LIGO Scientific Collaboration, Virgo Collaboration, 2021, LIGO Document P2000223-v7, Available at: https://dcc.ligo.org/LIGO-P2000223/public/
Tiwari V., Fairhurst S., 2021, ApJ, 913, L19
van den Heuvel E. P. J., Portegies Zwart S. F., de Mink S. E., 2017, MNRAS, 471, 4256
Veitch J. et al., 2015, Phys. Rev. D, 91, 042003
Venumadhav T., Zackay B., Roulet J., Dai L., Zaldarriaga M., 2020, Phys. Rev. D, 101, 08303030
Vink J. S., Higgins E. R., Sander A. A. C., Sabhahit G. N., 2021, MNRAS, 504, 146
Vitale S., Gerosa D., Farr W. M., Taylor S. R., 2020, preprint (arXiv:2007.05579)
Yang Y. et al., 2019, Phys. Rev. Lett., 123, 181101
Yang Y., Bartos I., Haiman Z., Kocsis B., Márka S., Tagawa H., 2020a, ApJ, 896, 138
Yang Y., Gayathri V., Bartos I., Haiman Z., Safarzadeh M., Tagawa H., 2020b, ApJ, 901, L34
Yoon S. C., Dierks A., Langer N., 2012, A&A, 542, A113
Zackay B., Dai L., Venumadhav T., Roulet J., Zaldarriaga M., 2019, preprint (arXiv:1910.09528)
Zevin M. et al., 2021, ApJ, 910, 152
Zevin M., Spera M., Berry C. P. L., Kalogera V., 2020b, ApJ, 899, L1
Ziosi B. M., Mapelli M., Branchesi M., Tormen G., 2014, MNRAS, 441, 3703

APPENDIX A

We show the parameter dependence of the $\bar{\chi}_{\rm typ}$ profile as a function of $m_{\rm chirp}$ using mock GW events, in which hierarchical mergers are assumed to be frequent. In Table A1, we list the model varieties we have investigated (models M1–M28). We additionally examine different choices of the number of detected mergers (models M14 and M15), the steps to create samples for hierarchical mergers (models M16–M19), pairing probability (models M20 and M21), the fraction of mergers in each step (models M22–M24), the escape velocity of the system (model M25), the power law for mass function (model M26), and the correlation between the steps and the redshift (models M27 and M28).

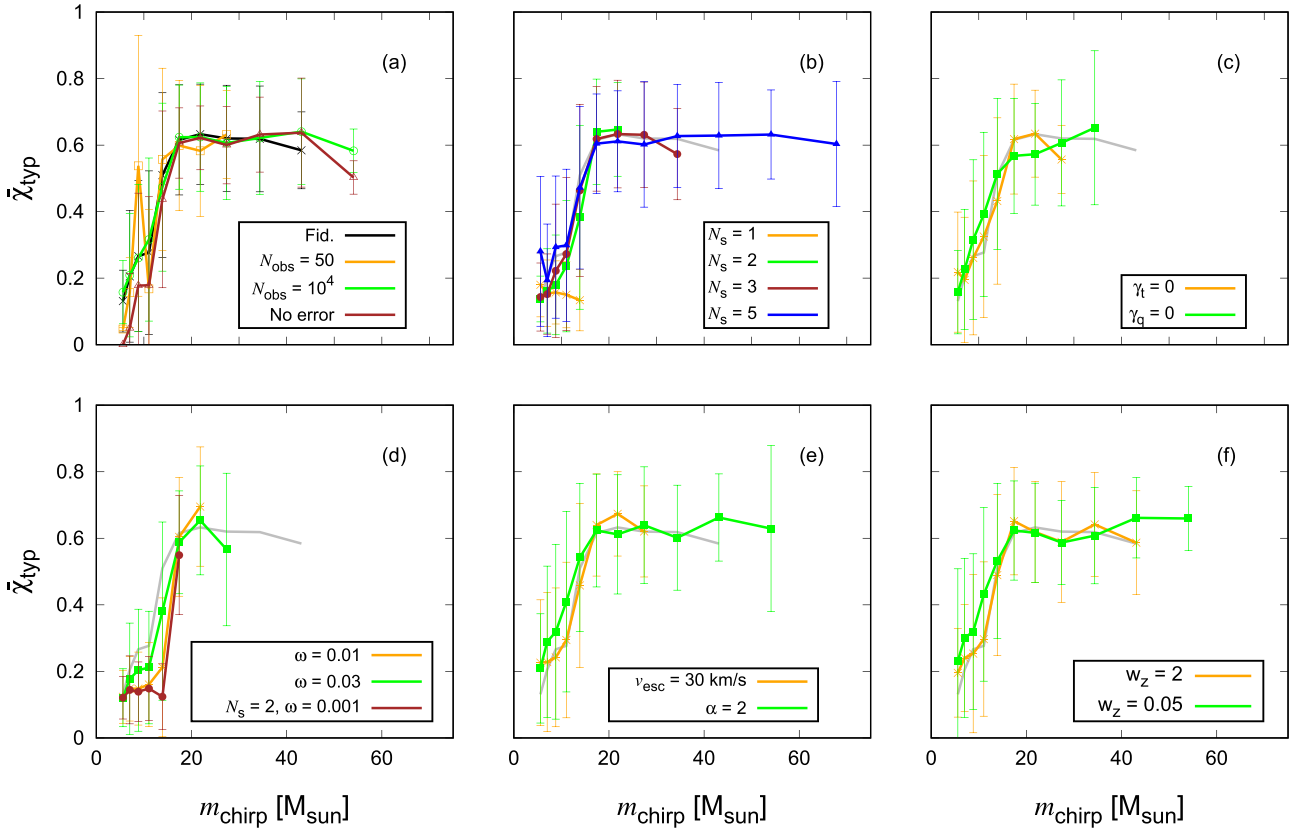

Figure A1. Same as Fig. 1, but present for models M1, M14–M28 (Table 2). We use $N_{\rm obs} = 10^3$ detectable mergers for models M1, M14–M28, while $N_{\rm obs} =$ 50 and 10^4 for models M14 and M15, respectively.

Table A1. Same as Table 2, but includes models M14–M28.

Model	Parameter	High-g fraction	High-g detection fraction	$m_{\rm chirp,\,max}[{\rm M}_\odot]$	$\bar{\chi}_{\rm p}$	$\sigma(\chi_{\rm p})$
M1	Fiducial	0.33	0.68	56	0.17	0.26
M2	Globular cluster (GC)	0.063	0.17	44	0.030	0.13
M3	Field binary (FB)	0	0	23	0	0
M4	Migration trap (MT)	0.31	0.80	42	0	0
M5	$a_{\rm uni}=1$	0.32	0.73	52	0.50	0.21
M6	$a_{\rm ave}=0.99$	0.31	0.70	51	0.75	0.20
M7	$a_{\rm ave}=0.66$, $a_{\rm uni}=0.1$	0.33	0.72	55	0.55	0.13
M8	$a_{\rm ave}=0.5$	0.33	0.74	65	0.46	0.12
M9	$m_{\rm max}=30\,{\rm M}_\odot$	0.35	0.73	70	0.18	0.26
M10	$N_{\rm obs}=50$, $N_{\rm s}=3$	0.25	0.62	28	0.13	0.24
M11	$N_{\rm obs}=50$, $N_{\rm s}=2$	0.15	0.28	24	0.077	0.19
M12	$N_{\rm obs}=50$, $N_{\rm s}=2$, $\omega=0.05$	0.077	0.18	19	0.040	0.14
M13	$N_{\rm obs}=50$, $N_{\rm s}=2$, $\omega=0.03$	0.046	0.14	19	0.023	0.11
M14	$N_{\rm obs}=50$	0.33	0.78	38	0.17	0.26
M15	$N_{\rm obs}=10^4$	0.33	0.77	60	0.17	0.26
M16	$N_{\rm s}=1$	0	0	15	0	0
M17	$N_{\rm s}=2$	0.15	0.40	23	0.077	0.19
M18	$N_{\rm s}=3$	0.25	0.61	42	0.13	0.24
M19	$N_{\rm s}=5$	0.38	0.80	72	0.20	0.27
M20	$\gamma_{\rm t}=0$	0.22	0.50	31	0.11	0.22
M21	$\gamma_{\rm q}=0$	0.39	0.71	36	0.19	0.25
M22	$\omega=0.01$	0.043	0.089	23	0.022	0.11
M23	$\omega=0.03$	0.13	0.37	33	0.066	0.18
M24	$N_{\rm s}=2$, $\omega=0.001$	0.0014	0.0040	17	0.001	0.02
M25	$v_{\rm esc}=30\,{\rm km\,s^{-1}}$	0.29	0.61	31	0.15	0.25
M26	$\alpha=2$	0.32	0.75	52	0.16	0.25
M27	$w_z=2$	0.33	0.73	46	0.17	0.26
M28	$w_z=0.05$	0.33	0.79	59	0.17	0.26

With smaller number of iteration steps ($N_{\rm s}$), the maximum $m_{\rm chirp}$ becomes smaller because the generations of BHs are limited by $N_{\rm s}$ (panel b of Fig. A1). Similarly, the maximum $m_{\rm chirp}$ decreases as $N_{\rm obs}$, $\gamma_{\rm t}$, ω, or $v_{\rm esc}$ decreases or $m_{\rm max}$ increases (panels a, c, d, and e of Fig. A1 and panel b of Fig. 1; Table A1). In these ways, the maximum $m_{\rm chirp}$ is influenced by a number of parameters, implying that the maximum $m_{\rm chirp}$ alone cannot constrain each of those parameters.

Here, we investigate the effect that mergers at larger iteration steps tend to occur at lower redshift because finite time needs to elapse between each generation and high-g mergers thus would take place after a significant delay compared to low-g mergers. To take this delay into account, we modify the redshift distribution of merging BHs as

$$p_z \propto \frac{{\rm d}V_{\rm c}}{{\rm d}z}\frac{1}{1+z}\exp\left(\frac{(t_{\rm L}(z)-\mu_t)^2}{2\sigma_t^2}\right), \tag{A1}$$

where $t_{\rm L}(z)$ is the look-back time, we set the average to $\mu_t = t_{\rm typ}\left(\frac{N_{\rm S}-N_i+1}{N_{\rm s}}\right)$ and the standard deviation to $\sigma_t = t_{\rm typ}w_t$, N_i is the number of steps that the ith merger is created, $t_{\rm typ}$ is the typical look-back time that mergers began to occur, which is set to 10 Gyr, and w_z is the parameter determining the strength of correlation between N_i and the time that mergers occur. A lower value of w_z makes mergers with high N_i occur at a lower z, and the fiducial model (equation 5) corresponds to $w_z=\infty$. The dependence of the $\bar{\chi}_{\rm typ}$ profile on w_z is shown in panel (f), suggesting that the correlation between the redshift and the generations of BHs has a negligible impact on the profile.

APPENDIX B: OBSERVED SPIN DISTRIBUTION

We presents the observed distributions of $\chi_{\rm p}$, $\chi_{\rm eff}$, and $\chi_{\rm typ}$ as a function of $m_{\rm chirp}$ in Fig. B1. Also, Fig. B2 compares the $m_{\rm chirp}$, $\chi_{\rm eff}$ and $\chi_{\rm p}$ distributions observed by LIGO/Virgo O1–O3a and those predicted by the model for the fiducial settings (Table 1), but $m_{\rm max}=30\,{\rm M}_\odot$ and $\alpha=2$, which is assessed to high Bayes factors for both $D_{\rm KL,\,crit}=0.05$ and the two parameters (Table 4). We can see that the observed distribution for these variables (blue and orange points) roughly follows the 90 and 99 percentile regions (black and grey lines) predicted by the model.

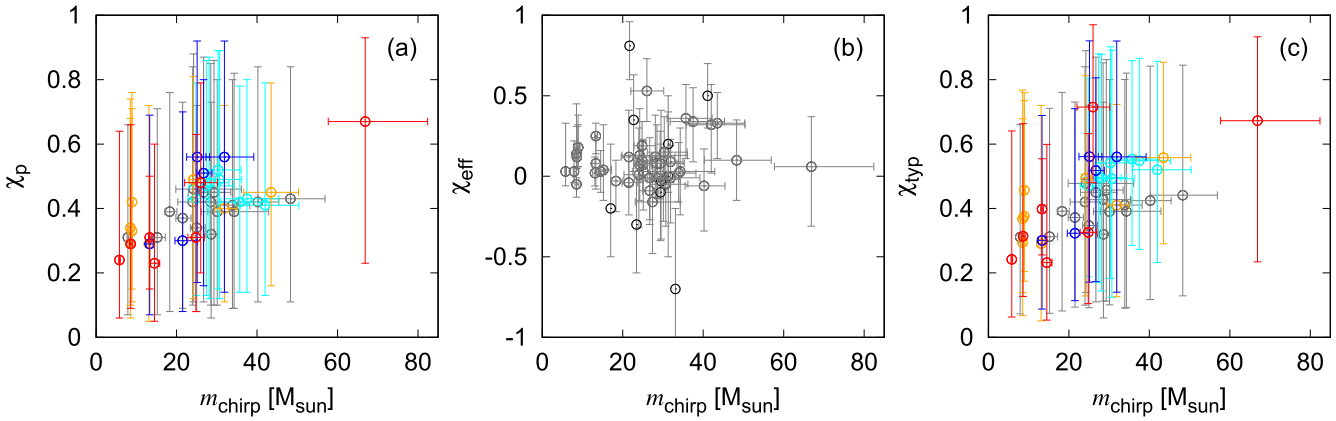

Figure B1. Observed distribution for spins reported in LIGO/Virgo O1–O3a. Panels (a)–(c) represent the distributions of χ_p, χ_{eff}, and χ_{typ}, respectively. Red, blue, orange, cyan, and grey circles represent events with the KL divergence between prior and posterior samples for χ_p to be $D_{KL} \geq 0.2$, $0.2 > D_{KL} \geq 0.15$, $0.15 > D_{KL} \geq 0.1$, $0.1 > D_{KL} \geq 0.05$, and $0.05 > D_{KL}$, respectively. Bars correspond to the 90 percentile credible intervals.

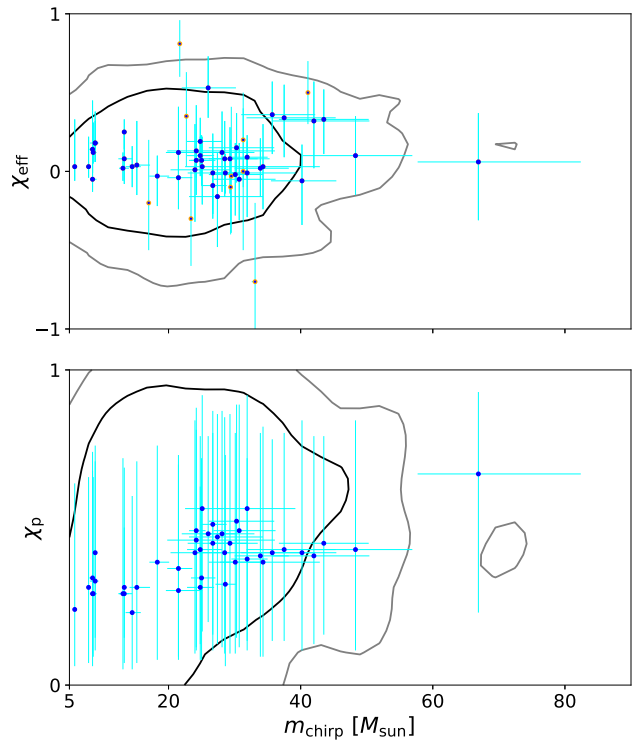

Figure B2. Comparisons between the χ_{eff} (upper) or χ_p (lower) and m_{chirp} distributions observed by LIGO/Virgo O1–O3a and those predicted by the model with $m_{max} = 30\,M_\odot$ and $\alpha = 2$ which have high Bayes factors (Table 4). Black and grey lines correspond to 90 and 99 credible intervals for the predicted distributions, and cyan bars correspond to the 90 percentile credible intervals for the observed variables. The orange points in the upper panel corresponds to the events reported in Venumadhav et al. (2020) and Zackay et al. (2019).

APPENDIX C: POSTERIOR DISTRIBUTIONS FOR SPIN PARAMETERS

We present the posterior distributions of the parameters characterizing the spin profile for the GW events with $D_{KL} \geq 0$ (Section 3.2.1) in Fig. C1.

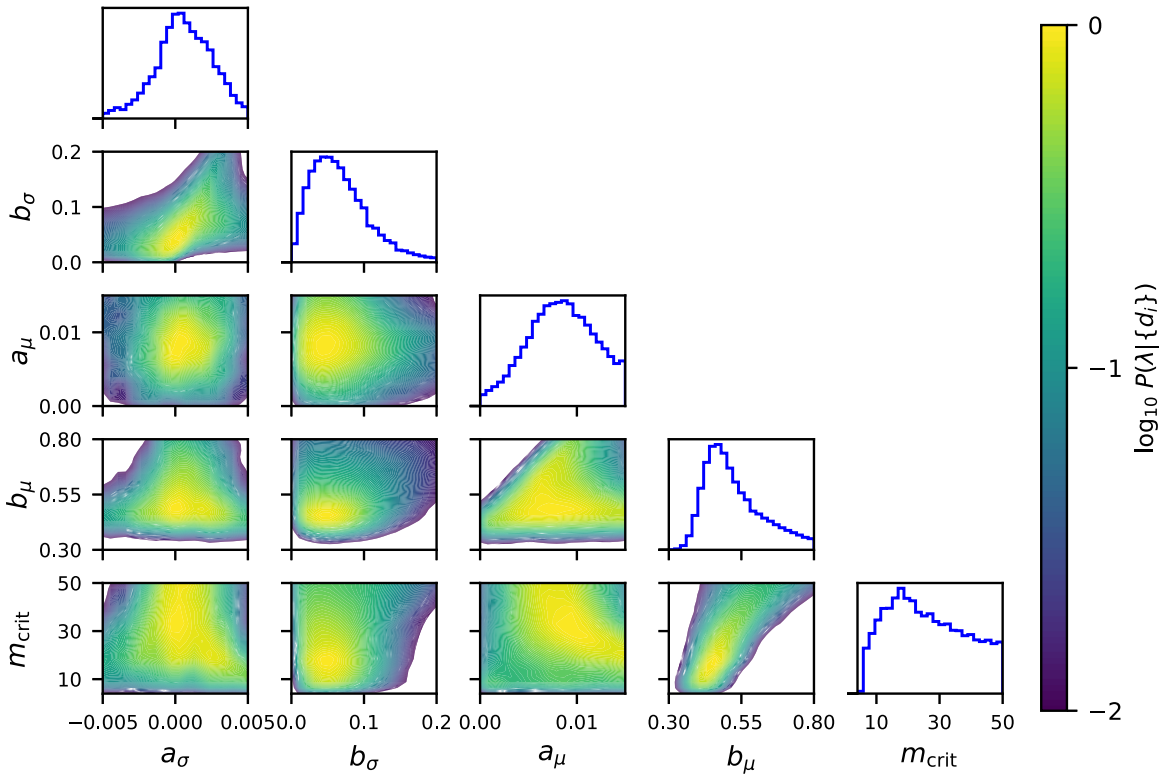

Figure C1. The posterior probability distributions for a_σ, b_σ, a_μ, b_μ, and $m_{\rm crit}$ for $D_{\rm KL,\,crit} = 0$ (Section 3.2.1).

This paper has been typeset from a TEX/LATEX file prepared by the author.