

Constraining the fraction of binary black holes formed in isolation and young star clusters with gravitational-wave data

YANN BOUFFANAIS,<sup>1,2</sup> MICHELA MAPELLI,<sup>1,2,3,4</sup> DAVIDE GEROSA,<sup>5</sup> UGO N. DI CARLO,<sup>1,2,6</sup> NICOLA GIACOBBO,<sup>1,2,3</sup>  
EMANUELE BERTI,<sup>7</sup> AND VISHAL BAIBHAV<sup>7</sup>

<sup>1</sup>*Dipartimento di Fisica e Astronomia ‘G. Galilei’, University of Padova, Vicolo dell’Osservatorio 3, I-35122, Padova, Italy*

<sup>2</sup>*INFN, Sezione di Padova, Via Marzolo 8, I-35131, Padova, Italy*

<sup>3</sup>*INAF–Osservatorio Astronomico di Padova, Vicolo dell’Osservatorio 5, I-35122, Padova, Italy*

<sup>4</sup>*Institut für Astro- und Teilchenphysik, Universität Innsbruck, Technikerstrasse 25/8, A-6020, Innsbruck, Austria*

<sup>5</sup>*School of Physics and Astronomy and Institute for Gravitational Wave Astronomy, University of Birmingham, Birmingham, B15 2TT, UK*

<sup>6</sup>*Dipartimento di Scienza e Alta Tecnologia, University of Insubria, Via Valleggio 11, I-22100, Como, Italy*

<sup>7</sup>*Department of Physics and Astronomy, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, US*

Submitted to ApJ

## ABSTRACT

Ten binary black-hole mergers have already been detected during the first two observing runs of advanced LIGO and Virgo, and many more are expected to be observed in the near future. This opens the possibility for gravitational-wave astronomy to better constrain the properties of black hole binaries, not only as single sources, but as a whole astrophysical population. In this paper, we address the problem of using gravitational-wave measurements to estimate the proportion of merging black holes produced either via isolated binaries or binaries evolving in young star clusters. To this end, we use a Bayesian hierarchical modeling approach applied to catalogs of merging binary black holes generated using state-of-the-art population synthesis and N-body codes. In particular, we show that, although current advanced LIGO/Virgo observations only mildly constrain the mixing fraction  $f \in [0, 1]$  between the two formation channels, we expect to narrow down the fractional errors on  $f$  to 10 – 20% after a few hundreds of detections.

*Keywords:* black hole physics – gravitational waves – Bayesian analysis

## 1. INTRODUCTION

The first two observing runs of the LIGO/Virgo collaboration (LVC) led to the detection of ten binary black holes (BBHs, Abbott et al. 2016, 2018a) and one binary neutron star (BNS, Abbott et al. 2017a,b). The next observing run will significantly boost this sample: several tens of new BBH detections and few BNS detections are expected in the coming months. The growing sample of merging BBHs is expected to provide key information on their mass, spin and local merger rate (Abbott et al. 2018b).

One of the main open questions about BBHs concerns their formation channel(s). Several possible scenarios have been proposed in the last decades. The *isolated evolution* of a massive binary star can lead to the formation of a merging BBH through a common envelope episode (Bethe & Brown 1998; Belczynski et al. 2002, 2014, 2016a; Dominik et al. 2013; Mennekens & Vanbevereren 2014; Spera et al. 2015; Eldridge & Stanway 2016; Eldridge et al. 2017; Mapelli et al. 2017; Mapelli & Giacobbo 2018; Stevenson et al. 2017b; Giacobbo & Mapelli 2018; Kruckow et al. 2018; Spera et al. 2019; Mapelli et al. 2019; Eldridge et al. 2019) or via chemically homogeneous evolution (Marchant et al. 2016; Mandel & de Mink 2016).

Alternatively, several *dynamical processes* can trigger the formation of a BBH and influence its subsequent evolution to the final merger (see Mapelli 2018 for a recent

yann.bouffanais@pd.infn.it

michela.mapelli@unipd.it

review on the subject). For example, the Kozai-Lidov dynamical mechanism (Kozai 1962; Lidov 1962) might significantly affect the formation of eccentric BBHs in triple stellar systems (e.g. Antonini & Perets 2012; Antonini & Rasio 2016; Kimpson et al. 2016; Antonini et al. 2017). Similarly, dynamical exchanges and three- or multi-body scatterings are expected to lead to the formation and dynamical hardening of BBHs in dense stellar systems, such as globular clusters (Portegies Zwart & McMillan 2000; O’Leary et al. 2006; Sadowski et al. 2008; Downing et al. 2010, 2011; Rodriguez et al. 2015, 2016a,b; Rodriguez & Loeb 2018; Askar et al. 2017; Samsing 2018; Samsing et al. 2018; Fragione & Kocsis 2018), nuclear star clusters (O’Leary et al. 2009; Antonini & Perets 2012; Antonini & Rasio 2016; Petrovich & Antonini 2017; Stone et al. 2017b,a; Rasskazov & Kocsis 2019) and young star clusters (Banerjee et al. 2010; Mapelli et al. 2013; Ziosi et al. 2014; Mapelli 2016; Banerjee 2017, 2018; Di Carlo et al. 2019; Kumamoto et al. 2019). Other formation mechanisms include black hole (BH) pairing in extreme gaseous environments (like AGN disks, e.g. McKernan et al. 2012, 2014, 2018; Bartos et al. 2017; Tagawa & Umemura 2018). Finally, primordial BHs of non-stellar origin may form binaries through dynamical processes (Carr & Hawking 1974; Carr et al. 2016; Sasaki et al. 2016; Bird et al. 2016; Inomata et al. 2017; Inayoshi et al. 2016; Scelfo et al. 2018).

Each formation channel leaves its specific imprint on the properties of BBHs. In particular, dynamically formed BBHs are expected to have larger masses than isolated BBHs (e.g. Di Carlo et al. 2019), because dynamical exchanges favour the formation of more massive binaries (Hills & Fullerton 1980). Several evolutionary processes in isolated binary systems (tides, mass transfer) tend to align the individual spins with the orbital angular momentum of the binary, while only supernova kicks can tilt the spins significantly in isolated binaries (Kalogera 2000; Gerosa et al. 2013, 2018; O’Shaughnessy et al. 2017). In contrast, dynamical exchanges are expected to reset any memory of previous alignments; thus, dynamically formed BBHs are expected to have isotropically oriented spins. Finally, dynamically formed BBHs (especially Kozai-Lidov triggered systems) might develop larger eccentricities than isolated BBHs. Eccentricities are larger and easier to measure at the low frequencies accessible to space-based interferometers such as LISA (Nishizawa et al. 2016; Breivik et al. 2016; Nishizawa et al. 2017), but in some cases they may be significant even in the advanced LIGO (aLIGO) and advanced Virgo (aVirgo) band (e.g. Antonini et al. 2017; Zevin et al. 2019).

Thus, BH masses, spins and eccentricities are key features to differentiate between binary formation channels. To achieve this goal, BBH populations predicted by models should be contrasted with gravitational-wave (GW) data, by means of a suitable model-selection framework. Several methodological approaches can be found in the literature (Stevenson et al. 2015, 2017a; Gerosa & Berti 2017; Vitale et al. 2017b; Zevin et al. 2017; Talbot & Thrane 2017, 2018; Taylor & Gerosa 2018; Abbott et al. 2018b; Fishbach et al. 2017; Fishbach et al. 2018; Wysocki et al. 2018; Roulet & Zaldarriaga 2019; Kimball et al. 2019). For example, Stevenson et al. (2017a) use of a hierarchical analysis in order to combine multiple GW observations of BBH spin-orbit misalignments, to give constraints on the fractions of BBHs forming through different channels. Similarly, Zevin et al. (2017) apply a hierarchical Bayesian model to mass measurements from mock GW observations. They compare populations obtained with isolated binary evolution and with Monte Carlo simulations of globular clusters and show that they can distinguish between the two channels with  $\mathcal{O}(100)$  GW observations. Taylor & Gerosa (2018) use banks of compact-binary population synthesis simulations to train a Gaussian-process emulator that acts as a prior on observed parameter distributions (e.g. chirp mass, redshift, rate). Based on the results of the emulator, a hierarchical population inference framework allows to extract information on the underlying astrophysical population. Alternative approaches consist in model-independent inference based on clustering of source parameters (e.g. Mandel et al. 2015, 2017; Powell et al. 2019).

Here, we follow a standard Bayesian model-selection approach (cf. e.g. Sesana et al. 2011; Gair et al. 2011; Gerosa & Berti 2017), properly including selection effects (Mandel et al. 2019) and posterior distributions, exploiting both full aLIGO/aVirgo data and mock samples for future forecasts. As for the astrophysical models, we compare BBHs from isolated binary evolution with dynamically formed BBHs. For the first time, we apply model selection to dynamically formed BBHs in young star clusters (Di Carlo et al. 2019). Young star clusters are intriguing dynamical environments for BBHs, because most massive stars (which are BH progenitors) form in young star clusters in the nearby Universe (Lada & Lada 2003; Portegies Zwart et al. 2010). On the other hand, simulating BBHs in young star clusters has a high computational cost, because it requires direct N-body simulations combined with binary population synthesis. Both isolated binaries and dynamically formed ones are evolved through the MOBSE population-synthesis code (Giacobbo et al. 2018), which includes state-of-the-art

modelling of stellar winds, supernova prescriptions and binary evolution.

## 2. DISTRIBUTIONS OF ASTROPHYSICAL SOURCES

### 2.1. Isolated formation channel of BBHs

We simulate isolated BBHs using the binary population-synthesis code MOBSE (Mapelli et al. 2017; Giacobbo et al. 2018). MOBSE includes single stellar evolution through polynomial fitting formulas as described in Hurley et al. (2000) and binary evolution processes (mass transfer, tidal evolution, common envelope, GW decay, etc.) as described in Hurley et al. (2002). The main differences between MOBSE and BSE are the following (cf. Giacobbo & Mapelli 2018 for additional details).

Mass loss by stellar winds of massive hot stars (O- and B-type stars, luminous blue variables and Wolf-Rayet stars) is implemented in MOBSE as  $\dot{M} \propto Z^\eta$  (Chen et al. 2015, and references therein), where  $Z$  is the stellar metallicity and

$$\eta = \begin{cases} 0.85, & \text{if } \Gamma_e < 2/3 \\ 2.45 - 2.4 \Gamma_e & \text{if } 2/3 \leq \Gamma_e \leq 1 \\ 0.05 & \text{if } \Gamma_e > 1, \end{cases} \quad (1)$$

where  $\Gamma_e = L_*/L_{\text{Edd}}$  is the Eddington factor,  $L_*$  is the current stellar luminosity, and  $L_{\text{Edd}}$  is the Eddington luminosity (Gräfenr et al. 2011). The mass of a compact object depends on the final mass and core mass of the progenitor star through fitting formulas which describe the outcome of electron-capture supernovae (see Giacobbo & Mapelli 2019), core-collapse supernovae (see Fryer et al. 2012) and pair-instability or pulsational pair-instability supernovae (see Spera & Mapelli 2017). In this paper, we adopt the delayed model for core collapse supernovae (see Fryer et al. 2012). These prescriptions enable us to obtain a BH mass distribution which is consistent with GW data from the first and second observational runs of aLIGO and aVirgo (Abbott et al. 2018a,b).

The natal kick of a neutron star is drawn from a Maxwellian distribution with 1-dimensional root-mean square  $\sigma = 15$  and  $265 \text{ km s}^{-1}$  for an electron-capture and a core-collapse supernova, respectively (see Hobbs et al. 2005 and Giacobbo & Mapelli 2019 for more details). The natal kick of a BH is calculated as  $v_{\text{BH}} = v_{\text{NS}} (1 - f_{\text{fb}})$ , where  $v_{\text{NS}}$  is a random number extracted from the same Maxwellian distribution as neutron stars born from core-collapse supernovae, while  $f_{\text{fb}}$  is the fraction of mass that falls back to a BH, estimated as in Fryer et al. (2012).

In this paper, we consider a sample of  $10^7$  binaries simulated with MOBSE with metallicity  $Z = 0.002 \simeq Z_\odot/10$

(the effect of varying the metallicity will be tackled in a forthcoming publication).

The primary mass is randomly drawn from a Kroupa (2001) initial mass function between  $m_1 = 5 M_\odot$  and  $150 M_\odot$ , while the secondary is randomly drawn from the mass ratio  $q$

$$\mathfrak{F}(q) \propto q^{-0.1} \quad \text{with } q = \frac{m_2}{m_1} \in [0.1 - 1]. \quad (2)$$

As suggested by observations (Sana et al. 2012), initial orbital periods  $P$  and eccentricities  $e$  are randomly drawn from

$$\mathfrak{F}(\mathcal{P}) \propto (\mathcal{P})^{-0.55} \quad \text{with } \mathcal{P} = \log_{10}(P/\text{day}) \in [0.15 - 5.5], \quad (3)$$

$$\mathfrak{F}(e) \propto e^{-0.42} \quad \text{with } 0 \leq e < 1. \quad (4)$$

For this paper we adopted common-envelope ejection efficiency  $\alpha = 3$ , while the envelope concentration  $\lambda$  is derived by MOBSE as described by Claeys et al. (2014).

From these population-synthesis simulations we obtain 31879 BBH mergers (hereafter referred to as “isolated BBHs”) which merge within a Hubble time  $t_{\text{H}} = 14 \text{ Gyr}$ .

### 2.2. Dynamical formation channel of BBHs in young star clusters

The dynamically formed BBHs were obtained by means of direct N-body simulations with NBODY6++GPU (Wang et al. 2015) coupled to MOBSE (Di Carlo et al. 2019). We have, therefore, the very same population-synthesis recipes in both the isolated binary simulations and the dynamical simulations.

The initial conditions were obtained with McLUSTER (Küpper et al. 2011). The distributions of dynamical BBHs discussed in this paper are obtained from 4000 simulations of young star clusters with fractal initial conditions. We chose to simulate young star clusters because most stars (and especially massive stars) are thought to form copiously in these environments (e.g. Lada & Lada 2003; Portegies Zwart et al. 2010). The assumption of fractal initial conditions mimics the clumpiness and asymmetry of observed star forming regions (e.g. Gutermuth et al. 2005). Each star cluster’s mass was randomly drawn from a distribution  $dN/dM_{\text{SC}} \propto M_{\text{SC}}^{-2}$ , consistent with the observed mass function of young star clusters in the Milky Way (Lada & Lada 2003). Thus, our simulated star clusters represent a synthetic young star-cluster population of Milky Way-like galaxies.

The initial binary fraction in each star cluster is  $f_{\text{bin}} = 0.4$ . While observed young star clusters can have larger values of  $f_{\text{bin}}$  (up to  $\sim 0.7$ , Sana et al. 2012),  $f_{\text{bin}} = 0.4$  is close to the maximum value considered in

state-of-the-art simulations, being  $f_{\text{bin}}$  the bottleneck of direct N-body simulations. Initial stellar and binary masses, orbital periods and orbital eccentricities are generated as described in Sec. 2.1, to guarantee a fair comparison. For the same reason, all the simulated star clusters have stellar metallicity  $Z = 0.002$ , the same as isolated binaries.

Each star cluster feels the tidal field of a Milky Way-like galaxy and is assumed to be on a circular orbit with radius similar to the Sun’s orbital radius. Star clusters are simulated for  $\sim 100$  Myr, corresponding to a conservative assumption for the lifetime of a young star cluster. We refer to Di Carlo et al. (2019) for a more detailed discussion of our dynamical models and assumptions.

From these dynamical simulations we obtain 229 BBHs (hereafter dynamical BBHs) which merge within a Hubble time  $t_{\text{H}} = 14$  Gyr. We stress that dynamical simulations are computationally more expensive than population-synthesis runs and our sample of merging BBHs is one of the largest ever obtained from direct N-body simulations with realistic binary evolution.

Dynamical BBHs belong to two families. About 47% of all merging BBHs in the simulated star clusters come from original binaries (hereafter, original BBHs), i.e. they form from the evolution of stellar binaries which were already present in the initial conditions. Such original binaries evolve in a star cluster, thus they are affected by close-by encounters with other stars (which can change their semi-major axis and eccentricity), but otherwise behave similarly to BBHs formed in isolation.

The remaining 53% of all merging BBHs in the simulated star clusters form via dynamical exchanges (hereafter, exchanged BBHs). Dynamical exchanges are three-body encounters between a binary system and a single object, during which the single object exchanges with one of the members of the binary system. BHs are tremendously efficient in acquiring companions through dynamical exchanges (Ziosi et al. 2014), because they are more massive than other stars in star clusters: exchanges favour the formation of more massive binaries (which are more energetically stable, Hills & Fullerton 1980). Because of their formation mechanism, exchanged BBHs are significantly more massive than both isolated BBHs and original BBHs (cf. Di Carlo et al. 2019).

### 2.3. Treatment of spins and redshift

#### 2.3.1. Spins

The initial magnitude and direction of BH spins is still a matter of debate (see e.g. Miller & Miller 2015 for a review). Overall, the dependence of the spin magnitude

of a BH on the spin magnitude of the progenitor star (or stellar core) is largely unknown.

As for the direction of the spins, most binary evolution processes in isolated binaries tend to favour the alignment of stellar spins with the orbital angular momentum of the binary. In isolated binaries, supernova kicks are the leading mechanism to substantially tilt the spin axes with respect to the orbital plane (Kalogera 2000). In star clusters, dynamical exchanges tend to reset the memory of the initial binary spin. Thus, we expect the spins of exchanged BBHs to be isotropically distributed. The spin direction of original BBHs in star clusters is expected to fall somewhere in between because, on the one hand, these BBHs participate in the dynamical evolution of the star cluster (thus, dynamical encounters can affect the initial spin orientation), while on the other hand they form from the evolution of stellar binaries (thus, binary evolution processes tend to realign the spins).

Model	Rms	Orientation	BBH sample
LSA	0.1	aligned	Isolated BBHs, Original BBHs
LSI	0.1	isotropic	Exchanged BBHs
HSA	0.3	aligned	Isolated BBHs, Original BBHs
HSI	0.3	isotropic	Exchanged BBHs

**Table 1.** Summary of our spin models. We implement two prescriptions (L: low; H: high) of the spin magnitude by varying the root mean square of their Maxwellian distribution, and two prescriptions for the spin orientations (A: aligned; I: isotropic). Exchanged BBHs are always assumed to have isotropic spins (LSI or HSI), while isolated BBHs and original BBHs are assumed to have aligned spins (LSA or HSA).

Given these considerable uncertainties on both spin magnitude and direction, we decided not to embed detailed spin models in our population synthesis and dynamical simulations. Spins are added to our simulations in post-processing, assuming simple toy models. Dimensionless spin magnitudes  $a$  (defined as  $a = |J|c/Gm_{\text{BH}}^2$ , where  $J$  is the BH spin,  $c$  is the speed of light,  $G$  is the gravitational constant and  $m_{\text{BH}}$  is the mass of the BH) are randomly drawn from a Maxwellian distribution with root mean square equal to 0.1. With this choice, the median spin is  $a \sim 0.15$  and the distribution quickly fades off for  $a > 0.4$ . Hereafter, we refer to this model as “low-spin” (LS). We assume this distribution for spin magnitudes because the results of the first two aLIGO/aVirgo runs disfavour distributions with large spin components aligned (or nearly aligned) with the orbital angular momentum (Abbott et al. 2018b).



For comparison, we also consider a second rather extreme case, in which spin magnitudes are drawn from a Maxwellian distribution with root mean square equal to 0.3 (we reject spin magnitudes  $a > 0.998$ ). With this choice, the median spin is  $a \sim 0.46$ . Hereafter, we refer to this model as “high-spin” (HS).

Regarding spin orientations, we assume that BH spins in both isolated BBHs and original BBHs are perfectly co-aligned with the orbital angular momentum of the binary (see e.g. [Rodríguez et al. 2016c](#)): there are large uncertainties on the kicks imparted on newly formed BHs, but recent work ([Gerosa et al. 2018](#)) shows that the fraction of BBHs with negative effective spins is at most  $\sim 20\%$ . Moreover, we neglect the effect of dynamical perturbations on original BBHs because their main properties are similar to isolated BBHs: they have nearly the same chirp mass, mass ratio and eccentricity distribution ([Di Carlo et al. 2019](#)).

Finally, BH spins in exchanged BBHs are randomly drawn isotropically over a sphere. Our assumptions for the spin models are summarized in Table 1.

### 2.3.2. Redshifts

The redshift parameter was not computed self-consistently in the set of astrophysical simulations generated for this study, because we only consider a fixed value for the metallicity, and the stellar metallicity is a crucial ingredient of redshift evolution ([Mapelli et al. 2017](#)). A self-consistent redshift evolution will be included in future work ([Baibhav et al.](#), in preparation).

Here we opted for excluding redshift information from our statistical analysis (cf. Sec. 4). However, we still need to prescribe a redshift probability distribution function to estimate selection effects. As a simple toy model, we assume that the redshift is distributed uniformly in comoving volume and source-frame time, i.e.

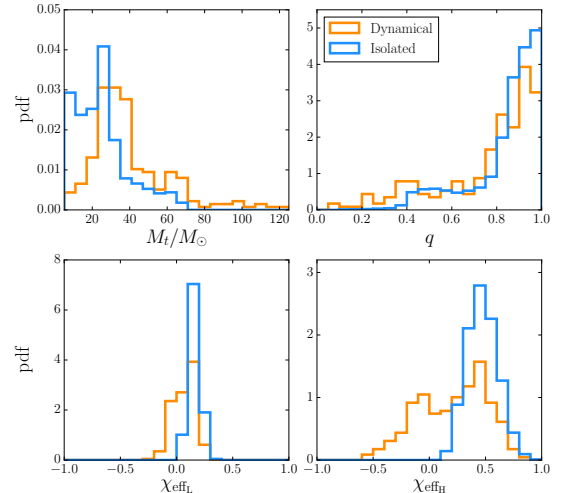
$$p(z) \propto \frac{1}{1+z} \frac{dV_c}{dz}. \quad (5)$$

We consider redshifts in the range  $z \in [0, 2]$  for both second- and third-generation detectors, postponing more accurate modelling to future work.

### 2.4. Catalog distributions

In Figure 1, we present our distributions from both dynamical (orange) and isolated (blue) catalogs. We plot distributions corresponding to total mass ( $M_t$ ), mass ratio ( $q$ ) and effective spins for both the LS ( $\chi_{\text{eff}_L}$ ) and HS ( $\chi_{\text{eff}_H}$ ) models.

The isolated model allows for BBHs with total mass in the range  $M_t \in [5, 70]M_\odot$ . On the other hand, the dynamical case presents massive BBHs with  $M_t > 70M_\odot$



**Figure 1.** Astrophysical population of merging BBHs from dynamical (orange) and isolated (blue) formation channels as presented in Section 2. We show distributions of total mass  $M_t$ , mass ratio  $q$  and effective spin parameters for low-spin ( $\chi_{\text{eff}_L}$ ) and high-spin ( $\chi_{\text{eff}_H}$ ) cases.

formed via dynamical interactions (exchanged BBHs). The dynamical model predicts BBHs with  $q < 0.2$ , which are not present in the isolated BBH catalogs. The physical reasons for the difference between the maximum mass of isolated BBHs and dynamical BBHs are thoroughly explained in [Di Carlo et al. \(2019\)](#). Here we summarize the main ingredients. First, BBHs with  $M_t > 70M_\odot$  form even in our isolated binaries, but they are too wide to merge within a Hubble time (see [Giacobbo et al. 2018](#)). In the dynamical simulations, these massive wide BBHs have a chance to shrink by dynamical interactions and to become sufficiently tight to merge within a Hubble time. Secondly, massive single BHs (with mass  $\gg 30 M_\odot$ ) form from collisions between stars (especially if one of the two colliding stars has already developed a Helium core). If these massive single BHs are in the field, they likely remain alone, while if they form in the core of a star cluster, they are very efficient in acquiring new companions through exchanges.

By construction, the isolated scenario only contains binaries with  $\chi_{\text{eff}} > 0$ , while dynamically-formed BHs are found with both positive and negative values for  $\chi_{\text{eff}}$  (with a preference for positive values).

## 3. STATISTICAL ANALYSIS

### 3.1. GW data analysis

#### 3.1.1. Detection probability

We estimate selection effects using the semi-analytic approach of [Finn & Chernoff 1993](#) (cf. also [Dominik et al. 2015](#); [Chen et al. 2017](#); [Taylor & Gerosa 2018](#)). We associate a detection probability  $p_{\text{det}}(\lambda) \in [0, 1]$  to

any given GW source with parameters  $\lambda$ . A source is detectable if its signal-to-noise ratio (SNR)

$$\rho = 4 \int_0^{+\infty} \frac{|\tilde{h}(f)|^2}{S_n(f)} df \quad (6)$$

exceeds a given threshold  $\rho_{th}$ , with  $\tilde{h}(f)$  being the gravitational waveform in the frequency domain and  $S_n(f)$  the one-sided noise power spectral density of the detector. To compute  $\tilde{h}(f)$ , we have used the IMRPhenomD model (Khan et al. 2016), that is a phenomenological waveform model describing the inspiral, merger and ringdown of a non-precessing BBH merger signal. We consider the noise power spectral density curves corresponding to both the aLIGO (Abbott et al. (2018)) and the Einstein Telescope (ET, Abbott et al. (2017)) at their design sensitivity. We implement a single-detector SNR threshold  $\rho_{th} = 8$ , which was shown to be a good approximation of more complex multi-detector analysis based on large injection campaigns (see Abadie et al. 2010; Abbott et al. 2016; Wysocki et al. 2018 for more details). Both waveforms and detector sensitivities were generated using PYCBC (Dal Canton et al. 2014; Usman et al. 2016).

For each binary in our catalogs, we estimate the optimal SNR,  $\rho_{opt}$ , using Eq. (6). This corresponds to a face-on source located overhead with respect to the detector. The SNR of a generic source is given by  $\rho = \omega \times \rho_{opt}$ , where  $\omega$  encapsulates all the dependencies on sky-location, inclination and polarization angle (Finn & Chernoff 1993; Finn 1996). A source located in a blind spot of the detector yields a value of  $\omega = 0$ , while an optimally oriented source has  $\omega = 1$ . The probability of detecting a source is then expressed as

$$p_{det}(\lambda) = \mathbb{P}(\rho \geq \rho_{thr}) \quad (7)$$

$$= \mathbb{P}(\omega \geq \rho_{thr}/\rho_{opt}) \quad (8)$$

$$= 1 - F_\omega(\rho_{thr}/\rho_{opt}), \quad (9)$$

where  $F_\omega$  is the cumulative distribution function of  $\omega$ . This function was computed via Monte Carlo methods as implemented in the python package GWDET (Gerosa 2019). The function  $F_\omega$  is set explicitly to 1 for  $\rho_{opt} < \rho_{thr}$ , which gives the expected detection probability  $p_{det}(\lambda) = 0$  for events which are too quiet to be observed.

### 3.1.2. Measurement errors

The noise contained in the data  $d$  of a GW detector results in errors on the measurement of the parameters  $\lambda$  of a GW source. From a Bayesian point of view, these errors are fully described by the posterior distribution  $p(\lambda|d)$ . We make use of posterior distributions for the first 10 GW events publicly released by Abbott et al. (2018a).

We also generate mock observations from our catalogs to forecast future scenarios with a growing number of events. In this case, running a full injection campaign to estimate measurement errors would be computationally too expensive and out of the scope of this study. For simplicity, we approximate posterior distributions with simple Gaussians (Gerosa & Berti 2017; Farr et al. 2017)

$$p(\lambda^i) = \mathcal{N}(\bar{\lambda}^i, \sigma^i), \quad (10)$$

where  $\bar{\lambda}^i$  is the value for the parameter  $\lambda^i$  extracted from our astrophysical models. The standard deviations  $\sigma^i$  are calibrated on results from full parameter estimation pipelines by Ghosh et al. (2016). Among their results, we selected runs making use of the IMRPhenomB template model with aligned spins and a 3-detector network. As the number of sources in this data set is relatively sparse ( $< 10^3$  sources), we implemented the same closest-neighbour fit presented by Gerosa & Berti (2017) to estimate the values of  $\sigma^i$  for any values  $M_t$ ,  $q$  and  $z$ . Errors on the spins parameters are not provided by Ghosh et al. (2016); for simplicity, we assume aLIGO observation will achieve a precision  $\sigma_{\chi_{eff}} = 0.1$  (Abbott et al. 2018a) regardless of the spin models and other source parameters.<sup>1</sup>

Measurement errors for ET are obtained by rescaling the aLIGO results using the SNR

$$\sigma_{ET} = \sigma_{aLIGO} \frac{\rho_{aLIGO}}{\rho_{ET}}, \quad (11)$$

as expected in the large-SNR limit (Poisson & Will 1995).

## 3.2. Bayesian modeling

### 3.2.1. Model rates

The general expression for the rate of a given model with parameters  $\theta$  can be written as

$$\frac{dN}{d\lambda}(\theta) = N(\theta)p(\lambda|\theta), \quad (12)$$

where  $N(\theta)$  is the total number of sources predicted by the model and  $p(\lambda|\theta)$  is the normalised model distribution or rate.

From the catalog of sources presented in Section 2.4, we approximate the normalised rates  $p$  using kernel density estimation (KDE) methods. Gaussian kernels with a bandwidth parameter of 0.05 on the data set  $\{M_t, q, \chi_{eff}\}$  are capable of accurately reproducing the distributions in Figure 1 for both formation channels and spin models.

<sup>1</sup> At the current sensitivity, prior choices also play an important role (Vitale et al. 2017a).

### 3.2.2. Hierarchical inference

Statistical inference is implemented with a standard Bayesian hierarchical model. Our analysis is based on the formalism already presented by [Loredo \(2004\)](#), [Mandel et al. \(2019\)](#) and [Taylor & Gerosa \(2018\)](#) and detailed in Appendix A. In a nutshell, the posterior distribution on the model parameters  $\theta$  marginalized over  $N(\theta)$  is

$$p(\theta|d) \propto p(\theta) \prod_{k=1}^{N_{\text{det}}} \frac{\int p(\lambda|\theta) p(d|\lambda)/p(\lambda) d\lambda}{\int p_{\text{det}}(\lambda) p(\lambda|\theta) d\lambda}, \quad (13)$$

where  $N_{\text{det}}$  is the number of entries in the detection catalog,  $p(\lambda|\theta)$  describes the astrophysical model,  $p(\theta)$  is the prior on each astrophysical model,  $p(d|\lambda)$  is the posterior of an individual GW event,  $p(\lambda)$  is the prior used in the single-event analysis, and  $p_{\text{det}}(\lambda)$  describes selection effects.

If the posterior  $p(\lambda|\theta)$  is provided in terms of Monte-Carlo samples  $\lambda_i$ , as in [Abbott et al. \(2018a\)](#), we can rewrite Eq. (13) as

$$p(\theta|d) \propto p(\theta) \prod_{k=1}^{N_{\text{det}}} \frac{\sum_i p(\lambda_i|\theta)/p(\lambda_i)}{\int p_{\text{det}}(\lambda) p(\lambda|\theta) d\lambda}. \quad (14)$$

For the case of our mock Gaussian posteriors, Eq. (13) becomes

$$p(\theta|d) \propto p(\theta) \prod_{k=1}^{N_{\text{det}}} \frac{\int p(\lambda|\theta) \mathcal{N}(\lambda_k, \sigma_k)/p(\lambda) d\lambda}{\int p_{\text{det}}(\lambda) p(\lambda|\theta) d\lambda}. \quad (15)$$

The denominator  $\beta(\theta) \equiv \int p_{\text{det}}(\lambda) p(\lambda|\theta) d\lambda$  does not depend on the event parameters  $\lambda$ , but only on the model  $\theta$ . In practice, we estimate  $\beta(\theta)$  by generating values for  $(M_t, q)$  using rejection sampling from our KDE approximation of the astrophysical rates, and extracting values for the aligned components of the spins and the redshift as described in sections 2.3.1 and 2.3.2. As the dynamical catalog is formed of both exchanged and originals BBHs, we consider the two sub-populations separately and combine them in the proportion predicted by our dynamical simulations.

## 4. RESULTS

### 4.1. Model selection: pure dynamical or pure isolated channel

We first apply the formalism presented in Section 3.2.2 to the case where the astrophysical model is such that all BBHs are assumed to form only via the isolated or dynamical channels. In this case, we can estimate what model best fits a given set of data by computing the odds ratio,

$$\mathcal{O}_{AB} = \frac{p(A|d)}{p(B|d)}, \quad (16)$$

where  $A$  and  $B$  either stand for isolated or dynamical and  $p$  is the model posterior distribution derived at the end of Section 3.2.2. Values  $\mathcal{O}_{AB} \gg 1$  indicate that model A is strongly favoured by the data, while model B is preferred for  $\mathcal{O}_{AB} \ll 1$ . It is somewhat indicative to relate values of the odds ratio to  $\sigma$ -levels of Gaussian measurements:

$$\mathcal{O} = \frac{1}{1 - \text{erf}(\sigma/\sqrt{2})}, \quad (17)$$

where erf is the error function.

#### 4.1.1. Mock observations

We want to assess the number of observations  $N_{\text{obs}}$  needed to discriminate between the two models assuming that one of them is a faithful representation of reality. To understand how each parameter impacts the analysis, we run our statistical pipeline assuming we only measure either  $\{M_t\}$ ,  $\{M_t, q\}$  or  $\{M_t, q, \chi_{\text{eff}}\}$ . In practice, this implies that the integral in Eq. (15) is evaluated on the selected variables while marginalising over the others. Regarding the production of mock data, we generated  $10^3$  sets of observations for each value of  $N_{\text{obs}}$  in order to produce a statistical estimation on our results. Each of the mock observations was sampled from the model and included in the observation set with probability  $p_{\text{det}}$ .

In Figure 2, we show values of the odds ratio  $\mathcal{O}_{AB}$  where the set of observed events is generated from model A, so that  $\mathcal{O}_{AB}$  is expected to increase as  $N_{\text{obs}}$  grows. We present results for observations generated from the dynamical (orange) and isolated (blue) models both for aLIGO/aVirgo (left panels) and ET (right panels).

Let us focus first on the top row, where only the total mass is considered in the analysis. In the case of aLIGO/aVirgo, the upper limit of the 90% credible interval presents high values  $\mathcal{O} > 10^{10}$  even for small  $N_{\text{obs}}$ . This is because the isolated model only predicts merging BBHs with  $M_t < 70M_{\odot}$ . As a result, any observation where the entire support of the total mass posterior distribution is above  $70M_{\odot}$  can only be described by the dynamical model.

Another interesting feature is that the lower bound of the credible interval has values of  $\mathcal{O} > 10^{10}$  starting at  $N_{\text{obs}} \sim 40 - 50$ , which is consistent with our models. In fact, as the number of dynamical BBHs with  $M_t > 70M_{\odot}$  in our dynamical catalog is equal to 16 (out of 229), the probability of not observing any massive BBHs in 50 observations generated from the dynamical model is equal to  $p = 1.9 \times 10^{-3}$ . It is also interesting to see that the results obtained are similar when observing with ET, suggesting that the differences in total mass range of our models is the dominant effect in the analysis.

In contrast, the evolution of the odds ratio for the isolated case presents a steady increase of  $\mathcal{O}$  with  $N_{\text{obs}}$ , as expected. This indicates that, unlike the dynamical case, there is no strong feature in the total mass spectrum of the isolated model that can drive the odds ratio towards very high values with a few observations. In this case, measurements with ET improve the analysis: the median number of observations yielding  $\mathcal{O} = 10^6$  ( $5\sigma$  level) decreases from 40 with aLIGO down to  $\sim 30 - 35$ .

The second row of Figure 2 presents results obtained when performing the analysis considering both the total mass and the mass ratio. We observe that including the mass ratio helps discriminating faster between the two models for all cases. When observations are generated from the dynamical model and observed with ET, the lower bound of the 90% credible interval reaches the  $5\sigma$  level for  $N_{\text{obs}} = 20$  while we had a value of  $N_{\text{obs}} = 40$  for the analysis considering only the total mass. Similarly, in 90% of the cases we find that only 30 observations generated from the isolated model and observed by ET gives values of odds ratio corresponding to a  $5\sigma$  level (60 observations were needed when considering only  $M_t$ ).

Finally, in the last two rows we present results obtained by simulating measurements with the total mass, the mass ratio and the effective spin parameter, adopting either the low-spin (LS, third row) or high-spin (HS, fourth row) models. When the dynamical model is true, only a few observations are needed to push the value of the lower bound of the credible interval towards  $\mathcal{O} > 10^{10}$  (similar results were found in [Stevenson et al. 2017a](#)). This results from our assumption for the spin models, as only the dynamical model has support for  $\chi_{\text{eff}} < 0$ , meaning that observations of negative values give a full support to the dynamical model. As the errors on the spins are already relatively low, with  $\sigma_{\chi_{\text{eff}}} = 0.1$  for aLIGO/aVirgo (see Section 3.1.2), we do not see much difference when repeating the analysis with even smaller error for ET. Finally, when the isolated model is true, the  $5\sigma$  level is attained in 90% of the cases only after only 15 observations for the LS model and 20 for the HS model. Measurements with ET are slightly better and bring the number of necessary observations down to  $\sim 10 - 15$  for both models.

#### 4.1.2. LVC observations

We now apply the formalism described in the previous Section to aLIGO/aVirgo BBH data from the catalog in [Abbott et al. \(2018a\)](#). We have used both the posterior and prior samples for each event released by the LVC, to compute the expression of the posterior in Eq. (14). In Table 2, we present the values obtained for  $\mathcal{O}$  and

$\sigma$  for the favoured model under the same parameter configurations presented in the last section.

For all simulations, the values of the odds ratio are  $\mathcal{O} \lesssim 10^4$ , corresponding to  $\lesssim 4\sigma$  and favour a dynamical origin. If we only take into account the mass parameters, the analysis done with  $\{M_t, q\}$  gives a value of  $\mathcal{O} = 1.7 \times 10^3$  (corresponding to  $3.4\sigma$ ) in favour of the dynamical model. Our isolated model has a hard upper-mass cutoff, which cannot accommodate the presence of GW170729. If we discard this event from the observed catalog, the odds ratio drops to  $\mathcal{O} = 1.6 \times 10^1$  when considering  $\{M_t, q\}$ .

Our results depend strongly on the underlying spin model. In the LS case, we get  $\mathcal{O} = 39$ , with slight preference for the dynamical model. This conclusion changes drastically in the HS case, as the value of the odds ratio is now larger than  $10^{16}$ . This is because in the HS case the dynamical model has a much wider range of possible negative values for  $\chi_{\text{eff}}$ , extending down to  $\chi_{\text{eff}} - 0.5$  (compared to  $\chi_{\text{eff}} - 0.2$  for LS case). As some of the events (namely GW1701014, GW170818, GW170823) have posterior distribution with some support for values of  $\chi_{\text{eff}} < -0.25$ , the HS dynamical model is significantly favored compared to the HS isolated case. When removing these events from the analysis, we found that the odds ratio is reduced by 10 orders of magnitude down to  $2.7 \times 10^6$ .

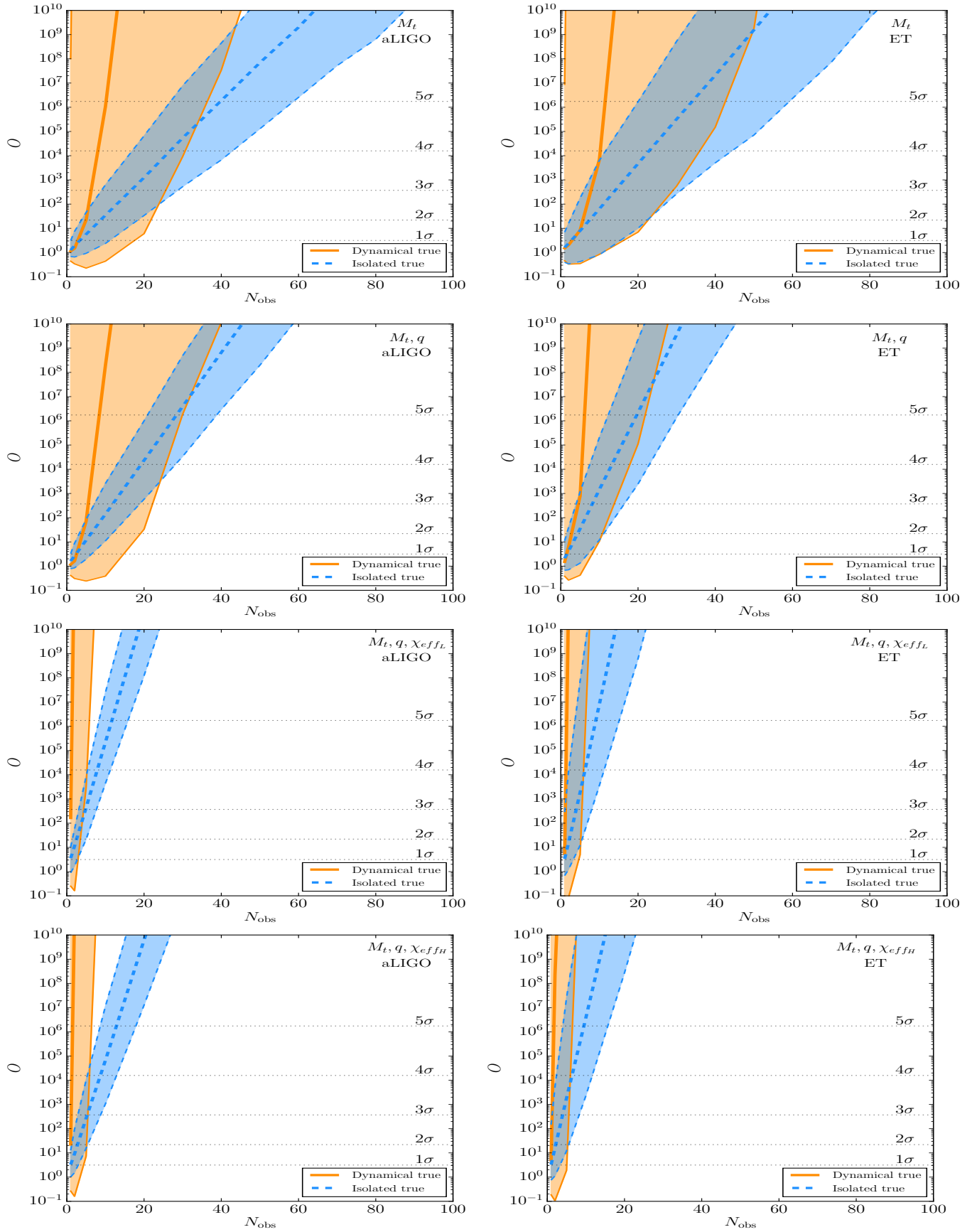
Parameters	Favoured model	$\mathcal{O}$	$\sigma$ -equivalent
$M_t$	dynamical	42	2.3
$M_t, q$	dynamical	$1.7 \times 10^3$	3.4
$M_t, q, \chi_{\text{eff}_L}$	dynamical	39	2.2
$M_t, q, \chi_{\text{eff}_H}$	dynamical	$4.1 \times 10^{16}$	x
$M_t, q$ [1]	dynamical	16	1.9
$M_t, q, \chi_{\text{eff}_H}$ [2]	dynamical	$2.7 \times 10^6$	5.1

**Table 2.** Summary of the results obtained when running Bayesian model selection on the set of 10 BBHs observed by aLIGO/aVirgo ([Abbott et al. 2018a](#)). We run our statistical analysis varying the set of parameters considered. The configurations [1] and [2] correspond to cases where we respectively exclude GW170729 and the set of events (GW1701014, GW170818, GW170823) from the analysis. Among these models, we prefer a dynamical origin mainly because of the hard cuts in the mass and spin distributions (cf. section 5).

#### 4.2. Mixture model

In the previous section, we have shown results obtained when comparing the two models assuming that one of them is true. While this pointed out interesting features





**Figure 2.** Evolution of the odds ratio as a function of the number of simulated observations generated from the dynamical (orange) and isolated model (blue). The odds ratios are defined such that the posterior of the true model is at the denominator. Equivalent  $\sigma$  levels are reported with dotted lines. The four rows correspond to analysis done with  $M_t$ ,  $q$ ,  $\chi_{\text{effL}}$  and  $\chi_{\text{effH}}$  (from top to bottom); while the left and right panels are associated to results with aLIGO/aVirgo and ET. For each scenario, the thick lines are the median values while the surfaces represent the 90% credible regions.

of our models, it is probably unrealistic to assume that *all* merging BBHs only come either from the dynamical or the isolated channel. In this section, we assume that the population of merging BBHs comes from a mixture of the two models parametrised by a “mixing fraction”  $f$  such that

$$r_{\text{MM}}(f) = f r_{\text{iso}} + (1 - f) r_{\text{dyn}}, \quad (18)$$

where  $r_{\text{MM}}(f)$  are the rates of the mixture model and  $r_{\text{iso}}$ ,  $r_{\text{dyn}}$  are the rates of the isolated and dynamical models, respectively. Similarly, the detection efficiency of the mixed model is also given by

$$\beta_{\text{MM}}(f) = f \beta_{\text{iso}} + (1 - f) \beta_{\text{dyn}} \quad (19)$$

where  $\beta_{\text{iso}}$ ,  $\beta_{\text{dyn}}$  are the detection efficiencies (cf. Sec. 3.2.2) for the isolated and dynamical models, respectively.

We want to estimate the posterior distribution of the mixing fraction  $f$ . Once again, we have considered two different cases with either fiducial or aLIGO/aVirgo data. In both cases, we have used a Metropolis-Hastings Monte Carlo algorithm to estimate the posterior distribution. We find that a Gaussian jump proposal with standard deviation  $\sigma = 0.5$  is sufficient to have good results and convergence when running  $10^6$  iterations chains.

#### 4.2.1. Mock observations

We generate mock observations assuming a “true” mixing fraction  $f_T = 0.7$ . To generate  $N_{\text{obs}}$  events from a mixed model, we first generated  $N_{\text{obs}}$  events both for isolated and dynamical models. For each entry of the mixed model, we draw a random number  $\epsilon \in \mathcal{U}[0, 1]$  and associate an event from the pre-processed isolated (dynamical) set if  $\epsilon < f_T$  ( $\epsilon > f_T$ ).

In Figure 3, we report the values of the medians (square), 90% (straight line) and 99% (dashed line) credible intervals for a set of mock observations with  $N_{\text{obs}} = \{10, 100, 500\}$ . The values correspond to the current number of events ( $N_{\text{obs}} = 10$ ), an optimistic prediction for O3 ( $N_{\text{obs}} = 100$ ) and a high-statistic case ( $N_{\text{obs}} = 500$ ). For simplicity, we restrict this study to the aLIGO/aVirgo detector case. For each set of observations we perform the analysis with the combination of parameters  $\{M_t\}$  (orange),  $\{M_t, q\}$  (purple),  $\{M_t, q, \chi_{\text{eff}_L}\}$  (green) and  $\{M_t, q, \chi_{\text{eff}_H}\}$  (red).

For any set of parameters, we observe a reduction of the width of the credible intervals for higher values of  $N_{\text{obs}}$ , as expected. In fact, for  $N_{\text{obs}} = 10$  the 99% credible interval spans almost the entire range of values for  $f$ , while for  $N_{\text{obs}} = 500$  it is reduced down to 0.2 for  $M_t$  and only  $\sim 0.15$  when including  $q$  and  $\chi_{\text{eff}}$  in the analysis. Another feature shown in Figure 3 is that the

width of the credible interval gets smaller when including more parameters, which is expected as more parameters provide more constraints on the model selection analysis.

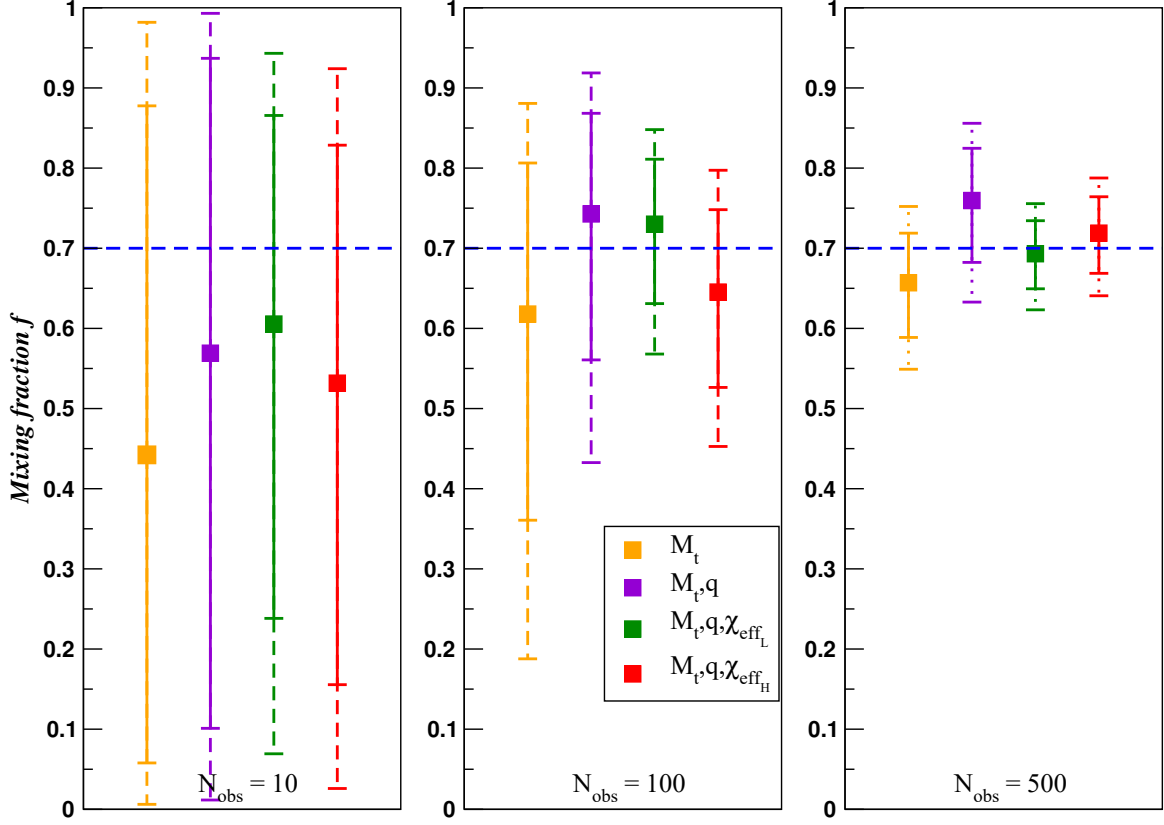
In conclusion, Figure 3 suggests that already with 100 detections (optimistic scenario for the end of O3), we can constrain the value of the fractional errors on the mixing fraction to an interval smaller than 0.5 using  $\{M_t, q\}$  and is in agreement with previous studies (Vitale et al. 2017b; Stevenson et al. 2017a; Zevin et al. 2017; Talbot & Thrane 2017). Furthermore, with even higher statistics of a few hundred detections, the fractional errors on the mixing fraction will go down to 20%, if we consider only the mass parameters. The inclusion of the effective spin parameter reduces this value even further down to 10%. Finally, we have repeated this study for a value of  $f_T = 0.3$ , and found similar predictions.

#### 4.2.2. LVC observations

In this section, we describe the results obtained when applying our mixed model analysis to aLIGO/aVirgo data (Abbott et al. 2018a). Figure 4 shows the posterior distributions obtained when doing the analysis using  $\{M_t\}$  (orange),  $\{M_t, q\}$  (purple),  $\{M_t, q, \chi_{\text{eff}_L}\}$  (green) and  $\{M_t, q, \chi_{\text{eff}_H}\}$  (red).

First, the two posterior distributions obtained when using the mass parameters are slightly shifted towards values of  $f < 0.5$  (pure dynamical scenario) with a median value of 0.40 and 0.36 for  $\{M_t\}$  and  $\{M_t, q\}$  respectively. In addition, the upper limit of the 99% credible interval is equal to 0.94 ( $\{M_t\}$  case) and 0.90 ( $\{M_t, q\}$  case), hence excluding our pure isolated scenario. As before, this is due to the fact that some of the detected BBHs have support for  $M_t > 70M_\odot$  while our isolated model strictly prevents the occurrence of these high masses. In particular, for GW170729, the lower 90% credible interval of  $M_t$  is equal to  $74.2M_\odot$ . For testing purposes, we ran an analysis where the total mass of GW170729 was not included (blue curve); in this case, the median was 0.55 and the upper limit of the 99% credible interval is much higher, with values of 0.99, hence giving more support towards the isolated scenario.

Including the effective spin parameter in the analysis has a significant impact on the posterior distributions. In the LS model, the distribution of the mixing fraction is centered towards higher values of  $f$  (with a median of 0.58), favouring the isolated scenario. Once again, this result is dominated by a single event, GW151226, that presents an effective spin parameter with a clearly positive median value  $\chi_{\text{eff}} = 0.18$  that is not well supported by the dynamical scenario. We re-ran the analysis for the LS model excluding this event (cyan curve) and found that the value of the median is reduced down to



**Figure 3.** Median (squares), 90% (thick lines) and 99% (dashed lines) credible intervals for the mixing fraction posterior distribution as a function of the set of parameters used for the analysis.  $f = 0$  ( $f = 1$ ) corresponds to the pure dynamical (isolated) model. In all cases, the fiducial data set were generated from a mixed model with mixing fraction  $f_T = 0.7$  and with number of observations  $N_{\text{obs}} = 10, 100, 500$  (from left to right).

a value of 0.48. It is interesting to highlight that while GW170729 has even higher values of effective spin (median of  $\chi_{\text{eff}} = 0.36$ ), this event gives only limited support to the isolated scenario as the masses are very high (unlike GW151226). In the HS case, the dynamical scenario is favoured with a median value for the mixing fraction equal to 0.27. This can be understood by the fact that the support for the effective spin parameter in the dynamical case extends between  $-0.6$  and  $0.7$  (see Figure 1), which is the range of all the events currently observed by aLIGO and aVirgo, while the isolated scenario now struggles to capture events with negative values of  $\chi_{\text{eff}}$ .

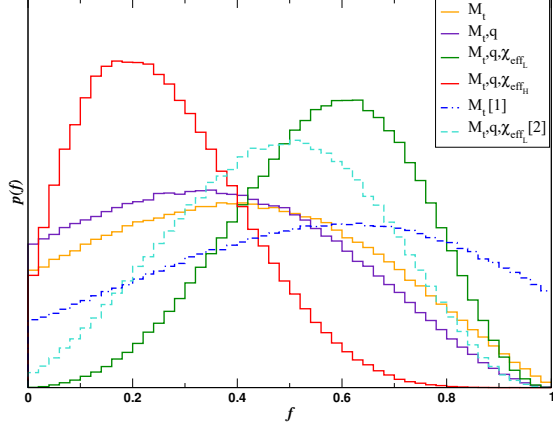
In conclusion, our results suggest that O1+O2 LVC data (Abbott et al. 2018a) exclude a pure isolated scenario as described by our population-synthesis simulations, which however include hard cutoffs on both mass and spin distributions. It is important to point out that the width of the posterior distribution is still quite large, and it is in agreement with the results obtained in the last section. Moreover, we stress that the models considered in this paper refer to a single metallicity: at lower progenitor metallicity even the isolated scenario includes higher mass BHs (with  $M_t$  up to  $80 - 90 M_{\odot}$ , Giacobbo

& Mapelli 2018). Thus, in a follow-up study (Bouffanais et al., in preparation), we will explore the importance of metallicity and other important population-synthesis parameters.

## 5. DISCUSSION

In this first paper, we have applied a Bayesian model-selection framework to discriminate BBHs formed in isolation versus those those formed in young star clusters. Under the assumption that there is only one “true” formation channel, O1+O2 LVC data prefer the dynamical formation channel at  $3.4\sigma$ , if we include in our analysis only the total mass  $M_t$  and mass ratio  $q$ . Similarly, if we adopt a more realistic mixture model, O1+O2 LVC data (Abbott et al. 2018a) exclude a purely isolated scenario as described by our population-synthesis simulations. We stress that this is a very model-dependent statement: our isolated binary formation model contains hard cutoffs in both binary masses ( $M_t < 70 M_{\odot}$ ) and spins ( $\chi_{\text{eff}} > 0$ ), which cannot accommodate some of the events, notably GW170729 and GW151226.

The effect of the mass cutoff is particularly important. We compared aLIGO data against two models, one with



**Figure 4.** Posterior distribution inferred from a  $10^6$  MCMC chain ran on the LVC observations. The legends [1] and [2] correspond to cases where the events GW170729 and GW151226 were respectively excluded from the analysis.  $f = 0$  ( $f = 1$ ) corresponds to the pure dynamical (isolated) model. The set of parameters considered for the analysis are indicated in the legend.

a mass cutoff at  $m_1, m_2 \sim 35M_\odot$  (isolated) and one without it (dynamical). Our analysis suggests preference for the model without cutoff. On the other hand, Abbott et al. (2018b) find strong evidence for an upper mass gap starting at  $\sim 45M_\odot$ —this constraint being driven by the lower limit on the largest BH mass in the sample. Crucially, they do not compare data against astrophysical simulations with a well specified set of assumptions, but rather fit a generic phenomenological population model. Going forward, this difference highlights the importance of accurately modeling tails of the predicted astrophysical distributions, that despite being responsible for a small number of events, might play a qualitative role in discriminating among different formation channels.

Several caveats need to be discussed about our models. Crucially, the simulations considered here assume a single metallicity  $Z = 0.1 Z_\odot$ . Isolated binaries with lower metallicity ( $Z \lesssim 0.01 Z_\odot$ ) simulated with MOBSE end up producing merging BBHs with total mass up to  $\sim 80 - 90 M_\odot$  (Mapelli et al. 2017; Giacobbo & Mapelli 2018). Thus, the effect of metallicity is somewhat degenerate with the effect of dynamics. We expect a pure isolated scenario to be consistent with O1 and O2 data if we include more metal-poor progenitors (down to  $Z \sim 0.0002 = 0.01 Z_\odot$ ).

Moreover, the simulations considered in this paper investigate only one model for core-collapse supernovae (from Fryer et al. 2012) and only one model for pair-instability and pulsational pair-instability supernovae. Furthermore, we assume a specific model for BH natal kicks, which are highly uncertain and can crucially impact the properties of merging systems (e.g. Belczynski et al.

2016b; Mapelli et al. 2017; Barrett et al. 2018; Wysocki et al. 2018; Gerosa et al. 2018). Even the prescriptions for common envelope affect the properties of merging systems significantly (here we consider only one value of  $\alpha = 3$ ).

It is worth mentioning here that MOBSE predicts a significant difference between the maximum mass of BHs and the maximum mass of *merging* BHs in isolated binaries (see e.g. Figure 11 of Giacobbo & Mapelli 2018). From progenitors with metallicity  $Z = 0.1 Z_\odot$  ( $Z = 0.01 Z_\odot$ ) we form BHs with individual mass up to  $\sim 55 M_\odot$  ( $\sim 65 M_\odot$ ), but only BHs with individual mass up to  $\sim 35 M_\odot$  ( $\sim 45 M_\odot$ ) merge within a Hubble time through isolated binary evolution. This behaviour is similar to what found from the independent population-synthesis code SEVN (Spera et al. 2019) and springs from the interplay between stellar radii and common envelope (see Giacobbo & Mapelli 2018 and Spera et al. 2019 for more details). Thus, if the progenitor metallicity is  $Z = 0.1 Z_\odot$ , merging BBHs from isolated binaries have  $M_t \leq 70 M_\odot$ , while non-merging BBHs from isolated binaries have  $M_t \leq 110 M_\odot$ .

In contrast, dynamically formed BBHs (especially BBHs formed by dynamical exchange) might be able to merge even if their total mass is larger than  $\sim 70 M_\odot$ , because common envelope is not the only way to shrink their orbits: dynamical exchanges and three-body encounters also contribute to reducing the binary semi-major axis and/or increasing the orbital eccentricity. This is the main reason why  $M_t$  is substantially larger for dynamical BBH mergers than for isolated BBH mergers.

In addition, dynamical BBHs can host BHs which form from the collision between two or more stellar progenitors (which is not possible in the case of isolated binaries). In our models, if a BH forms from the evolution of a collision product between an evolved star (with a well-developed Helium or Carbon-Oxygen core) and another star, it might have a significantly larger mass, because a collision product can end its life with a larger total (or core) mass and with a larger envelope-to-core mass ratio than a single star (see Di Carlo et al. 2019 for additional details). This further enhances the mass difference between isolated and dynamical merging BBHs. Moreover, BHs that form from the collisions of two or more stars are also allowed (in some rare cases) to have mass in the pair-instability mass gap, if their final Helium core is below the threshold for (pulsational) pair instability but their Hydrogen mass is larger than expected from single stellar evolution (in our models we impose that the BH mass is equal to the final progenitor mass, including



Hydrogen envelope, if there is a direct collapse<sup>2</sup>). From Di Carlo et al. (2019) we find that  $\lesssim 2\%$  of all merging BHs have mass in the pair-instability mass gap.

We also stress that the very edges of the pair-instability mass gap are not uniquely constrained. The lower boundary of the mass gap can be as low as  $\sim 40 M_\odot$  or as large as  $\sim 65 M_\odot$ , depending on details of the stellar evolution model (e.g. with/without rotation, Mapelli et al. in preparation) and of the pair-instability SN model.

Finally, in the current paper we assume a simplified evolution of the merger rate with redshift (uniform in comoving volume and source-frame time). The evolution with redshift is important (especially for ET) not only for the merger rate but also for the properties of merging systems, because of the influence of the cosmic star formation rate and of the metallicity evolution on BBHs (e.g. Mapelli et al. 2017; Lamberts et al. 2016). All these ingredients must be kept in mind when interpreting the results of our study. Thus, in a follow-up study we will include different metallicities, a self-consistent redshift evolution model, and we will consider a larger parameter space.

Previous papers already addressed the issue of dynamical formation versus isolated formation of BBHs within a model-selection approach (e.g. Stevenson et al. 2015, 2017a; Vitale et al. 2017b; Zevin et al. 2017). Most previous studies adopt just simple prescriptions for the dynamical evolution and do not consider a full set of dynamical simulations. Zevin et al. (2017) did compare results of a population-synthesis sample and a set of dynamical simulations. However, their approach is significantly different from ours as we make use of direct N-body simulations of young star clusters, while they consider hybrid Monte Carlo simulations of globular clusters. Globular clusters are massive ( $> 10^4 M_\odot$ ) old star clusters (most of them formed around 12 Gyr ago). They are site of intense dynamical processes: binary hardening and exchanges in globular clusters are very effective (see e.g. Wang et al. 2015; Rodriguez et al. 2015; Askar et al. 2017), but nowadays the stellar mass still locked up in globular clusters is a small fraction of the total stellar mass ( $< 1\%$ , Harris et al. 2013). In contrast, young star clusters are smaller systems (the systems we consider here have mass  $\sim 10^{3-4} M_\odot$ ) and are mostly short-lived ( $< 1$  Gyr), but they form continuously through the cosmic history and are expected to host the bulk of massive star formation (Lada & Lada 2003; Weidner & Kroupa 2006; Weidner et al. 2010). Thus, the importance of dynamics in a single young star cluster is lower than in a

single globular cluster, but the cumulative contribution of young star clusters to the dynamical formation of BBHs is a key factor (e.g. Di Carlo et al. 2019; Kumamoto et al. 2019).

From a more technical point of view, globular clusters are spherically symmetric relaxed systems. Hence, they can be simulated with a fast Monte Carlo approach (Hénon 1971; Joshi et al. 2000). In contrast, young star clusters are asymmetric and irregular systems, still on their way to relaxation (Portegies Zwart et al. 2010). Hence, we need more computationally expensive direct N-body simulations to model them realistically. Thus, our approach and the one followed by Zevin et al. (2017) are complementary both scientifically and numerically. To understand the dynamical formation of BBHs, we need to model both the globular cluster and the young star cluster environment. The final goal is to have a model-selection tool to distinguish isolated binary formation from dynamical formation, able to account for the many different flavours of dynamical formation (globular clusters, young star clusters, galactic nuclei and hierarchical triples). While we are still far from this goal, our work provides a new crucial piece of information in this direction.

## 6. SUMMARY

The formation channels of BBHs are still an open question. Here, we use a Bayesian model-selection framework and apply it to the isolated binary scenario versus the dynamical scenario of BBH formation in young star clusters. Young star cluster dynamics might be extremely important for BBHs, because the vast majority of massive stars (which are progenitors of BHs) form in young star clusters and OB associations (see e.g. Lada & Lada 2003; Portegies Zwart et al. 2010). However, only few studies focus on BBH formation in young star clusters, because this is a computational challenge (Ziosi et al. 2014; Mapelli 2016; Banerjee 2017, 2018; Kumamoto et al. 2019; Di Carlo et al. 2019). Here, we consider the largest sample of merging BBHs produced in a set of N-body simulations of young star clusters (Di Carlo et al. 2019). For the isolated binaries, we take a sample of  $> 3 \times 10^4$  merging BBHs simulated with the population synthesis code MOBSE (Giacobbo et al. 2018; Giacobbo & Mapelli 2018). MOBSE includes state-of-the-art models for stellar winds and BH formation. The same population-synthesis algorithm is used also in the N-body simulations, ensuring a fair comparison of the two scenarios (see Section 2).

We analyzed the two scenarios with a Bayesian hierarchical modeling approach capable of estimating which models best fit a given set of GW observations (see

<sup>2</sup> This assumption is still matter of debate given the high uncertainties on direct collapse (see e.g. Sukhbold et al. 2016).

Section 3). We looked at two different cases where we assumed that the underlying astrophysics is either described by a single model, or by a combination of the two models weighted by a mixing fraction parameter  $f$ . In both analyses, the statistical framework was applied on the combination of the mass parameters and the effective spin. The redshift dependence is inevitably entangled with the metallicity of the sources, which will be explored extensively elsewhere.

In terms of GW observations, we used both mock data and LVC observations during O1 and O2 (Abbott et al. 2018a). Our results with mock observations showed that the distributions of  $M_t$  and  $q$  already present strong features that can be used to differentiate between the two models. In fact, with 500 observations with aLIGO and aVirgo we could be able to restrict the values of the mixing fraction to an interval smaller than 0.5. With the inclusion of the effective spin parameter in the analysis, we found that this interval becomes even smaller with values close to 0.15.

Finally, this work is the first one that used the latest LVC data to perform Bayesian model selection approach in order to discriminate between BBHs formed via isolated or dynamical binaries. Our results showed that the current set of observations is not able to put a strong constraint on the mixing fraction of the two models. A pure isolated (dynamical) scenario in which all BBH progenitors have metallicity  $Z = 0.002$ , as described by our simulations, is barely consistent (still consistent)

with LVC data, because of the presence of massive BBHs such as GW170729. We stress that progenitor metallicity and dynamics have a somewhat degenerate effect on the maximum mass of merging BBHs: we expect the pure isolated scenario will be still consistent with O1+O2 data if we include more metal-poor progenitors (down to  $Z \sim 0.0002$ ). Thus, in a follow-up study, we will apply our methodology to a range of metallicities for both the dynamical and isolated scenarios.

Finally, given our estimations obtained with mock observations, we expect that after about a hundred detections (optimistic scenario for O3, middle panel of Figure 3), we should already be able to constrain the values of the mixing fraction in an interval smaller than 0.5.

## ACKNOWLEDGMENTS

MM and YB acknowledges financial support by the European Research Council for the ERC Consolidator grant DEMOBLACK, under contract no. 770017. YB would like to acknowledge networking support by the COST Action GWverse CA16104. EB and VB are supported by NSF Grant No. PHY-1841464, NSF Grant No. AST-1841358, NSF-XSEDE Grant No. PHY-090003, and NASA ATP Grant No. 17-ATP17-0225. This work has received funding from the European Unions Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 690904. Computational work was performed on the University of Birmingham’s BlueBEAR cluster and at the Maryland Advanced Research Computing Center (MARCC).

## APPENDIX

### A. HIERARCHICAL INFERENCE

Bayesian hierarchical modeling is a powerful approach that enables to generate predictions on the parameters of a given model. In this section, we present how we applied this technique in our context by following the formalism presented in Taylor & Gerosa (2018), Mandel et al. (2019).

Let us consider that data  $d$  of a GW detector contain an ensemble of GW signals indexed by  $k$ ,  $d = \{h_k\}_{k=1}^N$ , where  $N$  is the total number of signals. The posterior distribution of the model parameters  $\theta$  is given by Bayes’ theorem as,

$$p(\theta|\{h_k\}) = \frac{p(\{h_k\}|\theta)p(\theta)}{p(\{h_k\})}, \quad (\text{A1})$$

where  $p(\{h_k\}|\theta)$  is the likelihood of the model,  $p(\theta)$  is the prior on the model parameters and  $p(\{h_k\})$  is the evidence on the data. The likelihood of the model can be expressed in terms of the joint probability distribution over the signal parameters,

$$p(\{h_k\}|\theta) = \int_{\{\lambda_k\}} p(\{h_k\}, \{\lambda_k\}|\theta) d\{\lambda_k\}, \quad (\text{A2})$$

where the integration needs to be done over the parameters of all events  $\{\lambda_k\}$ . The previous joint distribution can be split into two contributions,

$$p(\{h_k\}, \{\lambda_k\}|\theta) = p(\{h_k\}|\{\lambda_k\}) \times p(\{\lambda_k\}|\theta), \quad (\text{A3})$$

where the left term is the GW events likelihood and the right term is the parametric likelihood of the model. The GW events likelihood can be computed via Bayes' theorem as

$$p(\{h_k\}|\{\lambda_k\}) = \frac{p(\{\lambda_k\}|\{h_k\})p(\{h_k\})}{p(\{\lambda_k\})}, \quad (\text{A4})$$

where  $p(\{\lambda_k\}|\{h_k\})$  is the posterior distribution for the combined set of events,  $p(\{\lambda_k\})$  is the prior distribution on the set of observed parameters, and  $p(\{h_k\})$  is the evidence on the data.

Now, to derive an expression for the parametric model likelihood in Eq. (A3), let us first assume that all quantities are binned along the parameters  $\lambda$ . In this case, the set of events parameters  $\{\lambda_k\}$  become a set of numbers of events contained in each bin,  $\{n_l\}$ , where  $l$  goes from 1 to the total number of bins  $N_{\text{bin}}$ . The probability of having a given distribution of events  $\{n_l\}$  is given by a Poisson process,

$$p(\{n_l\}|\theta, N(\theta)) = \prod_{l=1}^{N_{\text{bin}}} \frac{r_l(\theta)^{n_l} \times e^{-r_l(\theta)}}{n_l!}, \quad (\text{A5})$$

where  $N(\theta)$  is the total number of events predicted by the model and  $r_l(\theta)$  is given by

$$r_l(\theta) = N(\theta) \times p_l(\theta), \quad (\text{A6})$$

where  $p_l(\theta)$  is the binned version of the rate  $p(\lambda|\theta)$ .

Now if we take the limit where the bins become infinitely small, we arrive to a point where each bin will either have  $n_l = 1$  or  $n_l = 0$  events, which gives the following continuous limit,

$$p(\{\lambda_k\}, N(\theta)|\theta) = \left( \prod_{k=1}^N p(\lambda_k|\theta) \right) \times N(\theta)^N e^{-N(\theta)}. \quad (\text{A7})$$

As we are not interested in predicting the total number of events for a model, we marginalise over  $N(\theta)$  as,

$$p(\{\lambda_k\}|\theta) = \sum_{N=1}^{\infty} p(\{\lambda_k\}, N|\theta), \quad (\text{A8})$$

$$= \left( \prod_{k=1}^{N_{\text{det}}} p(\lambda_k|\theta) \right) \times K(N), \quad (\text{A9})$$

where the quantity  $K(N)$  is model-independent, and only a function of the total number of events.

Putting everything together in Eq A2, we find the following expression for the posterior distribution of the model

$$p(\theta|\{h_k\}) = K(N)p(\theta) \prod_{k=1}^N \int_{\lambda_k} \frac{p(\lambda_k|h_k)p(\lambda_k|\theta)}{p(\lambda_k)}. \quad (\text{A10})$$

There is a final caveat that need to be dealt with in this analysis. In fact, while  $N$  events are contained in the data, some of these events will be classified as detectable and the rest as non detectable. As in reality, we only have access to the  $N_{\text{det}}$  observed events, this induces a selection effect. To take this effect into account, one can show that the previous posterior distribution needs to be updated as,

$$p(\theta|\{h_k\}) = K(N_{\text{det}})p(\theta) \prod_{k_{\text{det}}=1}^{N_{\text{det}}} \int_{\lambda_{k_{\text{det}}}} \frac{p(\lambda_{k_{\text{det}}}|h_{k_{\text{det}}})p(\lambda_{k_{\text{det}}}|\theta)}{p(\lambda_{k_{\text{det}}})\beta(\theta)}, \quad (\text{A11})$$

where the subscript “det” stands for detected events and the factor  $\beta(\theta)$  is a normalisation factor defining the detection efficiency of a given model with parameters  $\theta$  as (Loredo 2004; Mandel et al. 2019),

$$\beta(\theta) = \int_{\rho[h(\lambda)] > \rho_{\text{thr}}} \int_{\lambda} p(h, \lambda | \theta) dh d\lambda, \quad (\text{A12})$$

$$= \int_{\rho[h(\lambda)] > \rho_{\text{thr}}} \int_{\lambda} p(h | \lambda) \times p(\lambda | \theta) dh d\lambda, \quad (\text{A13})$$

$$= \int_{\lambda} p_{\text{det}}(\lambda) p(\lambda | \theta) d\lambda. \quad (\text{A14})$$

where  $p_{\text{det}}$  is the detection probability defined in Section 3.1.1. In practice the previous integral is evaluated with Monte Carlo integration as

$$\int_{\lambda} p_{\text{det}}(\lambda) p(\lambda | \theta) d\lambda \approx \frac{1}{N} \sum_{i=1}^N p_{\text{det}}(\lambda_i), \quad (\text{A15})$$

where  $\lambda_i$  are samples from  $p(\lambda | \theta)$ .

## REFERENCES

- Abadie, J., et al. 2010, *Class. Quant. Grav.*, 27, 173001, doi: [10.1088/0264-9381/27/17/173001](https://doi.org/10.1088/0264-9381/27/17/173001)
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, *Physical Review X*, 6, 041015, doi: [10.1103/PhysRevX.6.041015](https://doi.org/10.1103/PhysRevX.6.041015)
- Abbott, B. P., et al. 2016, *Astrophys. J.*, 833, L1, doi: [10.3847/2041-8205/833/1/L1](https://doi.org/10.3847/2041-8205/833/1/L1)
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, *Physical Review Letters*, 119, 161101, doi: [10.1103/PhysRevLett.119.161101](https://doi.org/10.1103/PhysRevLett.119.161101)
- . 2017b, *ApJL*, 848, L12, doi: [10.3847/2041-8213/aa91c9](https://doi.org/10.3847/2041-8213/aa91c9)
- Abbott, B. P., et al. 2017, *Class. Quant. Grav.*, 34, 044001, doi: [10.1088/1361-6382/aa51f4](https://doi.org/10.1088/1361-6382/aa51f4)
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2018a, *ArXiv e-prints*. <https://arxiv.org/abs/1811.12907>
- . 2018b, *ArXiv e-prints*. <https://arxiv.org/abs/1811.12940>
- Abbott, B. P., et al. 2018, *Living Rev. Rel.*, 21, 3, doi: [10.1007/s41114-018-0012-9](https://doi.org/10.1007/s41114-018-0012-9), [10.1007/lrr-2016-1](https://doi.org/10.1007/lrr-2016-1)
- Antonini, F., & Perets, H. B. 2012, *ApJ*, 757, 27, doi: [10.1088/0004-637X/757/1/27](https://doi.org/10.1088/0004-637X/757/1/27)
- Antonini, F., & Rasio, F. A. 2016, *ApJ*, 831, 187, doi: [10.3847/0004-637X/831/2/187](https://doi.org/10.3847/0004-637X/831/2/187)
- Antonini, F., Toonen, S., & Hamers, A. S. 2017, *ApJ*, 841, 77, doi: [10.3847/1538-4357/aa6f5e](https://doi.org/10.3847/1538-4357/aa6f5e)
- Askar, A., Szkudlarek, M., Gondek-Rosińska, D., Giersz, M., & Bulik, T. 2017, *MNRAS*, 464, L36, doi: [10.1093/mnrasl/slz177](https://doi.org/10.1093/mnrasl/slz177)
- Banerjee, S. 2017, *MNRAS*, 467, 524, doi: [10.1093/mnras/stw3392](https://doi.org/10.1093/mnras/stw3392)
- . 2018, *MNRAS*, 473, 909, doi: [10.1093/mnras/stx2347](https://doi.org/10.1093/mnras/stx2347)
- Banerjee, S., Baumgardt, H., & Kroupa, P. 2010, *MNRAS*, 402, 371, doi: [10.1111/j.1365-2966.2009.15880.x](https://doi.org/10.1111/j.1365-2966.2009.15880.x)
- Barrett, J. W., Gaebel, S. M., Neijssel, C. J., et al. 2018, *Mon. Not. Roy. Astron. Soc.*, 477, 4685, doi: [10.1093/mnras/sty908](https://doi.org/10.1093/mnras/sty908)
- Bartos, I., Kocsis, B., Haiman, Z., & Márka, S. 2017, *ApJ*, 835, 165, doi: [10.3847/1538-4357/835/2/165](https://doi.org/10.3847/1538-4357/835/2/165)
- Belczynski, K., Buonanno, A., Cantiello, M., et al. 2014, *ApJ*, 789, 120, doi: [10.1088/0004-637X/789/2/120](https://doi.org/10.1088/0004-637X/789/2/120)
- Belczynski, K., Holz, D. E., Bulik, T., & O’Shaughnessy, R. 2016a, *Nature*, 534, 512, doi: [10.1038/nature18322](https://doi.org/10.1038/nature18322)
- Belczynski, K., Kalogera, V., & Bulik, T. 2002, *ApJ*, 572, 407, doi: [10.1086/340304](https://doi.org/10.1086/340304)
- Belczynski, K., Repetto, S., Holz, D. E., et al. 2016b, *ApJ*, 819, 108, doi: [10.3847/0004-637X/819/2/108](https://doi.org/10.3847/0004-637X/819/2/108)
- Bethe, H. A., & Brown, G. E. 1998, *ApJ*, 506, 780, doi: [10.1086/306265](https://doi.org/10.1086/306265)
- Bird, S., Cholis, I., Muñoz, J. B., et al. 2016, *Physical Review Letters*, 116, 201301, doi: [10.1103/PhysRevLett.116.201301](https://doi.org/10.1103/PhysRevLett.116.201301)
- Breivik, K., Rodriguez, C. L., Larson, S. L., Kalogera, V., & Rasio, F. A. 2016, *Astrophys. J.*, 830, L18, doi: [10.3847/2041-8205/830/1/L18](https://doi.org/10.3847/2041-8205/830/1/L18)
- Carr, B., Kühnel, F., & Sandstad, M. 2016, *PhRvD*, 94, 083504, doi: [10.1103/PhysRevD.94.083504](https://doi.org/10.1103/PhysRevD.94.083504)
- Carr, B. J., & Hawking, S. W. 1974, *MNRAS*, 168, 399, doi: [10.1093/mnras/168.2.399](https://doi.org/10.1093/mnras/168.2.399)
- Chen, H.-Y., Holz, D. E., Miller, J., et al. 2017, *arXiv e-prints*. <https://arxiv.org/abs/1709.08079>



- Chen, Y., Bressan, A., Girardi, L., et al. 2015, *MNRAS*, 452, 1068, doi: [10.1093/mnras/stv1281](https://doi.org/10.1093/mnras/stv1281)
- Claeys, J. S. W., Pols, O. R., Izzard, R. G., Vink, J., & Verbunt, F. W. M. 2014, *A&A*, 563, A83, doi: [10.1051/0004-6361/201322714](https://doi.org/10.1051/0004-6361/201322714)
- Dal Canton, T., et al. 2014, *Phys. Rev.*, D90, 082004, doi: [10.1103/PhysRevD.90.082004](https://doi.org/10.1103/PhysRevD.90.082004)
- Di Carlo, U. N., Giacobbo, N., Mapelli, M., et al. 2019, arXiv e-prints. <https://arxiv.org/abs/1901.00863>
- Dominik, M., Belczynski, K., Fryer, C., et al. 2013, *ApJ*, 779, 72, doi: [10.1088/0004-637X/779/1/72](https://doi.org/10.1088/0004-637X/779/1/72)
- Dominik, M., Berti, E., O’Shaughnessy, R., et al. 2015, *Astrophys. J.*, 806, 263, doi: [10.1088/0004-637X/806/2/263](https://doi.org/10.1088/0004-637X/806/2/263)
- Downing, J. M. B., Benacquista, M. J., Giersz, M., & Spurzem, R. 2010, *MNRAS*, 407, 1946, doi: [10.1111/j.1365-2966.2010.17040.x](https://doi.org/10.1111/j.1365-2966.2010.17040.x)
- . 2011, *MNRAS*, 416, 133, doi: [10.1111/j.1365-2966.2011.19023.x](https://doi.org/10.1111/j.1365-2966.2011.19023.x)
- Eldridge, J. J., & Stanway, E. R. 2016, *MNRAS*, 462, 3302, doi: [10.1093/mnras/stw1772](https://doi.org/10.1093/mnras/stw1772)
- Eldridge, J. J., Stanway, E. R., & Tang, P. N. 2019, *MNRAS*, 482, 870, doi: [10.1093/mnras/sty2714](https://doi.org/10.1093/mnras/sty2714)
- Eldridge, J. J., Stanway, E. R., Xiao, L., et al. 2017, *PASA*, 34, e058, doi: [10.1017/pasa.2017.51](https://doi.org/10.1017/pasa.2017.51)
- Farr, W. M., Stevenson, S., Coleman Miller, M., et al. 2017, *Nature*, 548, 426, doi: [10.1038/nature23453](https://doi.org/10.1038/nature23453)
- Finn, L. S. 1996, *Phys. Rev. D*, 53, 2878, doi: [10.1103/PhysRevD.53.2878](https://doi.org/10.1103/PhysRevD.53.2878)
- Finn, L. S., & Chernoff, D. F. 1993, *Phys. Rev.*, D47, 2198, doi: [10.1103/PhysRevD.47.2198](https://doi.org/10.1103/PhysRevD.47.2198)
- Fishbach, M., Holz, D. E., & Farr, B. 2017, *ApJL*, 840, L24, doi: [10.3847/2041-8213/aa7045](https://doi.org/10.3847/2041-8213/aa7045)
- Fishbach, M., Holz, D. E., & Farr, W. M. 2018, *Astrophys. J.*, 863, L41, doi: [10.3847/2041-8213/aad800](https://doi.org/10.3847/2041-8213/aad800)
- Fragione, G., & Kocsis, B. 2018, *Physical Review Letters*, 121, 161103, doi: [10.1103/PhysRevLett.121.161103](https://doi.org/10.1103/PhysRevLett.121.161103)
- Fryer, C. L., Belczynski, K., Wiktorowicz, G., et al. 2012, *ApJ*, 749, 91, doi: [10.1088/0004-637X/749/1/91](https://doi.org/10.1088/0004-637X/749/1/91)
- Gair, J. R., Sesana, A., Berti, E., & Volonteri, M. 2011, *Class. Quant. Grav.*, 28, 094018, doi: [10.1088/0264-9381/28/9/094018](https://doi.org/10.1088/0264-9381/28/9/094018)
- Gerosa, D. 2019, gwDET: Detectability of gravitational-wave signals from compact binary coalescences, doi: [doi.org/10.5281/zenodo.889966](https://doi.org/10.5281/zenodo.889966).
- Gerosa, D., & Berti, E. 2017, *Phys. Rev.*, D95, 124046, doi: [10.1103/PhysRevD.95.124046](https://doi.org/10.1103/PhysRevD.95.124046)
- Gerosa, D., Berti, E., O’Shaughnessy, R., et al. 2018, *PhRvD*, 98, 084036, doi: [10.1103/PhysRevD.98.084036](https://doi.org/10.1103/PhysRevD.98.084036)
- Gerosa, D., Kesden, M., Berti, E., O’Shaughnessy, R., & Sperhake, U. 2013, *PhRvD*, 87, 104028, doi: [10.1103/PhysRevD.87.104028](https://doi.org/10.1103/PhysRevD.87.104028)
- Ghosh, A., Del Pozzo, W., & Ajith, P. 2016, *Phys. Rev.*, D94, 104070, doi: [10.1103/PhysRevD.94.104070](https://doi.org/10.1103/PhysRevD.94.104070)
- Giacobbo, N., & Mapelli, M. 2018, *MNRAS*, 480, 2011, doi: [10.1093/mnras/sty1999](https://doi.org/10.1093/mnras/sty1999)
- . 2019, *MNRAS*, 482, 2234, doi: [10.1093/mnras/sty2848](https://doi.org/10.1093/mnras/sty2848)
- Giacobbo, N., Mapelli, M., & Spera, M. 2018, *MNRAS*, 474, 2959, doi: [10.1093/mnras/stx2933](https://doi.org/10.1093/mnras/stx2933)
- Gräfenner, G., Vink, J. S., de Koter, A., & Langer, N. 2011, *A&A*, 535, A56, doi: [10.1051/0004-6361/201116701](https://doi.org/10.1051/0004-6361/201116701)
- Gutermuth, R. A., Megeath, S. T., Pipher, J. L., et al. 2005, *ApJ*, 632, 397, doi: [10.1086/432460](https://doi.org/10.1086/432460)
- Harris, W. E., Harris, G. L. H., & Alessi, M. 2013, *ApJ*, 772, 82, doi: [10.1088/0004-637X/772/2/82](https://doi.org/10.1088/0004-637X/772/2/82)
- Hénon, M. H. 1971, *Ap&SS*, 14, 151, doi: [10.1007/BF00649201](https://doi.org/10.1007/BF00649201)
- Hills, J. G., & Fullerton, L. W. 1980, *AJ*, 85, 1281, doi: [10.1086/112798](https://doi.org/10.1086/112798)
- Hobbs, G., Lorimer, D. R., Lyne, A. G., & Kramer, M. 2005, *MNRAS*, 360, 974, doi: [10.1111/j.1365-2966.2005.09087.x](https://doi.org/10.1111/j.1365-2966.2005.09087.x)
- Hurley, J. R., Pols, O. R., & Tout, C. A. 2000, *MNRAS*, 315, 543, doi: [10.1046/j.1365-8711.2000.03426.x](https://doi.org/10.1046/j.1365-8711.2000.03426.x)
- Hurley, J. R., Tout, C. A., & Pols, O. R. 2002, *MNRAS*, 329, 897, doi: [10.1046/j.1365-8711.2002.05038.x](https://doi.org/10.1046/j.1365-8711.2002.05038.x)
- Inayoshi, K., Kashiyaama, K., Visbal, E., & Haiman, Z. 2016, *MNRAS*, 461, 2722, doi: [10.1093/mnras/stw1431](https://doi.org/10.1093/mnras/stw1431)
- Inomata, K., Kawasaki, M., Mukaida, K., Tada, Y., & Yanagida, T. T. 2017, *PhRvD*, 95, 123510, doi: [10.1103/PhysRevD.95.123510](https://doi.org/10.1103/PhysRevD.95.123510)
- Joshi, K. J., Rasio, F. A., & Portegies Zwart, S. 2000, *ApJ*, 540, 969, doi: [10.1086/309350](https://doi.org/10.1086/309350)
- Kalogera, V. 2000, *ApJ*, 541, 319, doi: [10.1086/309400](https://doi.org/10.1086/309400)
- Khan, S., Husa, S., Hannam, M., et al. 2016, *Phys. Rev.*, D93, 044007, doi: [10.1103/PhysRevD.93.044007](https://doi.org/10.1103/PhysRevD.93.044007)
- Kimball, C., Berry, C. P. L., & Kalogera, V. 2019, <https://arxiv.org/abs/1903.07813>
- Kimpson, T. O., Spera, M., Mapelli, M., & Ziosi, B. M. 2016, *MNRAS*, 463, 2443, doi: [10.1093/mnras/stw2085](https://doi.org/10.1093/mnras/stw2085)
- Kozai, Y. 1962, *AJ*, 67, 591, doi: [10.1086/108790](https://doi.org/10.1086/108790)
- Kroupa, P. 2001, *MNRAS*, 322, 231, doi: [10.1046/j.1365-8711.2001.04022.x](https://doi.org/10.1046/j.1365-8711.2001.04022.x)
- Kruckow, M. U., Tauris, T. M., Langer, N., Kramer, M., & Izzard, R. G. 2018, *MNRAS*, 481, 1908, doi: [10.1093/mnras/sty2190](https://doi.org/10.1093/mnras/sty2190)
- Kumamoto, J., Fujii, M. S., & Tanikawa, A. 2019, *MNRAS*, 486, 3942, doi: [10.1093/mnras/stz1068](https://doi.org/10.1093/mnras/stz1068)

- Küpper, A. H. W., Maschberger, T., Kroupa, P., & Baumgardt, H. 2011, *MNRAS*, 417, 2300, doi: [10.1111/j.1365-2966.2011.19412.x](https://doi.org/10.1111/j.1365-2966.2011.19412.x)
- Lada, C. J., & Lada, E. A. 2003, *ARA&A*, 41, 57, doi: [10.1146/annurev.astro.41.011802.094844](https://doi.org/10.1146/annurev.astro.41.011802.094844)
- Lamberts, A., Garrison-Kimmel, S., Clausen, D. R., & Hopkins, P. F. 2016, *MNRAS*, 463, L31, doi: [10.1093/mnras/rlw152](https://doi.org/10.1093/mnras/rlw152)
- Lidov, M. L. 1962, *Planet. Space Sci.*, 9, 719, doi: [10.1016/0032-0633\(62\)90129-0](https://doi.org/10.1016/0032-0633(62)90129-0)
- Lored, T. J. 2004, *AIP Conf. Proc.*, 735, 195, doi: [10.1063/1.1835214](https://doi.org/10.1063/1.1835214)
- Mandel, I., & de Mink, S. E. 2016, *MNRAS*, 458, 2634, doi: [10.1093/mnras/stw379](https://doi.org/10.1093/mnras/stw379)
- Mandel, I., Farr, W. M., Colonna, A., et al. 2017, *MNRAS*, 465, 3254, doi: [10.1093/mnras/stw2883](https://doi.org/10.1093/mnras/stw2883)
- Mandel, I., Farr, W. M., & Gair, J. R. 2019, *MNRAS*, 486, 1086, doi: [10.1093/mnras/stz896](https://doi.org/10.1093/mnras/stz896)
- Mandel, I., Haster, C.-J., Dominik, M., & Belczynski, K. 2015, *MNRAS*, 450, L85, doi: [10.1093/mnrasl/slv054](https://doi.org/10.1093/mnrasl/slv054)
- Mapelli, M. 2016, *MNRAS*, 459, 3432, doi: [10.1093/mnras/stw869](https://doi.org/10.1093/mnras/stw869)
- . 2018, arXiv e-prints. <https://arxiv.org/abs/1809.09130>
- Mapelli, M., & Giacobbo, N. 2018, *MNRAS*, 479, 4391, doi: [10.1093/mnras/sty1613](https://doi.org/10.1093/mnras/sty1613)
- Mapelli, M., Giacobbo, N., Ripamonti, E., & Spera, M. 2017, *MNRAS*, 472, 2422, doi: [10.1093/mnras/stx2123](https://doi.org/10.1093/mnras/stx2123)
- Mapelli, M., Giacobbo, N., Santoliquido, F., & Artale, M. C. 2019, *MNRAS*, doi: [10.1093/mnras/stz1150](https://doi.org/10.1093/mnras/stz1150)
- Mapelli, M., Zampieri, L., Ripamonti, E., & Bressan, A. 2013, *MNRAS*, 429, 2298, doi: [10.1093/mnras/sts500](https://doi.org/10.1093/mnras/sts500)
- Marchant, P., Langer, N., Podsiadlowski, P., Tauris, T. M., & Moriya, T. J. 2016, *A&A*, 588, A50, doi: [10.1051/0004-6361/201628133](https://doi.org/10.1051/0004-6361/201628133)
- McKernan, B., Ford, K. E. S., Kocsis, B., Lyra, W., & Winter, L. M. 2014, *MNRAS*, 441, 900, doi: [10.1093/mnras/stu553](https://doi.org/10.1093/mnras/stu553)
- McKernan, B., Ford, K. E. S., Lyra, W., & Perets, H. B. 2012, *MNRAS*, 425, 460, doi: [10.1111/j.1365-2966.2012.21486.x](https://doi.org/10.1111/j.1365-2966.2012.21486.x)
- McKernan, B., Ford, K. E. S., Bellovary, J., et al. 2018, *ApJ*, 866, 66, doi: [10.3847/1538-4357/aadae5](https://doi.org/10.3847/1538-4357/aadae5)
- Mennekens, N., & Vanbeveren, D. 2014, *A&A*, 564, A134, doi: [10.1051/0004-6361/201322198](https://doi.org/10.1051/0004-6361/201322198)
- Miller, M. C., & Miller, J. M. 2015, *PhR*, 548, 1, doi: [10.1016/j.physrep.2014.09.003](https://doi.org/10.1016/j.physrep.2014.09.003)
- Nishizawa, A., Berti, E., Klein, A., & Sesana, A. 2016, *Phys. Rev.*, D94, 064020, doi: [10.1103/PhysRevD.94.064020](https://doi.org/10.1103/PhysRevD.94.064020)
- Nishizawa, A., Sesana, A., Berti, E., & Klein, A. 2017, *Mon. Not. Roy. Astron. Soc.*, 465, 4375, doi: [10.1093/mnras/stw2993](https://doi.org/10.1093/mnras/stw2993)
- O’Leary, R. M., Kocsis, B., & Loeb, A. 2009, *MNRAS*, 395, 2127, doi: [10.1111/j.1365-2966.2009.14653.x](https://doi.org/10.1111/j.1365-2966.2009.14653.x)
- O’Leary, R. M., Rasio, F. A., Fregeau, J. M., Ivanova, N., & O’Shaughnessy, R. 2006, *ApJ*, 637, 937, doi: [10.1086/498446](https://doi.org/10.1086/498446)
- O’Shaughnessy, R., Gerosa, D., & Wysocki, D. 2017, *Physical Review Letters*, 119, 011101, doi: [10.1103/PhysRevLett.119.011101](https://doi.org/10.1103/PhysRevLett.119.011101)
- Petrovich, C., & Antonini, F. 2017, *ApJ*, 846, 146, doi: [10.3847/1538-4357/aa8628](https://doi.org/10.3847/1538-4357/aa8628)
- Poisson, E., & Will, C. M. 1995, *Phys. Rev.*, D52, 848, doi: [10.1103/PhysRevD.52.848](https://doi.org/10.1103/PhysRevD.52.848)
- Portegies Zwart, S. F., & McMillan, S. L. W. 2000, *ApJL*, 528, L17, doi: [10.1086/312422](https://doi.org/10.1086/312422)
- Portegies Zwart, S. F., McMillan, S. L. W., & Gieles, M. 2010, *ARA&A*, 48, 431, doi: [10.1146/annurev-astro-081309-130834](https://doi.org/10.1146/annurev-astro-081309-130834)
- Powell, J., Stevenson, S., Mandel, I., & Tino, P. 2019, <https://arxiv.org/abs/1905.04825>
- Rasskazov, A., & Kocsis, B. 2019, arXiv e-prints. <https://arxiv.org/abs/1902.03242>
- Rodriguez, C. L., Chatterjee, S., & Rasio, F. A. 2016a, *PhRvD*, 93, 084029, doi: [10.1103/PhysRevD.93.084029](https://doi.org/10.1103/PhysRevD.93.084029)
- Rodriguez, C. L., Haster, C.-J., Chatterjee, S., Kalogera, V., & Rasio, F. A. 2016b, *ApJL*, 824, L8, doi: [10.3847/2041-8205/824/1/L8](https://doi.org/10.3847/2041-8205/824/1/L8)
- Rodriguez, C. L., & Loeb, A. 2018, *ApJL*, 866, L5, doi: [10.3847/2041-8213/aae377](https://doi.org/10.3847/2041-8213/aae377)
- Rodriguez, C. L., Morscher, M., Pattabiraman, B., et al. 2015, *Physical Review Letters*, 115, 051101, doi: [10.1103/PhysRevLett.115.051101](https://doi.org/10.1103/PhysRevLett.115.051101)
- Rodriguez, C. L., Zevin, M., Pankow, C., Kalogera, V., & Rasio, F. A. 2016c, *ApJL*, 832, L2, doi: [10.3847/2041-8205/832/1/L2](https://doi.org/10.3847/2041-8205/832/1/L2)
- Roulet, J., & Zaldarriaga, M. 2019, *MNRAS*, 484, 4216, doi: [10.1093/mnras/stz226](https://doi.org/10.1093/mnras/stz226)
- Sadowski, A., Belczynski, K., Bulik, T., et al. 2008, *ApJ*, 676, 1162, doi: [10.1086/528932](https://doi.org/10.1086/528932)
- Samsing, J. 2018, *PhRvD*, 97, 103014, doi: [10.1103/PhysRevD.97.103014](https://doi.org/10.1103/PhysRevD.97.103014)
- Samsing, J., Askar, A., & Giersz, M. 2018, *ApJ*, 855, 124, doi: [10.3847/1538-4357/aaab52](https://doi.org/10.3847/1538-4357/aaab52)
- Sana, H., de Mink, S. E., de Koter, A., et al. 2012, *Science*, 337, 444, doi: [10.1126/science.1223344](https://doi.org/10.1126/science.1223344)
- Sasaki, M., Suyama, T., Tanaka, T., & Yokoyama, S. 2016, *Physical Review Letters*, 117, 061101, doi: [10.1103/PhysRevLett.117.061101](https://doi.org/10.1103/PhysRevLett.117.061101)

- Scelfo, G., Bellomo, N., Raccanelli, A., Matarrese, S., & Verde, L. 2018, JCAP, 9, 039, doi: [10.1088/1475-7516/2018/09/039](https://doi.org/10.1088/1475-7516/2018/09/039)
- Sesana, A., Gair, J., Berti, E., & Volonteri, M. 2011, Phys. Rev., D83, 044036, doi: [10.1103/PhysRevD.83.044036](https://doi.org/10.1103/PhysRevD.83.044036)
- Spera, M., & Mapelli, M. 2017, MNRAS, 470, 4739, doi: [10.1093/mnras/stx1576](https://doi.org/10.1093/mnras/stx1576)
- Spera, M., Mapelli, M., & Bressan, A. 2015, MNRAS, 451, 4086, doi: [10.1093/mnras/stv1161](https://doi.org/10.1093/mnras/stv1161)
- Spera, M., Mapelli, M., Giacobbo, N., et al. 2019, MNRAS, 485, 889, doi: [10.1093/mnras/stz359](https://doi.org/10.1093/mnras/stz359)
- Stevenson, S., Berry, C. P. L., & Mandel, I. 2017a, MNRAS, 471, 2801, doi: [10.1093/mnras/stx1764](https://doi.org/10.1093/mnras/stx1764)
- Stevenson, S., Ohme, F., & Fairhurst, S. 2015, ApJ, 810, 58, doi: [10.1088/0004-637X/810/1/58](https://doi.org/10.1088/0004-637X/810/1/58)
- Stevenson, S., Vigna-Gómez, A., Mandel, I., et al. 2017b, Nature Communications, 8, 14906, doi: [10.1038/ncomms14906](https://doi.org/10.1038/ncomms14906)
- Stone, N. C., Küpper, A. H. W., & Ostriker, J. P. 2017a, MNRAS, 467, 4180, doi: [10.1093/mnras/stx097](https://doi.org/10.1093/mnras/stx097)
- Stone, N. C., Metzger, B. D., & Haiman, Z. 2017b, MNRAS, 464, 946, doi: [10.1093/mnras/stw2260](https://doi.org/10.1093/mnras/stw2260)
- Sukhbold, T., Ertl, T., Woosley, S. E., Brown, J. M., & Janka, H.-T. 2016, ApJ, 821, 38, doi: [10.3847/0004-637X/821/1/38](https://doi.org/10.3847/0004-637X/821/1/38)
- Tagawa, H., & Umemura, M. 2018, ApJ, 856, 47, doi: [10.3847/1538-4357/aab0a4](https://doi.org/10.3847/1538-4357/aab0a4)
- Talbot, C., & Thrane, E. 2017, Phys. Rev., D96, 023012, doi: [10.1103/PhysRevD.96.023012](https://doi.org/10.1103/PhysRevD.96.023012)
- . 2018, Astrophys. J., 856, 173, doi: [10.3847/1538-4357/aab34c](https://doi.org/10.3847/1538-4357/aab34c)
- Taylor, S. R., & Gerosa, D. 2018, Phys. Rev., D98, 083017, doi: [10.1103/PhysRevD.98.083017](https://doi.org/10.1103/PhysRevD.98.083017)
- Usman, S. A., et al. 2016, Class. Quant. Grav., 33, 215004, doi: [10.1088/0264-9381/33/21/215004](https://doi.org/10.1088/0264-9381/33/21/215004)
- Vitale, S., Gerosa, D., Haster, C.-J., Chatziioannou, K., & Zimmerman, A. 2017a, Physical Review Letters, 119, 251103, doi: [10.1103/PhysRevLett.119.251103](https://doi.org/10.1103/PhysRevLett.119.251103)
- Vitale, S., Lynch, R., Sturani, R., & Graff, P. 2017b, Classical and Quantum Gravity, 34, 03LT01, doi: [10.1088/1361-6382/aa552e](https://doi.org/10.1088/1361-6382/aa552e)
- Wang, L., Spurzem, R., Aarseth, S., et al. 2015, MNRAS, 450, 4070, doi: [10.1093/mnras/stv817](https://doi.org/10.1093/mnras/stv817)
- Weidner, C., & Kroupa, P. 2006, MNRAS, 365, 1333, doi: [10.1111/j.1365-2966.2005.09824.x](https://doi.org/10.1111/j.1365-2966.2005.09824.x)
- Weidner, C., Kroupa, P., & Bonnell, I. A. D. 2010, MNRAS, 401, 275, doi: [10.1111/j.1365-2966.2009.15633.x](https://doi.org/10.1111/j.1365-2966.2009.15633.x)
- Wysocki, D., Gerosa, D., O’Shaughnessy, R., et al. 2018, PhRvD, 97, 043014, doi: [10.1103/PhysRevD.97.043014](https://doi.org/10.1103/PhysRevD.97.043014)
- Zevin, M., Pankow, C., Rodriguez, C. L., et al. 2017, ApJ, 846, 82, doi: [10.3847/1538-4357/aa8408](https://doi.org/10.3847/1538-4357/aa8408)
- Zevin, M., Samsing, J., Rodriguez, C., Haster, C.-J., & Ramirez-Ruiz, E. 2019, ApJ, 871, 91, doi: [10.3847/1538-4357/aaf6ec](https://doi.org/10.3847/1538-4357/aaf6ec)
- Ziosi, B. M., Mapelli, M., Branchesi, M., & Tormen, G. 2014, MNRAS, 441, 3703, doi: [10.1093/mnras/stu824](https://doi.org/10.1093/mnras/stu824)