MNRAS **000**, 1–22 (xxxx) Preprint 20 June 2019 Compiled using MNRAS LATEX style file v3.0

The effect of the metallicity-specific star formation history on double compact object mergers

Coenraad J. Neijssel,[1,7] ⋆ Alejandro Vigna-Gómez,[1,5,6] Simon Stevenson,[2,6] Jim W. Barrett,[1,9] Sebastian M. Gaebel,[1] Floor Broekgaarden,[3,4,5,6] Selma E. de Mink,[3,4] Dorottya Szécsi,[1,10] Serena Vinciguerra[1,7,8] and Ilya Mandel[5,6,1]

[1] *Birmingham Institute for Gravitational Wave Astronomy and School of Physics and Astronomy, University of Birmingham, Birmingham, B15 2TT, United Kingdom*
[2] *Centre for astrophysics and supercomputing, Swinburne University of Technology, Hawthorn VIC 3122, Australia*
[3] *Anton Pannekoek Institute for Astronomy, University of Amsterdam, Postbus 94249, 1090 GE Amsterdam, The Netherlands*
[4] *GRAPPA, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands*
[5] *Monash Centre for Astrophysics, School of Physics and Astronomy, Monash University, Clayton, Victoria 3800, Australia*
[6] *The ARC Center of Excellence for Gravitational Wave Discovery – OzGrav*
[7] *Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik, D-30167 Hannover, Germany*
[8] *Leibniz Universität Hannover, D-30167 Hannover, Germany*
[9] *Klarna Bank AB (publ). Sveavägen 46, 111 34 Stockholm*
[10] *I. Physikalisches Institut, Universität zu Köln, Zülpicher-Strasse 77, D-50937 Cologne, Germany*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We investigate the impact of uncertainty in the metallicity-specific star formation rate over cosmic time on predictions of the rates and masses of double compact object mergers observable through gravitational waves. We find that this uncertainty can change the predicted detectable merger rate by more than an order of magnitude, comparable to contributions from uncertain physical assumptions regarding binary evolution, such as mass transfer efficiency or supernova kicks. We statistically compare the results produced by the COMPAS population synthesis suite against a catalog of gravitational-wave detections from the first two Advanced LIGO and Virgo observing runs. We find that the rate and chirp mass of observed binary black hole mergers can be well matched under our default evolutionary model with a star formation metallicity spread of 0.39 dex around a mean metallicity $\langle Z \rangle$ that scales with redshift z as $\langle Z \rangle = 0.035 \times 10^{-0.23z}$, assuming a star formation rate of $0.01 \times (1+z)^{2.77}/(1+((1+z)/2.9)^{4.7})\,\mathrm{M}_\odot\ \mathrm{Mpc}^{-3}\ \mathrm{yr}^{-1}$. Intriguingly, this default model predicts that 80% of the approximately one binary black hole merger per day that will be detectable at design sensitivity will have formed through isolated binary evolution with only dynamically stable mass transfer, i.e., without experiencing a common-envelope event.

Key words: gravitational waves – binaries – cosmology

1 INTRODUCTION

There were 10 binary black hole (BBH) detections (Abbott et al. 2016; Abbott et al. 2018a) and a binary neutron star (BNS) (Abbott et al. 2017) in the first and second observing runs of the advanced Laser Interferometer Gravitational-wave Observatory (aLIGO) and Virgo gravitational-wave detectors. The intrinsic rate of BBH mergers is currently estimated by the LIGO-Virgo collaboration at 24–112 $\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}$, whereas for BNSs it is 110–3840 $\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}$ (Abbott et al. 2018b). These intrinsic rate estimates depend on the assumed shape of the mass and rate distribution of the double compact object (DCO) mergers, which remains uncertain. Multiple possible stellar origins exist for DCOs such as dynamical capture in open/globular/nuclear clusters, Lidov-Kozai resonances in hierarchical triples, chemically homogeneous evolution in compact stellar binaries, and mergers of primordial black

⋆ E-mail: cneijssel@star.sr.bham.ac.uk

holes (see Miller 2016; Mandel & Farmer 2018; Giacobbo & Mapelli 2018, for reviews). We focus on the merger rate of DCOs that come from isolated binary evolution. It appears that most of the massive stars ($M > 8$ $M_\odot$) in the field are born in binaries (Kiminki & Kobulnicky 2012; Sana et al. 2013; Moe & Di Stefano 2017). Once formed, these isolated binaries evolve without external influences and a fraction becomes DCOs. However, the exact physics of stellar and binary evolution and the resulting rates of DCO mergers are still uncertain (e.g., Dominik et al. 2015; Eldridge & Stanway 2016; Kruckow et al. 2018; Chruslinska et al. 2018).

The evolution of massive stars takes a few million years, but their inspiral as DCOs can span years to billions of years (e.g., Portegies Zwart & Yungelson 1998; Belczynski et al. 2002b; Eldridge & Stanway 2016; Mapelli et al. 2017). The detected mergers could therefore have formed at very high redshifts. Observations show that the star formation rate (SFR) changes significantly as a function of redshift (Madau & Dickinson 2014). At redshifts $z \gtrsim 2$ the SFR estimates become increasingly more sensitive to the assumed extinction, which is uncertain (Madau & Dickinson 2014; Strolger et al. 2004). The SFR determines the amount of stellar binaries formed and hence introduces an uncertainty on the rate of DCO formation.

Metallicity, and particularly the fraction of iron in the star at birth, significantly impacts the rate of mass loss through line-driven winds. Consequently, it has a significant effect on the DCO mass distribution and merger rate (Belczynski et al. 2010; Stevenson et al. 2017; Giacobbo et al. 2018). The metallicity of star forming gas depends on redshift, as subsequent generations of stars enrich the interstellar medium through winds and explosions with metals formed during their evolution. Galaxy catalogs, such as the Sloan Digital Sky Survey (Tremonti et al. 2004), show that there is an empirical correlation between the galaxy stellar mass and the mean metallicity of the galaxy. Furthermore, the galaxy stellar-mass function (GSMF) and the galaxy stellar mass – metallicity (MZ)-relation evolve with redshift. Different calibrations or galaxy samples lead to different results (Savaglio et al. 2005; Kewley & Ellison 2008; Furlong et al. 2015). Thus there is not only uncertainty in the overall SFR but also in the distribution of the metallicities in the star forming gas. Combined, these result in an uncertainty in the metallicity-specific star formation rate (MSSFR), which affects estimates of the rates and properties of DCO mergers.

Aim and overall method

Our aim is to assess how the uncertainty in the MSSFR affects predictions for the rate and distributions of DCO mergers. In this section, we introduce the key steps in the calculation of the redshift-dependent DCO merger distribution and the rate of detectable DCO mergers.

The time it takes for a binary to evolve its stars and then merge at t_m as a DCO due to the emission of gravitational waves is called the delay time (t_delay). The formation time t_f is related to the merger time t_m by $t_\mathrm{f} = t_\mathrm{m} - t_\mathrm{delay}$. We can calculate the rate of mergers at any given time as

$$\frac{d^3N_\mathrm{merge}}{dt_\mathrm{s}dV_\mathrm{c}dM_\mathrm{chirp}}(t_\mathrm{m}) = \int dZ \int dt_\mathrm{delay} \frac{d^3N_\mathrm{form}}{dt_\mathrm{delay}dM_\mathrm{SFR}dM_\mathrm{chirp}}(Z)\frac{d^3M_\mathrm{SFR}}{dt_s\ dV_c\ dZ}(t_\mathrm{f} = t_\mathrm{m} - t_\mathrm{delay}), \quad (1)$$

where Z is the metallicity, z is the redshift, t_s is the time in the source frame of the merger, and V_c is the comoving volume. The first term in the integrand is the number of DCOs per unit star forming mass M_SFR per unit delay time and per unit chirp mass $M_\mathrm{chirp} = M_1^{3/5}M_2^{3/5}(M_1+M_2)^{-1/5}$, where M_1, M_2 are the individual compact object masses. We compute this first term over a grid of metallicities by running the COMPAS population synthesis code. The second term is the MSSFR at the birth of the binary per unit time, volume, and metallicity, which we model analytically.

The second step is to calculate the distribution of observable DCO mergers. We do this by converting t_m to a redshift z and integrating the entire visible volume in shells of thickness dz. At each redshift we calculate the probability of detecting a binary (P_det) given its chirp-mass (M_chirp) and luminosity distance ($D_\mathrm{L}(z)$). The total observable merger distribution is then

$$\frac{d^2N}{dt_\mathrm{obs}dM_\mathrm{chirp}} = \int_0^{z_\mathrm{max}} dz \frac{dt_s}{dt_\mathrm{obs}}\frac{dV_\mathrm{c}}{dz}\times \frac{d^3N_\mathrm{merge}}{dt_\mathrm{s}\ dV_\mathrm{c}dM_\mathrm{chirp}}(z)P_\mathrm{det}(M_\mathrm{chirp}, D_\mathrm{L})\ , \quad (2)$$

where dV_c/dz is the differential comoving volume as a function of redshift and $dt_\mathrm{s}/dt_\mathrm{obs} = 1/(1+z)$ translates the rate to the observer frame (e.g. Hogg 1999). We assume a flat cosmology with $\Omega_\mathrm{M} = 0.308$ and a Hubble constant of $H_0 = 67.8$ km s-1 Mpc^{-1} (Ade et al. 2016). Altogether this general method is similar to works such as Langer & Norman (2006), Dominik et al. (2013), Mandel & de Mink (2016), Eldridge & Stanway (2016), Madau & Fragos (2017) and Chruslinska et al. (2018).

We sequentially solve the aforementioned equations, structuring the paper as follows:

– Section 2: COMPAS population synthesis code
We create a large sample of DCOs from a broad range of metallicities using the rapid population synthesis element of the COMPAS suite. We briefly describe the model assumptions used to evolve our massive stellar binaries.
– Section 3: $\frac{d^3N_\mathrm{form}}{dt_\mathrm{delay}dM_\mathrm{SFR}dM_\mathrm{chirp}}$ - DCO population
We show the results of our population synthesis of DCOs. We describe some of the key features such as their mass distribution at different formation metallicities in our simulation. We describe the three main formation channels for BBHs. We find a significant number of BBHs merging without experiencing a common-envelope event.
– Section 4: $\frac{d^3M_\mathrm{SFR}}{dt_s dV_c dZ}$ - MSSFR
We combine observations and simulations of galaxy stellar mass distributions with mass – metallicity relations to construct a MSSFR. These different prescriptions introduce an uncertainty into our DCO merger rate distributions. We propose a parametrised, smooth metallicity distribution, which facilitates the exploration of the MSSFR parameter-space
– Section 5: $\frac{d^3N_\mathrm{merge}}{dt_s dV_c dM_\mathrm{chirp}}$ - DCO Merger Distributions
We calculate the redshift-dependent DCO distribution by

convolving the MSSFR with our DCO population. We find that variation in MSSFR prescriptions significantly affects both the total rate and mass distributions of DCOs mergers.

– Section 6: $\frac{d^2N}{dt_{\rm obs}dM_{\rm chirp}}$ - Gravitational-Wave Detections
We apply selection effects of gravitational-wave detectors to our cosmic DCO populations. From this we get both rate and mass distributions of detectable BBH mergers for different MSSFR prescriptions. We use a Bayesian approach to compare the predictions of different MSSFR models against the observed sample of gravitational waves from BBH mergers. We find that the MSSFR significantly affects the predicted rate of gravitational-wave events from BBH mergers.

– Section 7: Discussion and conclusion
We review our findings and discuss future prospects.

2 COMPAS POPULATION SYNTHESIS CODE

We generate our population of DCOs by modelling isolated binary evolution with the population synthesis code COMPAS (Stevenson et al. 2017; Barrett et al. 2018; Vigna-Gómez et al. 2018; Stevenson et al. 2019). We use Monte Carlo simulations to empirically estimate the rate density of DCOs per unit star forming mass in delay time and chirp mass at each simulation metallicity:

$$\frac{d^3N_{\rm form}}{dt_{\rm delay}dM_{\rm SFR}dM_{\rm chirp}}(Z, t_{\rm delay}, M_{\rm chirp}).$$

In this section we briefly describe the parameter space of our simulation and our model assumptions for isolated binary evolution. The data will be made publicly available at `http://compas.science`.

2.1 Initial Distributions

The five initial conditions that describe a stellar binary are: the primary m_1 and secondary m_2 masses, the orbital separation a, the orbital eccentricity e, and the metallicity of the stars Z at zero-age main sequence (ZAMS). The mass of the initially more massive star, the primary, is drawn from an initial mass function (IMF) according to Kroupa (2001). The mass of the initially less massive secondary star is given by

$$m_2 = m_1 \times q, \tag{3}$$

where q is the initial mass ratio ($0 < q < 1$). We draw the mass ratio q from a flat distribution (Sana et al. 2012). We assume that the distribution of separations is flat-in-the-log ($0.1 < a/\mathrm{AU} < 1000$) (Opik 1924) and the orbits are all circular at birth. We assume that these distributions are both independent of each other as well as independent of metallicity. Recent studies such as Moe & Di Stefano (2017) suggest that the initial distributions might be correlated. de Mink & Belczynski (2015); Klencki et al. (2018) found that varying initial condition distributions affects DCO merger rates by factors of $\lesssim 2$.

For the metallicities of the binaries we use 30 grid points spread uniformly in log-space over a broad range of metal mass fractions $0.0001 \leq \mathrm{Z} \leq 0.03$. We evolve three million binaries with a total star forming mass of the order of $6.5 \times 10^7\ \mathrm{M}_\odot$ per grid-point.

To optimise the number of compact objects per binary simulated, whilst still leaving enough room in the parameter space to avoid boundary effects, we draw primaries with masses equal or bigger than 5 $\mathrm{M}_\odot$ (this represents a very naive version of importance sampling introduced by Broekgaarden et al. 2019). Our upper mass limit is 150 solar masses. In this mass range the power index of the IMF equals -2.3. Hence we need to correct for the 'true' amount of mass evolved in all stars (both single and binary). We calculate this by assuming a binary fraction of 70 per cent and a flat mass ratio for all stellar masses (Sana et al. 2012). This results in a total star forming mass per metallicity grid point of $\sim 3.1 \times 10^8\,\mathrm{M}_\odot$. It is this star forming mass that we use as our normalisation $dM_{\rm SFR}$.

2.2 Single Stellar Models

Stellar evolution in COMPAS is based on the stellar models by Pols et al. (1998). We use analytical fits to these models by Hurley et al. (2000, 2002) to rapidly evolve binaries. Our wind mass loss rates for stars with temperatures below 12500 K are prescribed by Hurley et al. (2000) and references therein. For hot massive stars (T > 12500K) we use the wind mass loss rates by Vink et al. (2001) as implemented in Belczynski et al. (2010). There is a region in the Hertzsprung-Russell diagram at low effective temperatures and high luminosities in which no stars are observed. The boundary of this region is called the Humphreys-Davidson limit (Humphreys & Davidson 1994). If a star enters this region we apply an additional wind mass loss rate of $1.5 \times 10^{-4}\ \mathrm{M}_\odot\ \mathrm{yr}^{-1}$ (Belczynski et al. 2010). From here onwards we refer to these winds as luminous blue variable (LBV) winds.

2.3 Mass Transfer Stability

The Roche lobe of a star defines the volume within which the self gravity of the star exceeds the tidal pull of its companion. We use the approximation of Eggleton (1983) for the Roche lobe radius. When a star expands, its radius may exceed its Roche lobe. At this moment, the star commences mass transfer onto the companion, Roche-lobe overflow (RLOF). If mass transfer results in the star further exceeding its Roche-lobe then the RLOF is unstable. We evaluate dynamical instability by comparing the radial response of the Roche-lobe to mass transfer $d\log(R_L)/d\log(m)$ against the response of the stellar radius to mass transfer $d\log(R_*)/d\log(m)$ (Paczyński & Sienkiewicz 1972; Hjellming & Webbink 1987; Soberman et al. 1997). We approximate the radial response of the star depending on its stellar type. The stellar types are defined in Hurley et al. (2000).

– main sequence (MS):
We use $d\log(R_*)/d\log(m) = 2.0$ for core hydrogen burning stars.

– Hertzsprung gap (HG):
We use $d\log(R_*)/d\log(m) = 6.5$ for so-called HG stars. Both MS and HG approximations follow our models in Vigna-Gómez et al. (2018), based on the work by Ge et al. (2015). More detailed models based on the evolutionary phase of the star and the amount of mass loss have been explored by Ge et al. (2015); Woods & Ivanova (2011); Pavlovskii et al. (2017a).

– Convective stars:
We use fits from Hjellming & Webbink (1987); Soberman et al. (1997) for the radial response to adiabatic mass loss of all evolved stars beyond HG. These fits are based on condensed polytropes for deeply convective stars and depend on the mass fraction of the core compared to the total mass of the star (Hjellming & Webbink 1987). We will investigate the applicability of these approximations in future work (Neijssel 2020).
– Stripped stars:
We make a special exception for mass transfer from exposed helium cores. We define this mass transfer to always be dynamically stable, yielding ultra-stripped stars based on Tauris et al. (2015, 2017). Vigna-Gómez et al. (2018) found that this assumption is necessary in order to recreate the observed Galactic double neutron stars in our models.

2.3.1 Stable Mass Transfer

If the mass transfer is dynamically stable, the companion star accretes a fraction β of the mass lost by the donor. In our model, this mass transfer efficiency β depends on the ratio of the thermal timescales $t_{\rm th}$ of the stars $\beta = \min\left(1, C \times t_{\rm th_1}/t_{\rm th_2}\right)$, where $0 \leq \beta \leq 1$, and $C = 10$ to allow for accretor radial expansion while adjusting to mass transfer (Paczyński & Sienkiewicz 1972; Hurley et al. 2002; Schneider et al. 2015). Any mass that is not accreted leaves the system instantaneously, taking away the specific angular momentum of the accretor (Hurley et al. 2002). For degenerate objects we assume the accretion is Eddington-limited, which results in a highly non-conservative mass transfer phase with $\beta \approx 0$.

2.3.2 Unstable Mass Transfer

If the mass transfer is unstable the envelope of the donor enfolds the entire binary in a common-envelope event (Paczynski 1976). This is a complex phase and we parametrise it in the so-called 'α–λ' formalism (see Ivanova et al. (2013) for a review). During a common-envelope event the two stars spiral in due to friction with the envelope and lose orbital energy and angular momentum. This loss of orbital energy can heat up and expel the envelope. To see if a binary is able to expel the common envelope, we compare the orbital energy against the binding energy of the envelope of the star (Webbink 1984). The efficiency α of converting orbital energy into heating up the envelope can vary (Livio & Soker 1988). We assume that all of the orbital energy goes into expelling the envelope (i.e. $\alpha = 1$). The binding energy of the envelope depends on the stellar structure of the star and is parametrised by λ (de Kool 1990). Our choices of λ are based on the binding energy fits by Xu & Li (2010) as implemented by Dominik et al. (2012).

Within the common envelope we define two scenarios for donor stars which are on the Hertzsprung-gap, following Belczynski et al. (2007). In the 'optimistic' scenario we evaluate the common-envelope evolution for Hertzsprung-gap stars using the 'α–λ' prescription. In the 'pessimistic' scenario we assume that unstable mass transfer from Hertzsprung-gap donors always results in a merger. The latter will therefore decrease the number of DCOs compared to the optimistic assumption. Common-envelope events with MS donors are assumed to lead to a prompt merger in all variations.

2.3.3 Supernovae

We use the 'delayed' model of Fryer et al. (2012) to determine the remnant mass from the pre-supernova (SN) mass of the star and its carbon-oxygen core. This model avoids an enforced mass gap between neutron stars (NSs) and black holes (BHs) (see also evidence that a mass gap is not consistent with microlensing observation unless BHs are assumed to receive substantial natal kicks (Wyrzykowski & Mandel 2019)). The explosion can be asymmetric and as a result impart a kick on the formed remnant. The kicks are drawn from a Maxwellian distribution with a one-dimensional standard deviation $\sigma = 265$ km s^{-1} based on the observations of isolated pulsars (Hobbs et al. 2005). If the progenitor either experiences an electron capture supernova or is ultra stripped by a NS companion, we lower the one-dimensional kick parameter to 30 km s^{-1} (Pfahl et al. 2002; Podsiadlowski et al. 2004; Tauris et al. 2015, 2017; Vigna-Gómez et al. 2018). The fraction f_b of mass that falls back onto the newly born compact object is prescribed by Fryer et al. (2012). All of the ejecta falls back (f_b=1) for carbon-oxygen core masses above 11 $M_\odot$. This natal kick is proportionally reduced based on the fallback fraction according to

$$V_{\rm kick} = (1 - f_b) V_{\rm kick,drawn}. \quad (4)$$

3 DCO POPULATION

In this section we describe the three main BBH formation channels. We focus on BBHs because they are the most common DCOs among already observed gravitational-wave events. More information on BNSs can be found in Vigna-Gómez et al. (2018) and the channels for black hole – neutron star binarys (BHNSs) are left for another study. We aslo show the metallicity, mass, mass-ratio, and delay time distributions for our model DCO population.

A 30 $M_\odot$ + 30 $M_\odot$ circular BBH needs a separation of $\lesssim 45\ R_\odot$ to merge in the age of the Universe, whereas the progenitor stars can expand up to hundreds of solar radii (Mandel & Farmer 2018). Therefore, progenitors of DCOs are expected to interact. This is not unlikely to happen for massive stars in binaries: observations show that most massive stars are likely to interact with a companion (Kiminki & Kobulnicky 2012; Sana et al. 2012). Only a small fraction of interacting massive binaries will form merging DCOs. This requires stars to avoid merger during mass transfer; to have sufficient mass to form compact objects; the binary must remain bound through SNe; and after the formation of a DCO, the binary must be tight enough to merge within the age of the Universe and create a detectable gravitational-wave event. Our main goal is to evaluate the DCOs that we can detect as gravitational-wave sources; hence we are only interested in the systems that merge within the age of the Universe. The results shown below assume the pessimistic common-envelope assumption. The optimistic assumption currently over-predicts the rates of BBHs (e.g. Dominik et al. 2012; Belczynski et al. 2016a); we show results using the optimistic assumption in Appendix C.

3.1 BBH Formation Channels

In our simulations, 97% of all the BBHs form through one of three distinct channels. Here we briefly summarise the evolutionary phases of the three main formation channels and the percentages of systems that remain after having experienced a given even/phase. The exact percentage and the ratio of the formation channels depends slightly on metallicity (see also figure 1). We focus on the systems that evolved at a metallicity of $Z = 0.1Z_{\odot}$,[1] and the percentage refers to the number of systems remaining divided by all of the systems evolved at this metallicity.

– Channel I –
This is the dominant, 'classical' channel of BBH formation as described in, e.g., van den Heuvel & De Loore (1973); Tutukov & Yungelson (1993); Lipunov et al. (1997); Belczynski et al. (2002a); Belczynski et al. (2016a); Stevenson et al. (2017).
-*Stable mass transfer*- The primary star expands sufficiently to engage in an episode of mass transfer (51.71%). The majority of these first mass transfer episodes will happen between a post-MS primary star and a MS companion (49.26%). The mass transfer is stable and strips the hydrogen envelope from the primary, leaving an exposed helium core with a main-sequence companion star (23.06%).
-*First supernova*- The exposed core is both massive enough to collapse into a BH and the binary survives the supernova (2.66%).
-*Unstable mass transfer*- The secondary star evolves and starts an episode of dynamically unstable mass transfer resulting in a common envelope (0.87%). The system is able to expel the envelope leaving a tighter binary (0.50%).
-*Second supernova*- The secondary also collapses into a BH and the binary system survives the second supernova (0.43%).
-*DCO merger*- The resulting BBH is then able to merge within the age of the Universe due to the emission of gravitational waves, which leaves 0.24% of all our evolved binaries merging as BBHs. Allowing for the optimistic common-envelope assumption, in which HG donors can survive a dynamically unstable mass transfer episode, increases the number of BBHs in this channel (0.39%).

– Channel II –
The second channel is similar to the 'classical' channel and goes through the same steps until the episode of mass transfer initiated by the secondary, which is dynamically stable in channel II. -*Stable mass transfer*- see channel I.
-*First supernova*- see channel I.
-*Stable mass transfer*- The secondary starts mass transfer as a post-MS star. The mass transfer is now dynamically stable and *does not result in a common-envelope phase* (1.35%).
-*Second supernova*- The secondary collapses into a BH without disrupting the binary (1.02%).
-*DCO merger*- Even without the common-envelope phase the BBH hardens (reduces orbital separation) sufficiently during the second mass transfer episode to spiral in and merge within the age of the Universe (0.15%).

Compared to Stevenson et al. (2017) we changed the radial response of HG donors to mass loss (see Sec. 2.3). In combination with our prescription for the angular momentum lost during non-conservative mass transfer onto a compact-object primary (see Sec. 2.3.1), the mass transfer is on average now stable for mass ratios up to $m_{\rm donor}/m_{\rm accretor} = 4.5$. Mass transfer from such donors that are significantly more massive than accretors can substantially harden the binary (van den Heuvel et al. 2017). With the increased stability these mass ratios are sufficiently extreme to allow the BBH to merge within the age of the Universe.

The stability of the second episode of mass transfer acts as a bifurcation point between channel I and channel II. Currently, channel II only happens for HG donors in our models, since we treat core-helium-burning donors as fully convective, making mass transfer from them less dynamically stable. The potential importance of this formation channel and the stability of mass transfer is discussed in previous studies (see for example Pavlovskii et al. 2017b; van den Heuvel et al. 2017, and references therein).

– Channel III –
The third channel for forming BBHs is similar to the double-core common-envelope channel introduced by (Brown 1995; Dewi et al. 2006). -*Unstable mass transfer*- . In this scenario both stars evolved beyond the HG before engaging in an episode of mass transfer (1.40%). This mass transfer is dynamically unstable (1.27%) and the binary survives the common-envelope ejection (0.71%). However, unlike the similar formation channel for BNSs (Vigna-Gómez et al. 2018), there is no further episode of mass transfer.
-*Two supernovae*- Both stars collapse in supernovae (non-simultaneously); 0.04% of binaries remain bound as a BBH.
-*DCO merger*- The DCO spirals in due to the emission of gravitational waves. In the end 0.03% of all binaries evolved go through this channel and merge within the age of the Universe.

The remaining three per cent of BBHs form through alternative channels. These include systems which have an additional moment of mass transfer after a common-envelope phase, or systems where the first moment of mass transfer is started by the secondary after the primary's supernova kick fortuitously tightened the binary.

3.2 Yield Per Metallicity

The yield of merging DCOs per unit star forming mass depends on the star formation metallicity, as shown in figure 1. As previously pointed out by Belczynski et al. (2010); Giacobbo et al. (2018); Spera et al. (2018), BBH yield is particularly sensitive to metallicity with a steep decline in BBH production at higher metallicities. Therefore, while BBHs are the dominant form of merging DCOs at sub-solar metallicities, they are more rare than BNSs and BHNSs at super-solar metallicities.

At higher metallicities, higher wind mass loss rates pre-

[1] In this study, we define the solar metallicity mass fraction as $Z_{\odot} = 0.0142$ and the solar oxygen abundance as $\log_{10}[{\rm O/H}]_{\odot} + 12 = 8.69$ based on Asplund et al. (2009); see appendix A5 for details.

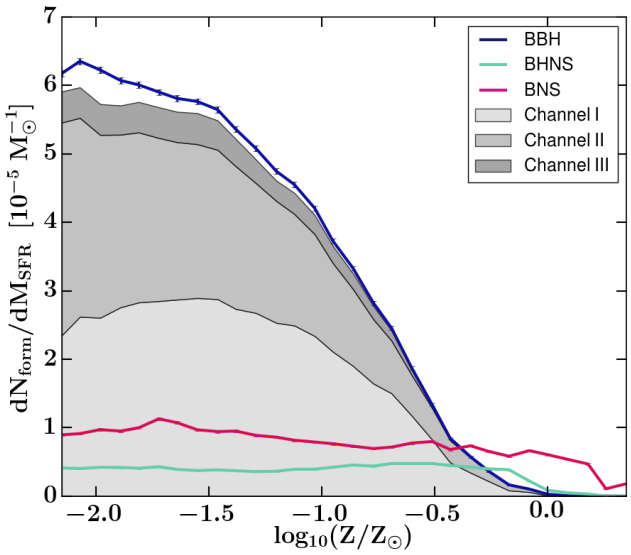

Figure 1. Yield of double compact objects with $t_{delay} < 14$ Gyr per unit star forming mass from COMPAS population synthesis. BBHs in blue, BHNSs in mint and BNSs in red. The curve under the BBH yield is shaded by the contribution of each channel (the white residual is due to rare alternative channels). The error bars show the sampling uncertainty of each simulation.

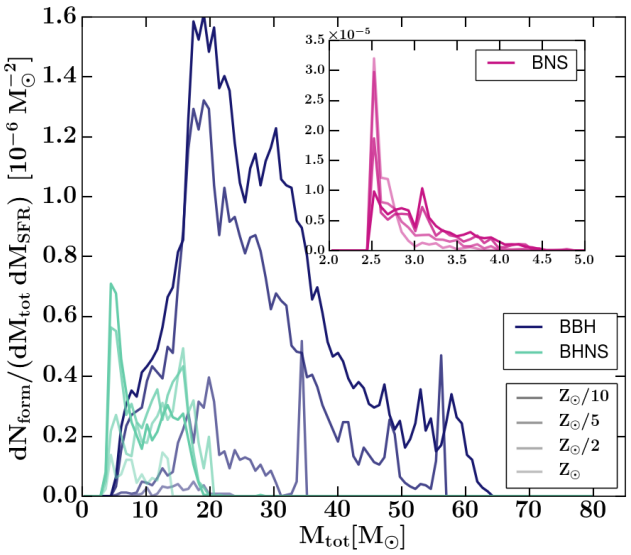

Figure 2. Total mass distributions for BBHs in blue, BHNSs in mint and BNSs in red from COMPAS simulations for a tenth, a fifth, a half, and solar metallicity (dark to light shade), for DCOs merging in $t_{delay} < 14$ Gyr. The integral under the curve is the yield plotted in Fig. 1. Higher metallicities yield lower total DCO masses, particularly for BBHs.

vent the growth of the carbon-oxygen core (Belczynski et al. 2010; Spera et al. 2015; Stevenson et al. 2017), leaving a less massive remnant. This affects the natal kicks imparted on the BHs. In the prescription of Fryer et al. (2012), stars with lower carbon-oxygen cores eject a larger fraction of their mass which results in larger natal kicks (see Eq. 4).

Therefore, we expect more potential BBH progenitors to be disrupted at higher metallicities. A smaller simulation without natal kicks does show a shallower drop-off of the BBH yield at higher metallicities. Nonetheless, there is still a lower yield at higher metallicities. This is largely due to the widening of binaries at higher metallicity, both directly through wind-driven mass loss and indirectly because reduced envelope masses limit the amount of orbital hardening during common-envelope ejection or stable mass transfer. In fact, if BH natal kicks are set to zero and all BBHs are accounted for, not just those merging in the age of the Universe, the BBH yield becomes almost independent of metallicity.

Lower-mass NS progenitors have lower mass loss rates, so the envelope mass is less sensitive to metallicity; moreover, their natal kicks are generally uncorrelated with metallicity. Hence it is not surprising that the yield of BNSs per unit solar mass evolved is less sensitive to metallicity, as also found by Giacobbo & Mapelli (2019).

3.3 Total Mass Distribution

Figure 2 shows the total mass distributions of DCOs merging within the age of the Universe for several metallicities. As discussed in the previous section, lower-metallicity stars with reduced wind-driven mass loss rates leave more massive remnants. For all metallicities the bulk of the BBH total masses lie between 15 & 35 $M_\odot$. More massive BBHs are suppressed by the IMF and wind-driven mass loss. The most massive binary black hole formed at a given metallicity is a function of both our assumptions about wind mass loss in massive stars, and our remnant prescription (these simulations do not include (pulsational) pair-instability supernova (PISN) – see Stevenson et al. 2019). Meanwhile, BHs with low masses get large kicks in the Fryer et al. (2012) prescription, and are therefore less likely to remain bound and form a BBH, explaining a dearth of BBHs with total mass below 15 $M_\odot$. The 'delayed' Fryer et al. (2012) remnant prescription does not enforce a mass gap between NSs and BHs, so we find some BBHs with total masses below 10 $M_\odot$ in our simulations, although these are relatively rare.

The presence of spikes in BBH masses, particularly in the highest mass bin at $Z = 0.5Z_\odot$ and $Z = 0.2Z_\odot$, are due to mass loss prescriptions, particularly LBV winds, that map a range of ZAMS masses to a single remnant mass (see Appendix. B). Similar features have been found in Dominik et al. (2015).

For BNSs we recover a similar total mass distribution as in Vigna-Gómez et al. (2018). As discussed in Vigna-Gómez et al. (2018), this distribution, driven by the Fryer et al. (2012) prescription, does not match the observed distribution of Galactic BNSs. For example, in our model, BNSs have total masses in the range 2.5–5.0 $M_\odot$, while observed Galactic BNSs with precise mass measurements have total masses in the narrower range 2.5–3.0 $M_\odot$ (Farrow et al. 2019).

3.4 Delay Times

The delay time is the time from the formation of the stars to their merger as a DCO. We follow Peters (1964) to estimate the time from DCO formation to merger through the emission of gravitational waves. The most massive binaries

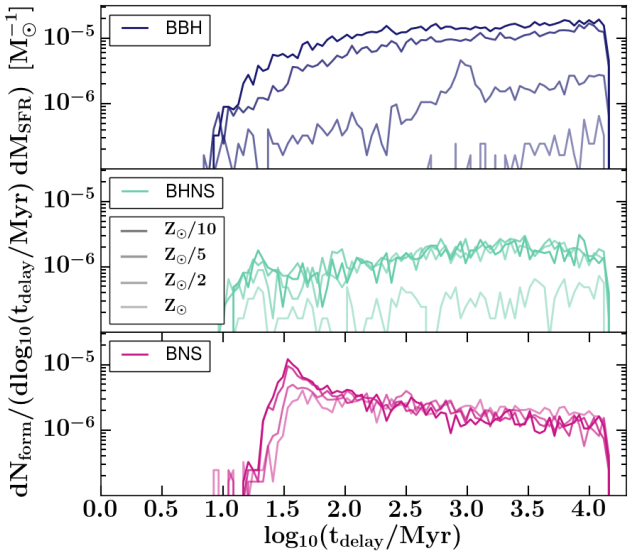

Figure 3. Delay time distributions up to $t_{\rm delay} = 14$ Gyr for BBHs in blue, BHNSs in mint and BNSs in red from COMPAS simulations for a tenth, a fifth, a half, and solar metallicity (dark to light shade).

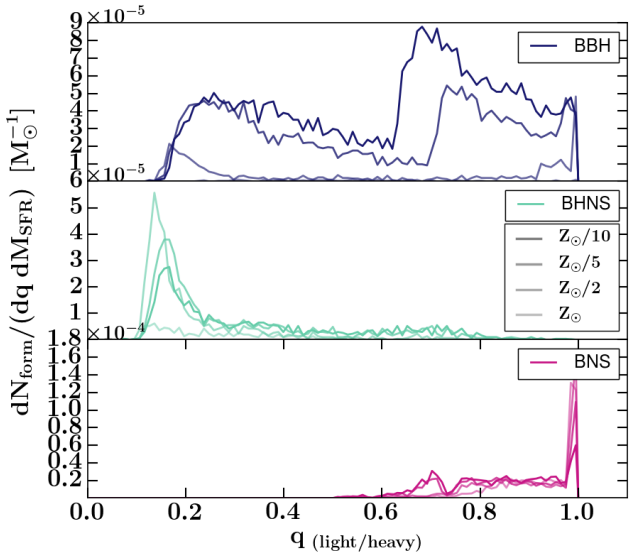

Figure 4. Mass ratio distributions for BBHs (blue), BHNSs (mint) and BNSs (red) merging in $t_{\rm delay} < 14$ Gyr from COMPAS simulations for a tenth, a fifth, a half, and solar metallicity (dark to light shade). For BBHs channel II contributes more at lower metallicities and therefore the mass ratio distribution at $0.1Z_\odot$ has a prominent feature around $q \approx 0.65$.

with the smallest separation at the formation of the DCO have the shortest inspiral times. The delay time distribution is roughly flat-in-the-log for all DCOs (see Fig. 3). Furthermore, for the pessimistic assumption it is not very sensitive to metallicity. These findings are similar to Dominik et al. (2012) and Mapelli et al. (2017).

3.5 Mass Ratios

Figure 4 shows the mass ratio distributions of DCOs merging within the age of the Universe at several metallicities. It is clear that the distributions differ between different types of DCOs and depend on metallicity.

The mass ratios of BNSs exhibit two dominant peaks. The mass ratio is close to unity if both NSs have the lowest allowed remnant mass from iron-core-collapse supernovae in the Fryer et al. (2012) 'delayed' prescription. The other peak is the ratio between this lowest iron-core-collapse remnant mass and the fixed remnant mass from electron-capture supernovae, 1.26 $M_\odot$ in our model. The prevalence of these peaks is enhanced by binary interactions. We do not expect extreme mass ratios given the limited spread in possible NS masses.

The BHNSs favour more extreme mass ratios. The average NS mass is 1.2 $M_\odot$ and the threshold between NS and BH is $2.5M_\odot$ in our models. This already results in a mass ratio of 0.5, but most of the BHs are heavier. Further details are outside the scope of this study.

The mass ratio distribution of BBHs depends on the formation channel. The classical channel I with a common-envelope phase occurs for a broad range of mass ratios between the donor star and the accreting BH. This channel yields a relatively flat mass-ratio distribution. Meanwhile, channel II, in which the mass transfer onto the BH is dynamically stable, has an upper limit of 4.5 for the mass ratio between the donor and the BH accretor. Mass ratios close to this limit are preferred as they provide the most orbital hardening. After this mass transfer, the stripped donor star collapses into a BH. This results in a BBH mass ratio around $q \approx 0.6$. If such an additional peak is observed in the mass ratio distribution of gravitational-wave events, its prominence and location could put a constrain on the ratio of formation channels, and, hence, the stability of mass transfer.

4 METALLICITY SPECIFIC STAR FORMATION RATE

We divide the calculation of the MSSFR into two independent factors, the SFR and the metallicity distribution:

$$\frac{d^3 M_{\rm SFR}}{dt_s dV_c dZ}(z) = \frac{d^2 M_{\rm SFR}}{dt_s dV_c}(z) \times \frac{dP}{dZ}(z). \quad (5)$$

In practice, the SFR and the metallicity distribution may be correlated (see for example Furlong et al. 2015); however, decoupling the SFR and the metallicity distribution is a convenient simplifying assumption that yields sufficient degrees of freedom given current observational constraints.

We discuss detailed models of the SFR and metallicity distribution in Appendix A. Here, we summarise the key approach to justify the shape of a phenomenological model that can be used for future inference. We highlight a particular choice of the model parameters that, coupled with our default binary evolution model, produce a good match to data from the first two observing runs of the advanced detector network (see section 6).

Figure 5 illustrates the SFR models we use. All models agree well at low redshift, $z \lesssim 2$, other than differences in calibration due to the assumed initial mass function (Madau

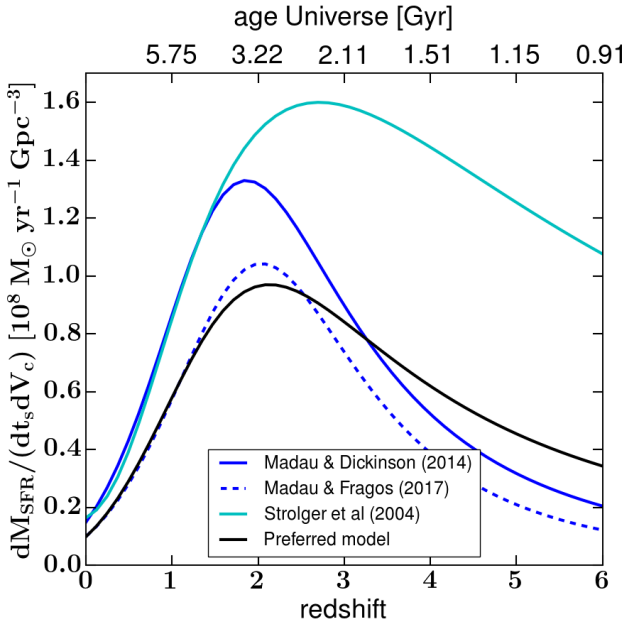

Figure 5. The star formation rate: mass per unit comoving volume per unit time as a function of redshift. The distributions shapes are similar up to a redshift of 1.5. The star formation rate estimate of Madau & Dickinson (2014) peaks slightly earlier, at a redshift of two, and then decreases steeply. Strolger et al. (2004) used an additional extinction correction and recover a higher star formation rate at higher redshifts. Madau & Fragos (2017) assume a slightly different IMF resulting in a lower normalisation. We also include our preferred phenomenological model for the star formation rate history.

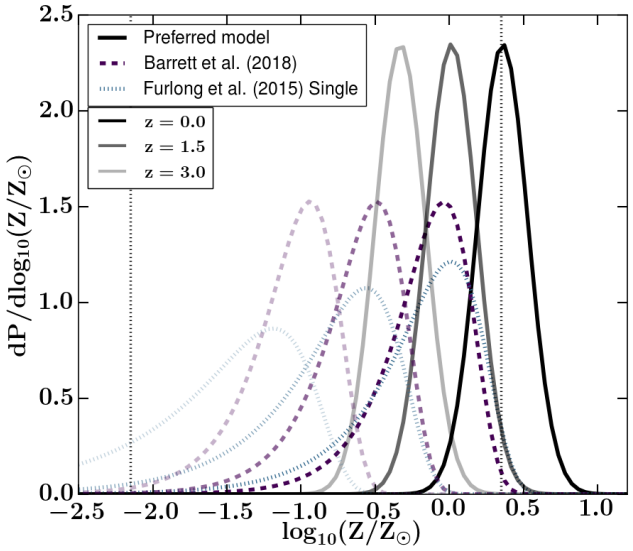

Figure 6. The star formation metallicity distribution. The shades (dark to light) denote the redshifts 0, 1.5, and 3. Our previous model of Barrett et al. (2018) convolves the MZ-relation of Langer & Norman (2006) with a redshift-independent GSMF of Panter et al. (2004) (purple dashed line). The blue dotted line instead uses the GSMF by Furlong et al. (2015). We also include our preferred model for the metallicity distribution of star formation (black solid). The vertical dotted lines denote the limits of our metallicity grid; portions of the distribution extending beyond these limits are included in the edge bins when integrating over metallicity.

& Fragos 2017). At higher redshift, Strolger et al. (2004) assume greater extinction and find a higher rate of star formation than Madau & Dickinson (2014). We follow the functional form of Madau & Dickinson (2014) in our phenomenological model:

$$\frac{d^2 M_{\rm SFR}}{dt_s dV_c} = a\frac{(1+z)^b}{1+[(1+z)/c]^d} \ \ {\rm M}_\odot \ {\rm year}^{-1} \ {\rm Mpc}^{-3}. \quad (6)$$

The entire parameter space would be 4-dimensional, but we find that all of these SFR prescriptions can be reasonably reproduced by setting $b = 2.77$, c=2.9 and letting a, d vary in the intervals [0.01–0.015] and [3.6–5.6], respectively. In section 6, we show that $a = 0.01$ and $d = 4.7$ yield a good match to gravitational-wave observations when coupled with the metallicity distribution model discussed below and our default binary evolution model.

The metallicity density function at each redshift is typically obtained by convolving a GSMF with a MZ relation. Both of these are subject to significant uncertainties, and we describe several GSMF fits (Panter et al. 2004; Furlong et al. 2015) and MZ relations (Savaglio et al. 2005; Langer & Norman 2006; Ma et al. 2015) in Appendix A. We show the metallicity distribution at several redshifts from a combination of some of these predictions in figure 6. This figure also shows our fiducial model – a log-normal distribution in metallicity

$$\frac{dP}{dZ}(z) = \frac{1}{Z\sigma\sqrt{2\pi}} e^{-\frac{(\ln(Z)-\mu(z))^2}{2\sigma^2}}, \quad (7)$$

with redshift-independent standard deviation σ in $\ln(Z)$ space around a redshift-dependent mean μ of $\ln(Z)$ given by

$$\langle Z \rangle = e^{(\mu+\frac{\sigma^2}{2})}. \quad (8)$$

We follow Langer & Norman (2006) in parametrising mean metallicity as

$$\langle Z(z) \rangle = Z_0 10^{\alpha z}, \quad (9)$$

where Z_0 is the mean metallicity at $z = 0$ and the parameter α has negative values, yielding lower mean metallicity at higher redshifts. Therefore the free parameters for the metallicity distribution are Z_0, α and σ. We show in section 6 that $Z_0 = 0.035$, $\alpha = -0.23$ and $\sigma = 0.39$ yield a good match to gravitational-wave observations when coupled with our other assumptions; this preferred model has a similar shape to the metallicity distribution inferred by Rafelski et al. (2012) from measurements of damped Lyman α galaxies.

5 DCO MERGERS

In this section we focus on the rate and mass distribution of DCO mergers as a function of redshift. We convolve the DCO population formed at each redshift (section 3) with the MSSFR (section 4), incorporating the delay time distribution according to equation 1. We do not yet take into account any selection effects. We find that the choice of MSSFR

affects the total merger rate as a function of redshift, the relative rate between different types of DCO, and the mass distribution. Additionally we show that our predicted distributions do not match the priors used by Abbott et al. (2016); Abbott et al. (2018b,a) for inference on gravitational-wave signals.

5.1 Rate and Redshift of Cosmic DCOs Mergers

Figure 7 shows the intrinsic rate of DCO mergers as a function of redshift for a few MSSFR combinations. In our preferred MSSFR model, the merger rate at redshift $z = 0$ is 49, 57, 20 Gpc^{-3} yr^{-1} for BBHs, BNSs, and BHNSs, respectively. These rates are lower than the other models considered in section 4, including our previous MSSFR model (Barrett et al. 2018). The main reason is that our new preferred model favours higher metallicities at low redshifts (see figure 6). This suppresses the yield of BBHs and shifts their peak merger rate to higher redshifts which is in principle measurable with future gravitational-wave observations (Fishbach et al. 2018; Vitale & Farr 2018). This is consistent with results from other MSSFR models: the variation with the Ma et al. (2015) MZ relation, which has a higher average metallicity than the Langer & Norman (2006) MZ relation used in Barrett et al. (2018), and consequently yields lower DCOs merger rates; while the Furlong et al. (2015) redshift-dependent GSMF, which allows for more low-mass galaxies than the Panter et al. (2004) GSMF assumed in Barrett et al. (2018), with correspondingly lower metallicities, yields higher DCO merger rates.

The effect of different metallicity distributions is smaller for the rates of BNSs and BHNSs, since their yield is less metallicity-dependent. We note that the change in the MSSFR affects not only the overall DCO merger rate, but also the ratio between different merger rates of different DCO types.

5.2 Mass Distribution and Redshift of Cosmic DCO Mergers

Figure 8 shows the normalised total mass distribution of BBH mergers at several redshifts for our preferred model MSSFR model. This is due to the convolution of the redshift dependence of the MSSFR with the delay time distribution. There is a significant contribution to low-redshift mergers from DCOs that formed at low metallicity and high redshift, with long delay times (see Fig.3). These low-metallicity systems give rise to high-mass BBH mergers (see Fig. 2). In fact, there is a greater tail of high-mass DCOs merging at redshift $z = 0$ than at higher redshifts in figure 8 (see also Dominik et al. (2015); Belczynski et al. (2016a)).

The mass distribution is sensitive to the metallicity of formation, and therefore depends on the assumed MSSFR prescription. We show the impact of the MSSFR on the mass distribution of DCOs merging at redshift $z = 0$ in figure 9. As with the BBH merger rate discussed in section 5.1, MSSFR models with lower metallicity (our previous model in Barrett et al. (2018), especially with the Furlong et al. (2015) GSMF variation) show enhanced high-mass tails relative to MSSFR models with higher metallicity (our preferred model)or reduced high-redshift, low-metallicity SFR

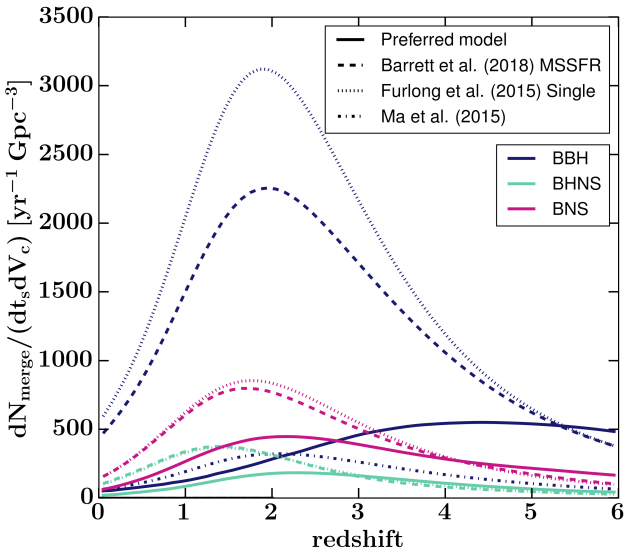

Figure 7. The intrinsic rate of DCO mergers per cubic Gpc per year. The colours denote different DCO types: BBHs in dark blue, BHNSs in mint, and BNSs in pink. The solid line is our preferred phenomenological model. The dashed line is the default model of Barrett et al. (2018), which combines the SFR of Madau & Dickinson (2014), the MZ-relation of Langer & Norman (2006), and the redshift-independent GSMF of Panter et al. (2004). The dotted line replaces the latter with the redshift-dependent single Schechter GSMF of Furlong et al. (2015). For clarity we only show the BBH distribution for the MZ-relation by Ma et al. (2015) (dot-dashed).

(the Ma et al. (2015) MZ relation combined with the SFR of Madau & Dickinson (2014), the redshift-independent GSMF of Panter et al. (2004)). The peaks at high masses in figures 8 and 9 are due to mass loss prescriptions, particularly LBV winds (for more details see Appendix B). These depend on metallicity, hence the prominence of the peaks varies depending on the MSSFR prescription.

5.3 Priors and Rate Estimates

The DCO merger rate inferred from gravitational-wave observations is sensitive to the assumed mass distribution (Abbott et al. 2018b,a). We show the mass priors assumed by Abbott et al. (2016) and Abbott et al. (2018b) in figure 8; it is clear that these are inconsistent with our predicted mass distribution. Abbott et al. (2018b) account for uncertainties in the shape of the BBH mass distribution by varying the slope of a power law distribution. However, as we show in Fig. 8, the mass distribution of BBHs might be more complex than a simple power law, and is furthermore a function of redshift, along with the merger rate itself. Therefore, DCO merger rates and mass distributions inferred from simple priors or phenomenological models should be treated with caution.

The complex dependence of the mass distribution of merging DCOs on both the binary evolution model (e.g. Dominik et al. 2013; Mapelli et al. 2017; Stevenson et al. 2017) and the MSSFR (this study), and the variation in the mass distribution and merger rate with redshift, makes

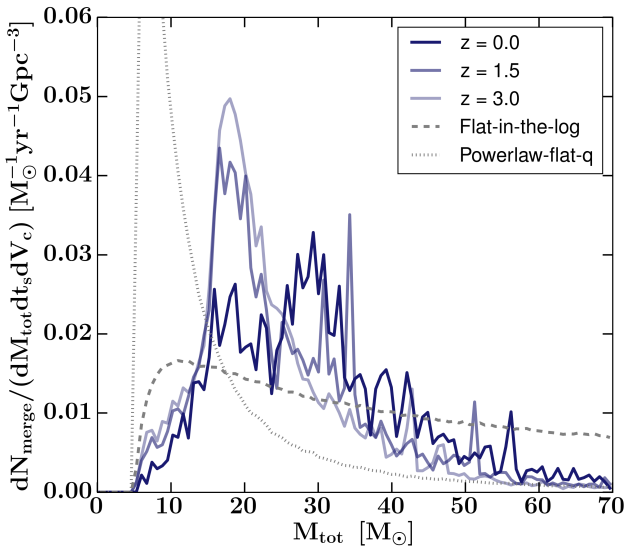

Figure 8. The normalised total mass distribution of BBH mergers at redshifts 0, 1.5, and 3 (shaded dark to light) for our preferred log model. The narrow significant spikes above 30 $M_\odot$ relate to the LBV systems (for more details see appendix B). For comparison, the dotted curve indicates the total mass distribution assuming that the more massive BH is sampled from a power law with index of -2.3 paired with a companion drawn from a flat mass ratio distribution. The dashed curve is a total mass distribution where both BH masses are sampled from a flat-in-the-log distribution. For the minimum mass we took 2.5 $M_\odot$ given that we have not introduced a mass gap. For the maximum mass we took 100 $M_\odot$. In our simulations such high BH masses are rare, but possible. Similar priors have been used in rate estimate studies such as Abbott et al. (2016) and Abbott et al. (2018b).

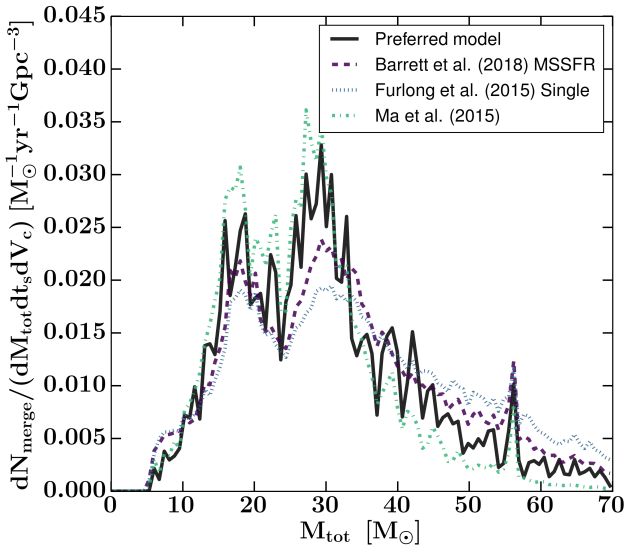

Figure 9. The normalised total mass distribution of BBH mergers at redshift $z = 0$. The MSSFR models are the same as in figure 7.

it challenging to propose alternative priors. Therefore, it is preferable to apply selection effects to the model population in order to compare model predictions against observations. This is the approach we take in the next section.

6 GRAVITATIONAL-WAVE DETECTIONS

This section focuses on the effect of the MSSFR on the predicted rates and mass distributions of detectable DCO mergers. We evaluate these using Eq. 2. We predict the total rate of detectable DCO mergers as a function of redshift and describe the mass distribution of BBH mergers. We carry out a Bayesian model comparison of different MSSFR prescriptions, taking into account both the number and the mass estimates of the 10 BBH mergers detected during the first and second observing runs of aLIGO (Abbott et al. 2016; Abbott et al. 2018a). We do not include in our analysis the 6 additional BBH candidates found in the same data set by Venumadhav et al. (2019) with an independent search pipeline and somewhat different data quality choices.

6.1 Selection Effects

For the selection effects we use the same method as described in Barrett et al. (2018). We use a single detector signal to noise ratio (SNR) threshold of 8 (Aasi et al. 2016), above which we assume that gravitational waves from the merger are detectable. To evaluate the SNR for a given DCO system, we compute the waveforms for the appropriate masses using a combination of IMRPhenomPv2 (Hannam et al. 2014; Husa et al. 2016; Khan et al. 2016) and SEOBNRv3 (Pan et al. 2014; Babak et al. 2017). We approximate the sensitivity of the second observing run (Abbott et al. 2018a) to be similar to the first observing run (Abbott et al. 2016). The fraction of systems with SNR above the threshold of 8 at a given distance (redshift), after sampling over the sky location and orientation of the binary (Finn & Chernoff 1993), yields the detection probability $P_{\rm det}(M_{\rm chirp}, D_L)$.

6.2 Rate and Redshift of Gravitational-Wave Detections

The rate of detectable DCO mergers depends on the underlying merger rate, which increases up to redshift $z \sim 2$ (Fig. 7). However, the detection probability drops off at higher redshift. These competing effects mean that the detection rate of BBH mergers this results in a peak rate at a redshift between 0.1–0.15 depending on the MSSFR model at the sensitivity of the first two observing runs. This is shown in Fig. 10. Note that this figure displays the number of detections per unit redshift per unit observer time, rather than per unit volume per unit source time as in Fig. 7 (see Eq. 2 for the additional factors of $dV_c/dz/(1+z)$). Because mergers involving less massive NSs cannot be observed as far as BBH mergers, detection rates of BHNS and BNS mergers per unit redshift peak at $z \approx 0.03$ and $z \approx 0.015$, respectively. As discussed in Sec. 5.1, sensitivity of the detection rate to MSSFR variations tracks the sensitivity of the DCO formation rate to metallicity (see Fig. 1).

Table 1 shows the observed rate per DCO type per

SFR	Variation MSSFR MZ	GSMF	Detection rate [yr^{-1}] BBH	BHNS	BNS	Likelihoods (log_{10}) $\mathcal{L}_{M_{chirp}}$	$\mathcal{L}_R$	$\mathcal{L}_{tot}$
	Preferred model		22.15	0.23	0.08	-32.32	-0.90	-33.22
Madau et al.	Ma et al. (2004)	1	18.43	0.4	0.11	-33.9	-0.97	-34.87
		2	94.35	0.51	0.12	-32.42	-8.86	-41.28
		3	113.92	0.52	0.13	-32.48	-11.9	-44.38
	Langer et al.	1	247.22	1.28	0.22	-32.24	-34.85	-67.09
		2	441.08	1.19	0.22	-32.61	-70.6	-103.21
		3	492.27	1.25	0.23	-32.77	-80.23	-113.0
	Langer et al., offset	1	28.72	0.23	0.09	-32.3	-1.07	-33.38
		2	120.3	0.35	0.11	-32.68	-12.93	-45.61
		3	148.74	0.35	0.11	-32.87	-17.62	-50.49
Strolger et al.	Ma et al. (2004)	1	32.93	0.52	0.12	-33.82	-1.31	-35.13
		2	203.93	0.6	0.14	-32.81	-27.14	-59.95
		3	208.21	0.61	0.14	-32.65	-27.9	-60.54
	Langer et al.	1	406.39	1.28	0.23	-32.44	-64.11	-96.55
		2	659.25	1.19	0.24	-32.98	-111.92	-144.9
		3	710.91	1.25	0.24	-33.09	-121.79	-154.87
	Langer et al., offset	1	89.79	0.33	0.11	-32.46	-8.18	-40.63
		2	267.34	0.43	0.12	-33.2	-38.48	-71.68
		3	292.76	0.43	0.12	-33.22	-43.1	-76.33

Table 1. Rate estimates and likelihoods per MSSFR variation. The numbers in the column GSMF refer to 1=Panter et al. (2004), 2=Furlong et al. (2015) (single Schechter function), 3=Furlong et al. (2015) (double Schechter function). The detection rates are estimated for a year of coincident observing with the sensitivity of the first observing run of aLIGO. The likelihoods account for BBH detections during the first and second observing runs, assuming the same sensitivity (Abbott et al. 2016; Abbott et al. 2018a). The total log likelihood $\mathcal{L}_{tot}$ is the sum of the log likelihoods of the chirp-mass distribution $\mathcal{L}_{M_{chirp}}$ and the rate $\mathcal{L}_R$.

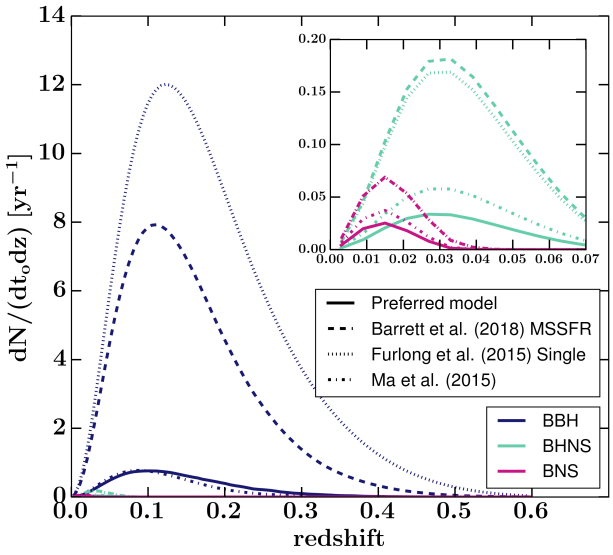

Figure 10. The number of DCO mergers per year per unit redshift at the sensitivity of the first two advanced detector observing runs. The top-right panel is an enlargement with the same axes to focus in on the merger rates for BHNSs and BNSs. The MSSFR models are the same as in Fig. 7.

year. The combined observing time of the first two observing runs is about 166 days: 48 days of coincident data for the first and 118 days for the second observing run (Abbott et al. 2016; Abbott et al. 2018a). Thus, 10 detections translate to an observed detection rate of 22 BBH mergers per year. Most of our variations significantly overestimate the observed rate. As previously mentioned, variations that favour a higher SFR (Strolger et al. 2004), lower metallicities (Langer & Norman 2006) or lower galaxy stellar masses (Furlong et al. 2015) predict a higher detection rate. We find that by changing the MSSFR alone we can vary the predicted rate of detectable BBH mergers by more than an order of magnitude.

All of the predictions for detectable BNS mergers are lower than of one in four years of observing time, suggesting that GW170817 was a fortuitous event. MSSFR models with the highest rates predict more than one detectable BHNS merger in one year observing time, however, these are generally inconsistent with observations in their BBH merger rate predictions.

6.3 Mass Distribution of Detectable BBH Mergers

The top panel of Fig. 11 shows the predicted chirp-mass distributions of detectable BBH mergers for several MSSFR variations. We use the chirp masses of the BBHs mergers here since these are typically better observationally constrained than the total masses.

Mergers of more massive DCOs emit louder gravitational-wave signal that can be detected to greater distances. Therefore, the mass distribution of detectable BBHs is shifted to higher masses relative to the intrinsic mass distribution of Fig. 9. The impact of MSSFR variations on the shape of the distribution follows the discussion in Sec. 5.2. However, the selection effects emphasise the peak due to the LBV winds at chirp masses around 25 $M_\odot$ (see also appendix B). Although the 'delayed' remnant mass model of Fryer et al. (2012) used in our simulations does

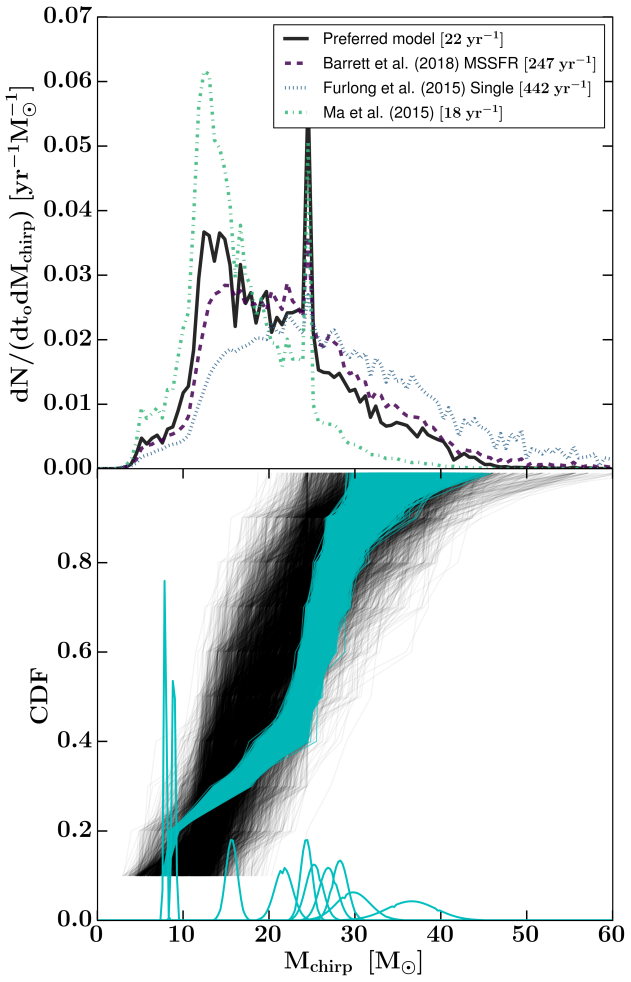

Figure 11. Top panel: Predicted chirp-mass distribution of BBH mergers detectable at the sensitivity of the first two observing runs. The masses are in the source reference frame. The MSSFR models are the same as in Fig. 7. The numbers in the label are the predicted annual detection rate. Bottom panel: Approximate Gaussian posteriors (see appendix C) for BBH mergers detected during the first and second observing runs (Abbott et al. 2018a), from left to right: GW170608, GW151226, GW151012, GW170104, GW170814, GW170809, GW170818, GW150914, GW170823, GW170729. The cyan area shows randomly sampled cumulative density functions from the posteriors, indicating the spread due to the measurement uncertainty. The black lines are cumulative density functions when 10 events are randomly drawn from the preferred model.

not enforce a mass gap between NS and BH masses, we find that low-mass merging BBHs are very rare, especially after selection effects are applied. We do not expect significant numbers of detections in the mass gap for any of the MSSFR variations.

The bottom panel of Fig. 11 shows that our preferred MSSFR model predicts a chirp mass distribution of detectable BBH mergers that is consistent with the detection from the first two advanced detector observing runs. Approximate Gaussian posteriors (see appendix C) for the ten detections are shown in cyan at the bottom of the plot. We construct observed CDFs by taking a random sample from each of these ten posteriors. The set of cyan curves indicates the range of observed CDFs consistent with measurement uncertainty. Meanwhile, each black curve represents a CDF constructed by sampling from the predicted distribution of detectable BBH events under the preferred MSSFR model. The visual consistency between the black and cyan regions indicates a successful graphical predictive check of the model.

6.4 Bayesian Comparison of MSSFR Models

We showed that the choice of the MSSFR affects both the detectable rates and mass distributions of DCO mergers. Here we quantitatively compare these models against observations during the first and second observing run (Abbott et al. 2018a). We consider the total rate of events and the relatively well-measured chirp masses. We do not consider other properties such as relatively poorly measured mass ratios or source redshifts given the narrow range of redshifts reached to date. Bavera et al. (2019) compare a possible model for BBH spins evolving through channel I (see section 3.1) using the COMPAS data presented here against observations. In this analysis (as in Barrett et al. (2018)), the total log likelihood $\mathcal{L}_{\rm tot}$ is the sum of the rate log likelihood $\mathcal{L}_{\rm R}$ and the likelihood of the normalised chirp-mass distribution $\mathcal{L}_{\rm M_{chirp}}$:

$$\mathcal{L}_{\rm tot} = \mathcal{L}_{\rm M_{chirp}} + \mathcal{L}_{\rm R}. \quad (10)$$

The rate likelihood assumes a Poisson distribution where the MSSFR model gives the expected number of detections over the duration of the first two observing runs. The chirp-mass likelihood is the product over the ten events of the probabilities of making individual detections given the predicted chrip-mass distribution (see appendix C). A difference of 1 in log likelihoods, corresponding to a factor of 10 in the likelihoods, implies that the higher-likelihood model is preferred over the lower-likelihood model by a factor of 10 (i.e., has an odds ratio of 10 : 1, assuming both models are equally probable a priori). Table 1 shows the total likelihoods for the pessimistic common-envelope assumptions. A longer list of variations, including the optimistic common-envelope assumption, can be found in tables C2, C1.

The rate likelihoods differ significantly given our range in rate estimates. Many of the MSSFR models greatly overestimate the rates and are strongly disfavoured under the assumed model of binary evolution. Meanwhile, despite the visual difference in the shape of the chirp-mass distribution (see Fig. 11), the difference in the chirp-mass likelihoods is small. More detections will make it possible to jointly explore MSSFR and evolutionary models using the observed chirp-mass distributions (Barrett et al. 2018). Given our binary evolution model, higher star-formation metallicities at low redshifts are preferred to match the observed BBH rate and chirp-mass distribution.

In section 4 we introduced a 5-parameter phenomenological model of the MSSFR. With suitable parameter choices this generic model can match all of the detailed models considered here, while providing the convenience of a continuous, smooth parametrisation that is useful for inference. We also introduced a particular choice of these 5 parameters – our preferred model – that yields a good match to both the number of BBH mergers detected during the first two

observing runs (10.06 predicted vs. 10 observed) and their chirp mass distribution (section 6.3 and the bottom panel of Fig. 11). As table 1 shows, this preferred model also yields the highest likelihood among all considered models. This preferred model favours a SFR similar to Madau & Fragos (2017), which includes the contribution from stars in binaries. However, we do favour a higher SFR at high redshifts, where metallicity is lower, to enhance the fraction of massive BBH merger events. We caution, however, that the MSSFR parameters in the preferred model are chosen ad hoc, with some 'Fingerspitzengefühl'. Future analyses should jointly infer the parameters of the MSSFR and parameters describing the binary evolution model, using gravitational waves and other observational constraints.

7 CONCLUSION AND DISCUSSION

We showed that assuming different MSSFR within observational constraints can vary the rate of BBH mergers by more than an order of magnitude within a fixed stellar and binary evolution model[2] and affect the ratio between BBH and BNS detection rates. This is comparable to the impact of uncertainties on evolutionary physics such as wind mass loss rates, conservativeness of mass transfer, the efficiency of common envelope evolution and BH natal kicks (Dominik et al. 2012; Kruckow et al. 2018; Giacobbo & Mapelli 2018).

The sensitivity to MSSFR is predominantly driven by the impact of metallicity on the yield of BBHs per unit star forming mass. This is consistent with earlier findings (e.g., Dominik et al. 2015; Chruslinska et al. 2018). In particular, Chruslinska et al. (2018) also find that a higher average metallicity is required in order to not over-predict the BBH merger rate.

Here, we explored the impact of the MSSFR while keeping the binary evolution model unchanged. In practice, joint inference on stellar and binary physics and the MSSFR is required to fully interpret observations (e.g., Chruslinska et al. 2018). For example, (pulsational) PISNe (see for example Woosley (2017) and references therein) can prevent the formation of BHs with masses between around 50 and 130 $M_\odot$, i.e., with chirp masses between 45 and 115 $M_\odot$ for equal-mass binaries. Abbott et al. (2018b) find that existing gravitational-wave detections show evidence for a maximum black hole mass of around $\sim 45\,M_\odot$, consistent with population synthesis studies such as Belczynski et al. (2016b); Spera & Mapelli (2017); Stevenson et al. (2019). However in Fig. 11, we show that it is possible to reproduce such a limit within the evolutionary model of this paper, which does not include pulsation PISNe, by choosing a suitable MSSFR alone (Madau & Dickinson (2014); Ma et al. (2015); Langer & Norman (2006)). A similar argument can be made for the presence or absence of a mass gap between NSs and BHs. Beyond gravitational-wave observations, other observational constraints such as the epoch of reionisation (Stanway et al. 2016) and X-ray binaries (Madau & Fragos 2017) can further help to lift the degeneracy between binary physics and the MSSFR.

[2] There is a further uncertainty from the definition of solar metallicity, see appendix A5.

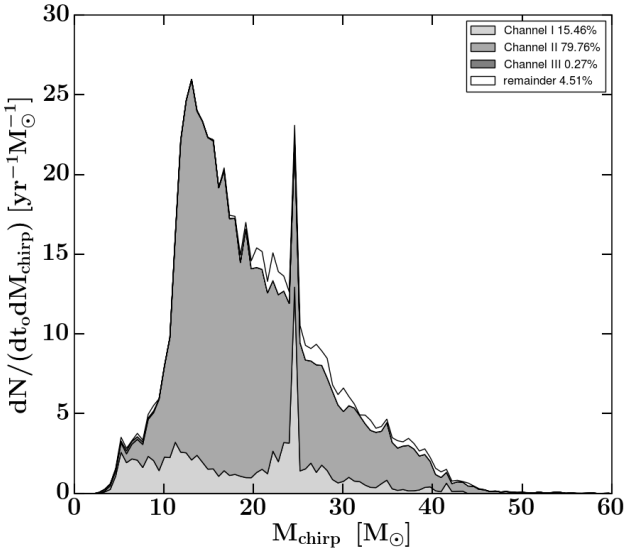

Figure 12. The predicted chirp mass distribution of detectable BBH mergers at advanced detector design sensitivity within our preferred MSSFR model, coloured by the BBH formation channels. Almost 80% of the expected 380 detectable BBHs per year are formed through channel II (dynamically stable mass transfer) within our evolutionary model.

We introduced a phenomenological description of the MSSFR with 5 continuous parameters (section 4) to facilitate the joint exploration of the MSSFR and parametrised evolutionary assumptions. We also proposed a particular choice of the MSSFR model parameters that represents a good match to the gravitational-wave detections made during the first two observing runs of advanced LIGO and Virgo. Looking ahead, we can apply this preferred MSSFR model to make predictions for the detection rate and chirp mass distribution at design sensitivity, shown in figure 12. We predict 380 detections BBH detections per year, or approximately one detection per day, within our default evolutionary model.

Figure 12 also highlights the importance of the dynamically stable mass transfer channel without a common envelope phase for the formation of detectable merging BBHs. We find that channel II may be responsible for 80% of all detected BBH mergers. This highlights the importance of mass transfer stability criteria, which merit further investigation. Meanwhile, the narrow chirp mass spike at around 25 $M_\odot$ is due to the operation of LBV mass loss at a particular metallicity (cf. Dominik et al. 2015). While we expect that a finer metallicity grid or interpolation between metallicities would lead to a smoother chirp mass distribution, this again highlights the importance of highly uncertain LBV winds for these predictions (Mennekens & Vanbeveren 2014). Finally, the sampling accuracy of predictions (e.g., the time delay distribution for BHNS in Figure 3) could be improved with more efficient importance sampling techniques (Broekgaarden et al. 2019).

Our predictions suggest that approximately one thousand detections could be reached within a couple of years of operation of advanced detectors operating at design sensitivity. Barrett et al. (2018) showed that this will be sufficient to

constrain the binary evolutionary parameters to a fractional accuracy of a few percent. Our phenomenological MSSFR model can be incorporated into this hierarchical modelling framework to be enable joint inference on binary evolution and the cosmic history of star formation.

ACKNOWLEDGEMENTS

C. J. N. thanks the University of Birmingham for financial support, C. P. L. Berry, S. Bavera and A. Vecchio for helpful discussions, D. J. Stops for technical support, and of course above all T. F. Pauw for everything. A. V. G. acknowledges support from Consejo Nacional de Ciencia y Tecnologia (CONACYT). S. S. is supported by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (OzGrav), through project number CE170100004. S. M. is supported by STFC grant ST/M004090/1. D. Sz. accepts funding from the Alexander von Humboldt Foundation. This paper used the `Astropy` library (Astropy Collaboration et al. 2013, 2018), matplotlib (Hunter 2007), and numpy (Oliphant 2006).

REFERENCES

Aasi J., et al., 2016, Living Reviews in Relativity, 19
Abbott B., et al., 2016, Physical Review X, 6, 041015
Abbott B. P., et al., 2017, Physical Review Letters, 119, 161101
Abbott B. P., et al., 2018b, preprint, (`arXiv:1811.12940`)
Abbott B. P., et al., 2018a, preprint, (`arXiv:1811.12907`)
Ade P. A., et al., 2016, Astronomy & Astrophysics, 594, A13
Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Astropy Collaboration et al., 2013, A&A, 558, A33
Astropy Collaboration et al., 2018, AJ, 156, 123
Babak S., Taracchini A., Buonanno A., 2017, Phys. Rev. D, 95, 024010
Baldry I., Glazebrook K., Driver S., 2008, Monthly Notices of the Royal Astronomical Society, 388, 945
Barrett J. W., Gaebel S. M., Neijssel C. J., Vigna-Gómez A., Stevenson S., Berry C. P. L., Farr W. M., Mandel I., 2018, MNRAS, 477, 4685
Bavera S. S., et al., 2019, in prep
Belczynski K., Bulik T., Kalogera V., 2002a, ApJ, 571, L147
Belczynski K., Kalogera V., Bulik T., 2002b, ApJ, 572, 407
Belczynski K., Taam R. E., Kalogera V., Rasio F. A., Bulik T., 2007, The Astrophysical Journal, 662, 504
Belczynski K., Bulik T., Fryer C. L., Ruiter A., Valsecchi F., Vink J. S., Hurley J. R., 2010, The Astrophysical Journal, 714, 1217
Belczynski K., Holz D. E., Bulik T., O'Shaughnessy R., 2016a, Nature, 534, 512
Belczynski K., et al., 2016b, A&A, 594, A97
Broekgaarden F. S., et al., 2019, preprint, (`arXiv:1905.00910`)
Brown G. E., 1995, ApJ, 440, 270
Chruslinska M., Nelemans G., Belczynski K., 2018, Monthly Notices of the Royal Astronomical Society
Dewi J. D. M., Podsiadlowski P., Sena A., 2006, Mon. Not. Roy. Astron. Soc., 368, 1742
Dominik M., Belczynski K., Fryer C., Holz D. E., Berti E., Bulik T., Mandel I., O'Shaughnessy R., 2012, ApJ, 759, 52
Dominik M., Belczynski K., Fryer C., Holz D. E., Berti E., Bulik T., Mandel I., O'Shaughnessy R., 2013, ApJ, 779, 72
Dominik M., et al., 2015, ApJ, 806, 263
Duncan K., et al., 2014, Monthly Notices of the Royal Astronomical Society, 444, 2960
Eggleton P. P., 1983, ApJ, 268, 368
Eldridge J. J., Stanway E. R., 2016, MNRAS, 462, 3302
Farrow N., Zhu X.-J., Thrane E., 2019, Astrophys. J., 876, 18
Finn L. S., Chernoff D. F., 1993, Phys. Rev. D, 47, 2198
Fishbach M., Holz D. E., Farr W. M., 2018, ApJ, 863, L41
Fryer C. L., Belczynski K., Wiktorowicz G., Dominik M., Kalogera V., Holz D. E., 2012, ApJ, 749, 91
Furlong M., et al., 2015, Monthly Notices of the Royal Astronomical Society, 450, 4486
Ge H., Webbink R. F., Chen X., Han Z., 2015, ApJ, 812, 40
Giacobbo N., Mapelli M., 2018, Monthly Notices of the Royal Astronomical Society, 480, 2011
Giacobbo N., Mapelli M., 2019, Mon. Not. Roy. Astron. Soc., 482, 2234
Giacobbo N., Mapelli M., Spera M., 2018, MNRAS, 474, 2959
Hannam M., Schmidt P., Bohé A., Haegel L., Husa S., Ohme F., Pratten G., Pürrer M., 2014, Phys. Rev. Lett., 113, 151101
Hjellming M. S., Webbink R. F., 1987, ApJ, 318, 794
Hobbs G., Lorimer D. R., Lyne A. G., Kramer M., 2005, MNRAS, 360, 974
Hogg D. W., 1999, arXiv Astrophysics e-prints,
Humphreys R. M., Davidson K., 1994, PASP, 106, 1025
Hunter J. D., 2007, Computing In Science & Engineering, 9, 90
Hurley J. R., Pols O. R., Tout C. A., 2000, Monthly Notices of the Royal Astronomical Society, 315, 543
Hurley J. R., Tout C. A., Pols O. R., 2002, Monthly Notices of the Royal Astronomical Society, 329, 897
Husa S., Khan S., Hannam M., Pürrer M., Ohme F., Forteza X. J., Bohé A., 2016, Phys. Rev. D, 93, 044006
Ivanova N., et al., 2013, The Astronomy and Astrophysics Review, 21, 59
Kewley L. J., Ellison S. L., 2008, ApJ, 681, 1183
Khan S., Husa S., Hannam M., Ohme F., Pürrer M., Forteza X. J., Bohé A., 2016, Phys. Rev. D, 93, 044007
Kiminki D. C., Kobulnicky H. A., 2012, ApJ, 751, 4
Klencki J., Moe M., Gladysz W., Chruslinska M., Holz D. E., Belczynski K., 2018, A&A, 619, A77
Kroupa P., 2001, Monthly Notices of the Royal Astronomical Society, 322, 231
Kruckow M. U., Tauris T. M., Langer N., Kramer M., Izzard R. G., 2018, MNRAS, 481, 1908
Langer N., Norman C., 2006, The Astrophysical Journal Letters, 638, L63
Leitherer C., et al., 1999, ApJS, 123, 3
Lipunov V. M., Postnov K. A., Prokhorov M. E., 1997, MNRAS, 288, 245
Livio M., Soker N., 1988, ApJ, 329, 764
Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2015, Monthly Notices of the Royal Astronomical Society, 456, 2140
Madau P., Dickinson M., 2014, Annual Review of Astronomy and Astrophysics, 52, 415
Madau P., Fragos T., 2017, ApJ, 840, 39
Mandel I., Farmer A., 2018, preprint, (`arXiv:1806.05820`)
Mandel I., de Mink S. E., 2016, Monthly Notices of the Royal Astronomical Society, 458, 2634
Mandel I., Farr W. M., Gair J. R., 2019, MNRAS,
Mapelli M., Giacobbo N., Ripamonti E., Spera M., 2017, MNRAS, 472, 2422
Mennekens N., Vanbeveren D., 2014, A&A, 564, A134
Miller M. C., 2016, General Relativity and Gravitation, 48, 95
Moe M., Di Stefano R., 2017, ApJS, 230, 15
Neijssel C. J., 2020, in preparation
Oliphant T., 2006, NumPy: A guide to NumPy, USA: Trelgol Publishing, `http://www.numpy.org/`
Oliphant T. E., 2007, Computing in Science & Engineering, 9
Opik E., 1924, Publications of the Tartu Astrofizica Observatory, 25

Paczynski B., 1976, in Eggleton P., Mitton S., Whelan J., eds, IAU Symposium Vol. 73, Structure and Evolution of Close Binary Systems. p. 75
Paczyński B., Sienkiewicz R., 1972, Acta Astron., 22, 73
Pan Y., Buonanno A., Taracchini A., Kidder L. E., Mroué A. H., Pfeiffer H. P., Scheel M. A., Szilágyi B., 2014, Phys. Rev. D, 89, 084006
Panter B., Heavens A. F., Jimenez R., 2004, Monthly Notices of the Royal Astronomical Society, 355, 764
Pavlovskii K., Ivanova N., Belczynski K., Van K. X., 2017a, MNRAS, 465, 2092
Pavlovskii K., Ivanova N., Belczynski K., Van K. X., 2017b, MNRAS, 465, 2092
Perez F., Granger B. E., Hunter J. D., 2011, Computing in Science & Engineering, 13, 13
Peters P. C., 1964, Physical Review, 136, B1224
Pfahl E., Rappaport S., Podsiadlowski P., Spruit H., 2002, ApJ, 574, 364
Podsiadlowski P., Langer N., Poelarends A. J. T., Rappaport S., Heger A., Pfahl E., 2004, ApJ, 612, 1044
Pols O. R., Schröder K. P., Hurley J. R., Tout C. A., Eggleton P. P., 1998, MNRAS, 298, 525
Portegies Zwart S. F., Yungelson L. R., 1998, A&A, 332, 173
Rafelski M., Wolfe A. M., Prochaska J. X., Neeleman M., Mendez A. J., 2012, ApJ, 755, 89
Salpeter E. E., 1955, ApJ, 121, 161
Sana H., et al., 2012, Science, 337, 444
Sana H., et al., 2013, A&A, 550, A107
Savaglio S., et al., 2005, The Astrophysical Journal, 635, 260
Schechter P., 1976, The Astrophysical Journal, 203, 297
Schneider F. R. N., Izzard R. G., Langer N., de Mink S. E., 2015, ApJ, 805, 20
Scott D. W., 2015, Multivariate density estimation: theory, practice, and visualization. John Wiley & Sons
Soberman G. E., Phinney E. S., van den Heuvel E. P. J., 1997, A&A, 327, 620
Spera M., Mapelli M., 2017, MNRAS, 470, 4739
Spera M., Mapelli M., Bressan A., 2015, MNRAS, 451, 4086
Spera M., Mapelli M., Giacobbo N., Trani A. A., Bressan A., Costa G., 2018, preprint, (arXiv:1809.04605)
Stanway E. R., Eldridge J. J., Becker G. D., 2016, MNRAS, 456, 485
Stevenson S., Vigna-Gómez A., Mandel I., Barrett J. W., Neijssel C. J., Perkins D., de Mink S. E., 2017, Nature Communications, 8
Stevenson S., Sampson M., Powell J., Vigna-Gómez A., Neijssel C. J., Szécsi D., Mandel I., 2019, preprint, p. arXiv:1904.02821 (arXiv:1904.02821)
Strolger L.-G., et al., 2004, The Astrophysical Journal, 613, 200
Tauris T. M., Langer N., Podsiadlowski P., 2015, MNRAS, 451, 2123
Tauris T. M., et al., 2017, ApJ, 846, 170
Taylor P., Kobayashi C., 2015, MNRAS, 448, 1835
Tremonti C. A., et al., 2004, The Astrophysical Journal, 613, 898
Tutukov A. V., Yungelson L. R., 1993, MNRAS, 260, 675
Venumadhav T., Zackay B., Roulet J., Dai L., Zaldarriaga M., 2019, preprint, (arXiv:1904.07214)
Vigna-Gómez A., et al., 2018, MNRAS, 481, 4009
Vink J. S., de Koter A., Lamers H. J. G. L. M., 2001, Astronomy and Astrophysics, 369, 574
Vitale S., Farr W. M., 2018, preprint, (ArXiv:1808.00901)
Webbink R., 1984, The Astrophysical Journal, 277, 355
Woods T. E., Ivanova N., 2011, ApJ, 739, L48
Woosley S. E., 2017, ApJ, 836, 244
Wyrzykowski Ł., Mandel I., 2019, preprint, (arXiv:1904.07789)
Xu X.-J., Li X.-D., 2010, ApJ, 716, 114
de Kool M., 1990, ApJ, 358, 189
de Mink S. E., Belczynski K., 2015, The Astrophysical Journal, 814, 58
van den Heuvel E. P. J., De Loore C., 1973, A&A, 25, 387
van den Heuvel E. P. J., Portegies Zwart S. F., de Mink S. E., 2017, MNRAS, 471, 4256

APPENDIX A: METALLICITY SPECIFIC STAR FORMATION RATE

A1 Cosmological Star Formation Rate - SFR

We consider several prescriptions for the cosmological SFR as a function of redshift z. The first is from Madau & Dickinson (2014):

$$\frac{d^2M_{\rm SFR}}{dt_s dV_c}(z) = 0.015\frac{(1+z)^{2.7}}{1+[(1+z)/2.9]^{5.6}} \ \ {\rm M}_\odot\ {\rm yr}^{-1}\ {\rm Mpc}^{-3}. \tag{A1}$$

At higher redshifts the observations become more sensitive to extinction which is not exactly known. Strolger et al. (2004) construct a fit for the SFR using a different extinction correction, as

$$\frac{d^2M_{\rm SFR}}{dt_s dV_c}(t) = 0.182\times \tag{A2}$$
$$\left(t^{1.26}e^{-t/1.865} + 0.071\ e^{0.071(t-t_0)/1.865)}\right) {\rm M}_\odot\ {\rm yr}^{-1}\ {\rm Mpc}^{-3},$$

where $t(z)$ is the age of the Universe at redshift z in Gyrs, and t_0 is the current age of the Universe, which they set to 13.47 Gyrs. These two SFR models agree at low redshifts, $z \lesssim 2$, where both models peak; however, the model of Strolger et al. (2004) has a shallower drop off at higher redshifts (see Fig. 5). Simulations so far have not independently constrained the SFR at high redshifts. There are, for example, additional uncertainties such as the role of active galactic nuclei and feedback on the interstellar medium (Taylor & Kobayashi 2015).

Madau & Fragos (2017) use an updated SFR compared to Madau & Dickinson (2014). A key difference is assuming a broken power-law IMF by Kroupa (2001) instead of the classic power-law by Salpeter (1955). This increases the relative number of massive stars and therefore lowers the overall SFR normalisation by a factor of 0.66. The shape of the Madau & Fragos (2017) and Madau & Dickinson (2014) SFR models is similar, and we generally use the Madau & Dickinson (2014) prescription in our analysis. However, we adjust the low-redshift normalisation of our preferred model to approximately match the more recent estimate of Madau & Fragos (2017).

A2 Galaxy Stellar Mass to Metallicity - MZ-Relation

As described in section 4, we can construct star-forming metallicity density functions by convolving the galaxy stellar mass distribution with the MZ relation, which connects the galaxy stellar mass (M_*) and metallicity. We describe the MZ relations considered in this work in this subsection, and the GSMFs in the next one.

Stellar metallicities are assumed to match the metallicity of the interstellar gas of their surroundings at their birth.

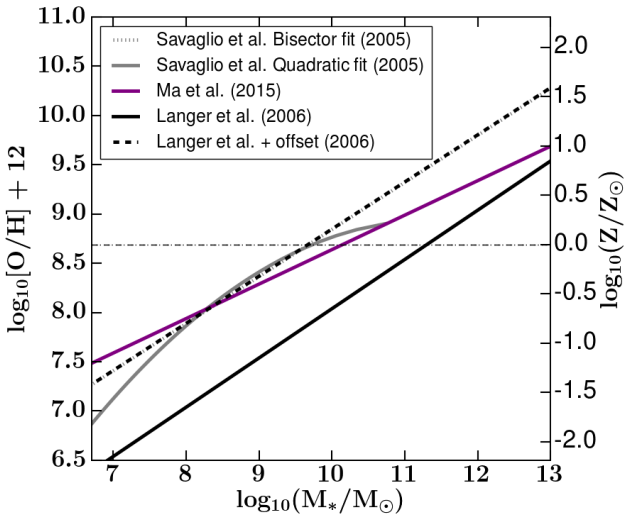

Figure A1. Comparison between different MZ relations at a redshift of 0.7, at which the relation of Savaglio et al. (2005) was determined. We also show our introduced offset to Langer & Norman (2006) which overlaps with the bisector fit of Savaglio et al. (2005). Note that the extrapolation of the quadratic fir of Savaglio et al. (2005) beyond their upper limit of $\log_{10}(\mathrm{M}_*) = 11$ results in a turnover. The dot-dashed horizontal line is our definition of the relation between solar metallicity and solar oxygen number density (Asplund et al. 2009).

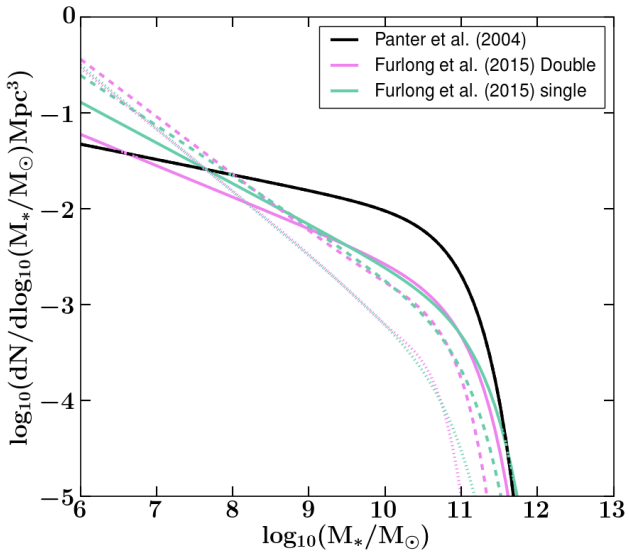

Figure A2. Comparison between the galaxy stellar mass density functions at redshifts $z = 0$ (solid), $z = 1.5$ (dashed), and $z = 3$ (dotted). The relation by Panter et al. (2004) is independent of redshift, therefore there is only a single curve (black). The double Schechter function by Furlong et al. (2015) (pink) has a steeper drop off at higher galaxy stellar masses compared to their single Schechter function (mint). Both power-law slopes of Furlong et al. (2015) are steeper than the Schechter function of Panter et al. (2004). This shifts the distribution toward lower galaxy stellar masses, which translate to lower metallicities.

Observations are typically given in terms of the ratio of the number density of oxygen and hydrogen in the gas, generally written as $\log_{10}[\mathrm{O/H}] + 12$. Conversions to metallicity depend on the assumed solar abundances. In this study we define the solar metallicity mass fraction as $Z_\odot = 0.0142$ and the solar oxygen abundance of $\log_{10}[\mathrm{O/H}]_\odot + 12 = 8.69$ based on Asplund et al. (2009), but see appendix A5. Ma et al. (2015) discuss some of the uncertainties in the slopes and offsets in the MZ relation, including the use of different observational samples or metallicity diagnostics, or the use of different simulation resolutions and feedback mechanisms in theoretical models (e.g., Taylor & Kobayashi 2015).

In Barrett et al. (2018) we used the prescriptions of Langer & Norman (2006), who in turn use a MZ relation from Savaglio et al. (2005). This MZ relation is derived from a fit of 56 galaxies in the Gemini Deep Deep Survey with a mean redshift of around 0.7. Savaglio et al. (2005) provide a quadratic and linear bisector fit, the latter being

$$\log_{10}[\mathrm{O/H}] + 12 = 0.478\ \log_{10}\left(\frac{M_*}{M_\odot}\right) + 4.062. \tag{A3}$$

We use the bisector fit because it is a monotonically increasing function of galaxy mass. The large differences at higher masses between the fits are largely due to the inclusion or exclusion of just four high-mass galaxies (Savaglio et al. 2005), illustrating the uncertainty at the extreme ends of MZ relations. Langer & Norman (2006) approximate this fit with a simplified MZ relation:

$$\frac{M_*}{M_x} = \left(\frac{Z}{Z_\odot}\right)^2, \tag{A4}$$

where $M_\mathrm{x} = 7.64 \times 10^{10}\,\mathrm{M}_\odot$ (Panter et al. 2004). Langer & Norman (2006) assume that the mean metallicity decreases exponentially with redshift as,

$$\langle Z \rangle = Z_\odot 10^{-0.3z}. \tag{A5}$$

When we translate this back into a MZ relation we find that there is difference between the approximate Langer & Norman (2006) MZ relation and the fit of Savaglio et al. (2005) (see Fig. A1). We introduce an offset to the model of Langer & Norman (2006) in order to recover the relation by Savaglio et al. (2005). This offset together with the original redshift scaling results in a high mean metallicity at redshift zero, but we keep this as an alternative model to look at its effects.

The second MZ-relation we consider is a theoretical model due to Ma et al. (2015). They combine cosmological simulations with stellar population synthesis models and a variety of feedback mechanisms to trace the evolution of the interstellar gas, for galaxy stellar masses ranging between $4 \leq \log_{10}(M_*/\mathrm{M}_\odot) \leq 11$ and redshifts between 0–6. Ma et al. (2015) give the MZ-relationship as

$$\log_{10}\left(\frac{Z_\mathrm{gas}}{Z_\odot}\right) = 0.35\left[\log_{10}\left(\frac{M_*}{M_\odot}\right) - 10\right] + 0.93e^{-0.43z} - 1.05. \tag{A6}$$

A3 Galaxy Stellar Mass Density Function - GSMF

The GSMF is empirically constructed by converting the luminosity of a sample of galaxies into a stellar mass, assuming

a mass-to-light ratio. Although samples and methods differ between compilations, Baldry et al. (2008) show that for galaxies within the mass range of $8.5 \leq \log_{10}(M_*/\mathrm{M}_\odot) \leq 12$ at redshift $z < 0.1$, there is good agreement on the shape of the GSMF.

The general shape is that of a Schechter function (Schechter 1976):

$$\Phi_{M_*}(z)dM = \phi_1(z) \left(\frac{M_*}{M_c(z)}\right)^{-\alpha(z)} e^{-\frac{M_*}{M_c(z)}} dM, \tag{A7}$$

where α determines the slope of the GSMF at the low-mass end, M_c is the turnover mass, and ϕ_1 the overall normalisation. However, a double Schechter function appears to better fit the extreme mass ends of the GSMF (Baldry et al. 2008; Furlong et al. 2015):

$$\Phi_{M_*}(z)dM = e^{-\frac{M_*}{M_c(z)}} \times \left[\phi_1(z) \left(\frac{M_*}{M_c(z)}\right)^{-\alpha_1(z)} + \phi_2(z) \left(\frac{M_*}{M_c(z)}\right)^{-\alpha_2(z)}\right] dM. \tag{A8}$$

The double Schechter function fit determined from the EAGLE simulations by Furlong et al. (2015) is able to reproduce the empirical observations of Duncan et al. (2014). We performed a linear fit to the tabulated coefficient values in the appendix of Furlong et al. (2015) (see their table A1) to recover both a single and double Schechter GSMF. Their results are for redshifts in the range $0.1 < z < 4$ and we linearly interpolate the coefficients within that range. We also extrapolate for lower and higher redshifts. In order to avoid unphysical behaviour, we set ϕ_2, which is zero at $z = 0.1$ to also be zero at all redshifts below 0.1; fix $\alpha_2 = -1.79$ at $z \leq 0.5$; and enforce $\alpha \geq -1.99$ everywhere. This allows us to extrapolate the Furlong et al. (2015) GSMF over the full range $z \in [0, 6.5]$.

The GSMFs has an overall normalisation which in principle carries information on the star formation history, although Furlong et al. (2015) note that the normalisation of their fits is imperfect at the highest redshifts, while the slope remains well fitted. However, we use a simplified model in which the SFR is independent of the GSMF, allowing us to independently parametrise and test the SFR and the metallicity distribution. Consequently, the normalisation coefficients ϕ are relevant only for describing the ratio between the two Schechter functions in Eq.A8.

Figure A2 shows the different GSMF relations at a few redshifts. For comparison we also use a *redshift-independent* single Schechter function of Panter et al. (2004) as used in Langer & Norman (2006) and our previous work (Barrett et al. 2018).

Even though the Panter et al. (2004) GSMF is redshift-independent, the metallicity distribution still changes due to the redshift dependence in the MZ relation. Meanwhile, the Furlong et al. (2015) GSMF is redshift-dependent: as galaxies grow over time, the mass distribution shifts toward higher masses at lower redshifts (Duncan et al. 2014). Conversely, the masses are lower at higher redshifts, favouring lower metallicity. Coupled with a redshift-dependent MZ relation, this further reduces mean metallicity at higher redshifts.

A4 Metallicity Specific Star Formation Rate - MSSFR

The MZ relation allows us to convert the GSMF into a metallicity distribution dP/dZ (the last term of Eq. 5). In practice, when integrating over metallicity, we sum over discrete bins. We convert the edges of those bins into limits on galaxy stellar masses in order to determine the fraction of star formation that happens in a given metallicity bin as the fraction of the GSMF that falls into the appropriate mass range at a given redshift.

We convert the number density of Eq. A8 into a mass density by multiplying by M_*. The form of this equation makes it possible to carry out the mass integral analytically, with the amount of mass at $M_* \leq M_x$ given through the incomplete gamma functions $\hat{\Gamma}$:

$$\int_0^{M_x} M_* \Phi_{M_*} dM_* = \Phi_1 \hat{\Gamma}(\alpha_1 + 2, \frac{M_x}{M_c}) + \Phi_2 \hat{\Gamma}(\alpha_2 + 2, \frac{M_x}{M_c}) \tag{A9}$$

The fraction of mass in the range between $M_x \leq M_* \leq M_y$ can be obtained from the above equation after normalisation with the complete gamma function Γ. Figure 6 shows several of the resulting star formation metallicity distributions at a few redshifts.

We compute the MSSFR by multiplying the metallicity distribution at a given redshift by the SFR at that redshift (Eq.5). Altogether we test the effect of 18 variations (2 SFR $\times$ 3 MZ $\times$ 3 GSMF), as well as our preferred MSSFR model. The two SFR variations differ mostly at redshifts above 2. The MZ relations span the range between extra-solar and sub-solar metallicities at $z = 0$. The GSMFs variants include a static redshift-independent fit and two redshift-dependent fits, which evolve toward higher galaxy stellar masses at lower redshifts (see tables C1, C2).

A5 Definition of Solar Values

In this study we defined the solar metallicity mass fraction as $Z_\odot = 0.0142$ and the solar oxygen abundance as $\log_{10}[\mathrm{O/H}]_\odot + 12 = 8.69$ based on Asplund et al. (2009). However, the assumed solar values differ between papers, so our choice is not always consistent with the fits used.

In particular, Ma et al. (2015) assume a mass fraction of $Z_\odot = 0.02$ and a specific iron mass fraction of 0.00173 to obtain an oxygen abundance of $\log_{10}[\mathrm{O/H}]_\odot + 12 = 9.0$. Savaglio et al. (2005) assume an oxygen abundance of 8.69, but mention that systematics can lead to uncertainties in the range between 8.7 and 9.1. Meanwhile, their single stellar models for their galaxy models assume a mass fraction of $Z_\odot = 0.02$ (Leitherer et al. 1999). On the other hand, Furlong et al. (2015) use a solar mass fraction of $Z_\odot = 0.0127$.

We evaluate the impact of the assumed solar metallicity and oxygen abundances on our predictions by varying these within a single MSSFR model. Figure A3) shows the predicted number of BBH detections per year at the sensitivity of the first observing run over a two-dimensional grid of solar metallicities and oxygen abundances, while keeping all other model parameters fixed. We see that these uncertainties alone could change the predicted values by up to a factor of 3.

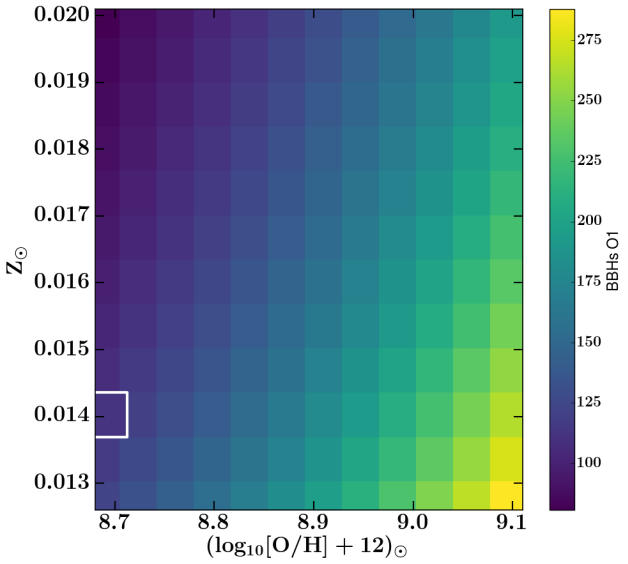

Figure A3. The predicted number of BBH detections per year at the sensitivity of the first observing run for different choices of the solar mass fraction $Z_\odot$ and oxygen abundance $(\log_{10}[O/H]+12)_\odot$. The MSSFR model is based on the SFR of Madau & Dickinson (2014), with the MZ relation of Ma et al. (2015) and the double Schechter GSMF of Furlong et al. (2015). The same (pessimistic) evolutionary model is assumed. The white square denotes the point in parameter space we assumed for this study.

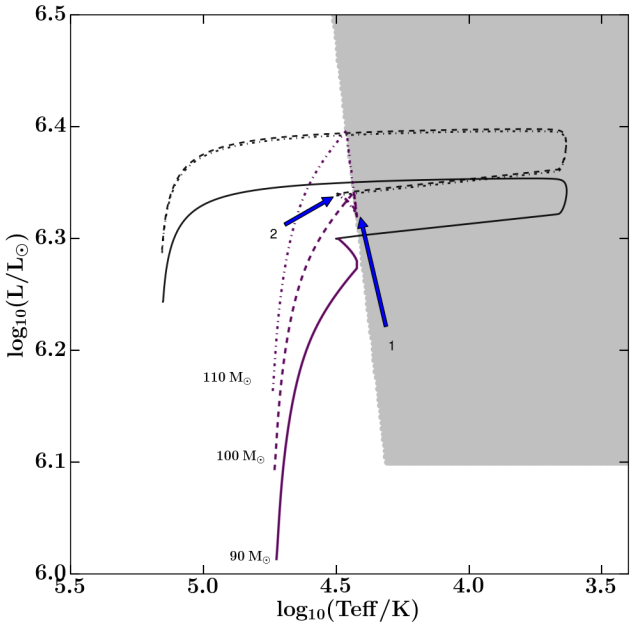

Figure B1. Single stellar tracks of stars with initial masses of 90 $M_\odot$, 100 $M_\odot$, and 110 $M_\odot$. The metallicity of the stars is $Z = Z_\odot/3$. The purple portions of the tracks indicate the core hydrogen burning phase (main sequence), while subsequent evolution is shown in black. The region in which we apply the LBV-wind mass loss rate is shaded. At point 1 the 100 $M_\odot$ and 110 $M_\odot$ stars turn off the main sequence. At point 2 they start evolving onto the Hertzsprung gap. The two tracks evolve identically from point 1, resulting in the same remnant mass.

APPENDIX B: REMNANT MASSES OF SINGLE STARS

In COMPAS models, metallicity impacts the masses of compact remnants by influencing stellar evolutionary tracks and the rates of wind-driven mass loss. The Fryer et al. (2012) recipes for calculating the remnant mass are based on the mass of the carbon-oxygen core and the total stellar mass at the moment of the supernova. If the carbon-oxygen core mass exceeds 16 $M_\odot$, there is assumed to be no explosion and the remnant mass is the same as the total mass of the star before the supernova.

However, within our models, all single stars above a certain initial mass yield the same remnant mass at a given metallicity.[3] This is driven by our implementation of LBV-wind mass loss (see Sec. 2).

Figure B1 shows three tracks of very massive single stars. The shaded region is where we apply the LBV-wind mass loss rates (Belczynski et al. 2010). When stars are on the main sequence (core hydrogen burning phase), they evolve on a nuclear timescale, which is not sufficiently fast to overcome the LBV winds and pass through the Humphreys-Davidson limit (Humphreys & Davidson 1994) into the shaded region. At point 1 the stars start to turn off the main sequence. By this time stars with initial masses of 100 and 110 solar masses have the same mass. At point 2 they begin to evolve onto the Hertzsprung gap. It is at this point that the analytical fits of Hurley et al. (2000) define a core mass. This core mass only depends on the current properties of the stars, so the two stars continue evolving identically. Their faster, thermal-timescale evolution now allows them to pass through the Humphreys-Davidson limit and enter the shaded region. They end up having the same remnant mass.

The process described above can yield sharp peaks in the BBH mass distribution. Every binary in which both stars go through this LBV phase on the main sequence will end up with the same total BBH mass. This is the maximum total mass for a given metallicity. The lower the initial mass for this LBV-wind mass loss, i.e., the higher the metallicity, the more binaries will have degenerate remnant masses. Hence the sharpest peaks in the total BBH mass distribution are for metallicities around a third solar (see Fig. 2). This feature of the COMPAS and StarTrack implementation of LBV winds also explains the asymptote of the maximal remnant mass in figure 1 of Belczynski et al. (2010) and the peaks in the highest mass bins of Dominik et al. (2012).

[3] These simulations do not include PISNe or pulsational PISNe. Future COMPAS analyses will incorporate them (Stevenson et al. 2019).

APPENDIX C: STATISTICS

In this appendix, we describe our procedures for computing the likelihood of a given MSSFR model given the observed number of detections and their chirp masses, and describe the use of bootstrapping to estimate the Monte Carlo simulation uncertainty.

C1 Evaluating model likelihoods

We can write the total likelihood $\mathcal{L}_{\rm tot}(d|M)$ of observing the data set d, which consists of $N_{\rm obs}$ detections with individual data d_i, given a model M that predicts N_M expected detections with a probability distribution of source properties P_M, as (e.g., Mandel et al. 2019)

$$\mathcal{L}_{\rm tot}(d|M) = \frac{N_M^{N_{obs}}}{N_{obs}!} e^{-N_M} \prod_{i=1}^{N_{obs}} p(d_i|P_M). \tag{C1}$$

Here, we focus on the chirp mass $\mathcal{M}_c$, as the parameter which is best constrained by gravitational wave observations and is directly predicted by COMPAS simulations. Writing the preceding equation in logarithmic form, the log-likelihood of a particular MSSFR model is

$$\log_{10}\left(\mathcal{L}_{\rm tot}(d|M)\right) = \log_{10}\left(\mathcal{L}(N_{\rm obs}|N_M)\right) + \sum_{i=1}^{N_{\rm obs}} \log_{10}\left(\mathcal{L}(\mathcal{M}_{c,i}|p_M(\mathcal{M}_c))\right), \tag{C2}$$

where $\mathcal{M}_{c,i}$ is the measured chirp-mass of the i'th gravitational wave observation and $p_M(\mathcal{M}_c)$ is chirp-mass distribution characterising the MSSFR model M. The first term is abbreviated as $\mathcal{L}_{\rm R}$ in table C2. The second term, $\mathcal{L}_{\rm M_{chirp}} \equiv \mathcal{L}(\mathcal{M}_{c,i}|p_M(\mathcal{M}_c))$ is the probability of detecting a chirp mass $\mathcal{M}_{c,i}$ given the chirp-mass distribution predicted from the MSSFR model M.

COMPAS Monte Carlo simulations yield a discrete set of chirp masses and their respective rates. A kernel density estimator is used to turn this set of discrete data points into an approximated continuous function. We do this by approximating each of $N_{\rm sim}$ chirp masses produced by the COMPAS simulation as a 1-dimensional Gaussian centred on the simulated chirp-mass value $\mathcal{M}_j$. All Gaussians have the same bandwidth σ, determined using the default 'Scott's rule' (Scott 2015) of the Gaussian kernel density estimator in the `scipy` package (Oliphant 2007; Perez et al. 2011). Each simulated data point j contributes to the overall probability density function proportionally to its observing rate R_j, estimated in Eq. 2. Therefore we re-weigh each data point by R_j and normalise by the total rate $R_{\rm tot}$,

$$p_M(\mathcal{M}_c) = \frac{1}{R_{\rm tot}} \sum_{j=1}^{N_{\rm sim}} R_j \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\mathcal{M}_c - \mathcal{M}_j)^2}{2\sigma^2}}. \tag{C3}$$

For a single perfect detection, the likelihood of observing a chirp-mass $\mathcal{M}_c$ would be given by Eq. C3. In practice, gravitational-wave measurements suffer from observational uncertainty, although these are typically small for chirp masses. Chirp-mass posteriors of individual detections were not yet available when this work started; therefore, we reconstruct them as symmetric Gaussian distributions with 90 per cent confidence intervals matching those reported in Abbott et al. (2016); Abbott et al. (2018a). The reported error-bars are asymmetric, so the median of our reconstructed posterior is slightly shifted compared to the original. Given the accuracy of chirp mass measurement, we make two further simplifications. We ignore the impact of the priors used in Abbott et al. (2018a) (which is reasonable inasmuch as the posterior is determined by the sharply peaked likelihood function), and do not reweigh by those priors; and we ignore the selection effects on the chirp mass for the purpose of population analysis, since the selection function does not vary significantly over the range of likelihood support (see Mandel et al. (2019) for a discussion of both issues). With these simplifications, the likelihood of observing a particular gravitational wave event i, characterised by the approximated Gaussian posterior of the chirp mass $p_i(\mathcal{M}_c)$, given a MSSFR model M, becomes

$$\mathcal{L}(\mathcal{M}_{c,i}|p_M(\mathcal{M}_c)) = \int_0^\infty p_i(\mathcal{M}_c)\; p_M(\mathcal{M}_c)\; d\mathcal{M}_c. \tag{C4}$$

Figure C1 shows our constructed posterior for GW150914 (red); part of the chirp-mass distribution estimated from the MSSFR model which combines the SFR of Madau & Dickinson (2014), the MZ relation of Ma et al. (2015) and the GSMF of Furlong et al. (2015) (blue); and the convolution between the two (black). The integral of this convolution is our estimate of the likelihood $\mathcal{L}(\mathcal{M}_{c,i}|p_M(\mathcal{M}_c))$.

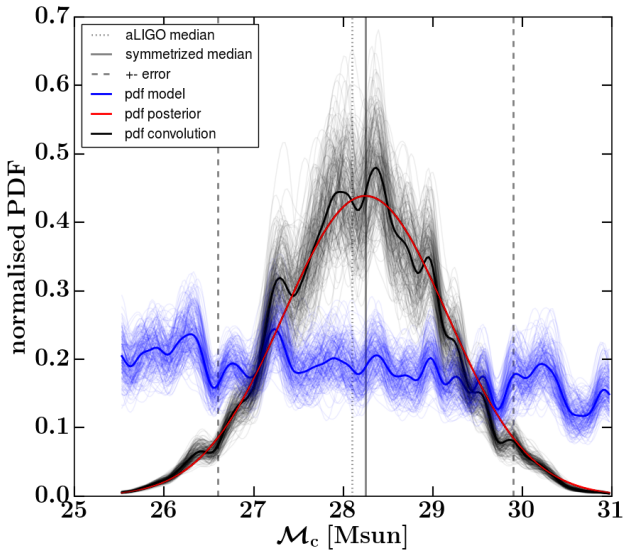

Figure C1. An example of the likelihood calculation for GW150914. The dashed vertical lines show the 90 per cent confidence interval from Abbott et al. (2016). The dotted vertical line is the median from Abbott et al. (2016) and the solid vertical line is the median after symmetrising. The red curve is a mock gaussian posterior. The blue curve is part of the normalised chirp mass distribution obtained by applying a one dimensional KDE to the results of the COMPAS simulation. The black curve shows the Gaussian likelihood convolved with the model. The fainter lines show scatter in the chirp-mass distribution and the convolution as estimated with bootstrapping.

C2 Bootstrapping

Our simulation is based on a Monte Carlo sampling of binaries. We estimate the sampling uncertainty on all derived quantities via bootstrapping: we uniformly re-sample a set with the same total number of binaries from our already evolved initial set of binaries (with replacement), including systems which did not form a DCO. The central value in tables C1, C2 corresponds to the original sample, while the error bars correspond to the 5th and 95th percentile rates and likelihoods from bootstrapping.

MSSFR Variation			BBH Rates		BHNS Rates		BNS Rates	
SFR	MZ	GSMF	$z = 0$ merg. $\mathrm{Gpc^{-3}\ yr^{-1}}$	O1 det. $\mathrm{yr^{-1}}$	$z = 0$ merg. $\mathrm{Gpc^{-3}\ yr^{-1}}$	O1 det. $\mathrm{yr^{-1}}$	$z = 0$ merg. $\mathrm{Gpc^{-3}\ yr^{-1}}$	O1 det. $\mathrm{yr^{-1}}$
Pessimistic								
Preferred model			$49.00^{+1.93}_{-1.68}$	$21.80^{+0.47}_{-0.50}$	$56.87^{+1.80}_{-1.89}$	$0.07^{+0.00}_{-0.00}$	$20.00^{+1.34}_{-1.03}$	$0.23^{+0.02}_{-0.015}$
Madau et al.	Ma et al. (2004)	1	$63.07^{+1.86}_{-1.76}$	$18.43^{+0.42}_{-0.4}$	$32.22^{+1.7}_{-1.43}$	$0.4^{+0.02}_{-0.02}$	$85.97^{+2.47}_{-2.41}$	$0.11^{+0.0}_{-0.0}$
		2	$158.56^{+2.07}_{-2.48}$	$94.35^{+1.37}_{-1.39}$	$40.73^{+1.37}_{-1.13}$	$0.51^{+0.02}_{-0.02}$	$90.87^{+2.39}_{-2.15}$	$0.12^{+0.0}_{-0.0}$
		3	$174.71^{+2.28}_{-2.7}$	$113.92^{+1.12}_{-1.22}$	$42.14^{+1.43}_{-1.22}$	$0.52^{+0.02}_{-0.02}$	$91.42^{+2.37}_{-2.24}$	$0.13^{+0.0}_{-0.0}$
	Langer et al.	1	$448.84^{+4.28}_{-5.24}$	$247.22^{+2.63}_{-2.78}$	$95.47^{+1.85}_{-2.12}$	$1.28^{+0.03}_{-0.03}$	$144.57^{+2.08}_{-2.3}$	$0.22^{+0.0}_{-0.0}$
		2	$563.44^{+4.1}_{-5.9}$	$441.08^{+5.72}_{-6.24}$	$91.76^{+1.73}_{-1.8}$	$1.19^{+0.03}_{-0.03}$	$143.81^{+1.93}_{-2.1}$	$0.22^{+0.0}_{-0.0}$
		3	$589.13^{+4.27}_{-6.05}$	$492.27^{+5.85}_{-6.63}$	$96.32^{+1.9}_{-1.9}$	$1.25^{+0.03}_{-0.03}$	$147.1^{+1.92}_{-2.3}$	$0.23^{+0.0}_{-0.0}$
	Langer et al., offset	1	$59.17^{+1.27}_{-1.34}$	$28.72^{+0.46}_{-0.45}$	$18.98^{+1.01}_{-0.77}$	$0.23^{+0.01}_{-0.01}$	$72.09^{+3.21}_{-3.01}$	$0.09^{+0.0}_{-0.0}$
		2	$151.84^{+1.88}_{-2.23}$	$120.3^{+2.12}_{-2.23}$	$28.67^{+0.92}_{-0.88}$	$0.35^{+0.01}_{-0.01}$	$79.27^{+2.67}_{-2.52}$	$0.11^{+0.0}_{-0.0}$
		3	$167.64^{+1.92}_{-2.29}$	$148.74^{+2.0}_{-2.05}$	$29.4^{+0.97}_{-0.98}$	$0.35^{+0.01}_{-0.01}$	$79.32^{+2.57}_{-2.51}$	$0.11^{+0.0}_{-0.0}$
Strolger et al.	Ma et al. (2004)	1	$101.98^{+3.01}_{-2.88}$	$32.93^{+0.83}_{-0.84}$	$42.11^{+2.4}_{-1.81}$	$0.52^{+0.03}_{-0.02}$	$91.66^{+2.71}_{-2.49}$	$0.12^{+0.0}_{-0.0}$
		2	$255.46^{+4.49}_{-5.7}$	$203.93^{+5.17}_{-4.76}$	$49.05^{+2.1}_{-1.68}$	$0.6^{+0.02}_{-0.02}$	$98.87^{+2.59}_{-2.72}$	$0.14^{+0.0}_{-0.0}$
		3	$271.91^{+4.09}_{-4.67}$	$208.21^{+3.29}_{-3.38}$	$50.49^{+1.98}_{-1.66}$	$0.61^{+0.03}_{-0.02}$	$98.72^{+2.4}_{-2.72}$	$0.14^{+0.0}_{-0.0}$
	Langer et al.	1	$574.91^{+6.58}_{-8.19}$	$406.39^{+5.25}_{-5.66}$	$99.35^{+2.32}_{-2.43}$	$1.28^{+0.03}_{-0.04}$	$153.78^{+2.74}_{-2.9}$	$0.23^{+0.0}_{-0.0}$
		2	$688.91^{+9.03}_{-10.42}$	$659.25^{+15.27}_{-14.58}$	$95.9^{+2.65}_{-2.62}$	$1.19^{+0.03}_{-0.03}$	$153.93^{+3.49}_{-3.4}$	$0.24^{+0.01}_{-0.01}$
		3	$714.15^{+8.77}_{-11.17}$	$710.91^{+15.17}_{-14.29}$	$100.12^{+2.67}_{-2.69}$	$1.25^{+0.03}_{-0.03}$	$157.4^{+3.8}_{-3.43}$	$0.24^{+0.01}_{-0.01}$
	Langer et al., offset	1	$132.55^{+3.38}_{-3.54}$	$89.79^{+1.42}_{-1.42}$	$27.88^{+1.96}_{-1.27}$	$0.33^{+0.02}_{-0.01}$	$79.81^{+3.16}_{-3.29}$	$0.11^{+0.0}_{-0.0}$
		2	$259.14^{+5.65}_{-6.53}$	$267.34^{+8.2}_{-7.55}$	$36.67^{+1.58}_{-1.57}$	$0.43^{+0.02}_{-0.02}$	$88.47^{+3.1}_{-3.58}$	$0.12^{+0.01}_{-0.0}$
		3	$276.57^{+5.32}_{-6.48}$	$292.76^{+7.71}_{-6.55}$	$37.36^{+1.64}_{-1.56}$	$0.43^{+0.02}_{-0.02}$	$88.34^{+2.88}_{-3.29}$	$0.12^{+0.0}_{-0.0}$
Optimistic								
Preferred model			$190.85^{+3.94}_{-3.99}$	$36.80^{+0.65}_{-0.65}$	$158.8^{+3.13}_{-3.49}$	$0.23^{+0.00}_{-0.01}$	$56.68_{-1.87}+2.21$	$0.51^{+0.02}_{-0.02}$
Madau et al.	Ma et al. (2004)	1	$291.23^{+4.93}_{-5.28}$	$46.27^{+0.7}_{-0.72}$	$89.32^{+3.19}_{-2.47}$	$0.85^{+0.03}_{-0.02}$	$231.08^{+4.36}_{-4.8}$	$0.33^{+0.01}_{-0.01}$
		2	$408.74^{+4.03}_{-5.52}$	$128.37^{+1.44}_{-1.48}$	$102.22^{+2.67}_{-2.07}$	$1.04^{+0.03}_{-0.02}$	$227.38^{+4.13}_{-4.46}$	$0.33^{+0.01}_{-0.01}$
		3	$431.55^{+4.5}_{-5.58}$	$148.55^{+1.26}_{-1.38}$	$104.83^{+2.71}_{-2.18}$	$1.06^{+0.03}_{-0.02}$	$226.34^{+3.71}_{-4.21}$	$0.33^{+0.01}_{-0.01}$
	Langer et al.	1	$938.86^{+6.16}_{-7.01}$	$332.4^{+2.92}_{-2.77}$	$201.14^{+3.44}_{-3.1}$	$2.33^{+0.04}_{-0.04}$	$209.7^{+2.85}_{-2.73}$	$0.33^{+0.0}_{-0.0}$
		2	$1002.87^{+5.33}_{-6.55}$	$517.59^{+5.41}_{-6.39}$	$187.17^{+2.79}_{-2.49}$	$2.16^{+0.04}_{-0.03}$	$210.0^{+2.47}_{-2.48}$	$0.33^{+0.0}_{-0.0}$
		3	$1042.93^{+5.27}_{-6.97}$	$572.54^{+5.69}_{-6.41}$	$193.78^{+2.98}_{-2.65}$	$2.25^{+0.04}_{-0.03}$	$208.21^{+2.55}_{-2.49}$	$0.33^{+0.0}_{-0.0}$
	Langer et al., offset	1	$190.7^{+3.5}_{-3.65}$	$43.44^{+0.52}_{-0.52}$	$59.21^{+2.69}_{-2.15}$	$0.52^{+0.02}_{-0.01}$	$236.54^{+6.48}_{-6.59}$	$0.34^{+0.01}_{-0.01}$
		2	$317.64^{+3.16}_{-4.02}$	$141.94^{+2.28}_{-2.08}$	$74.72^{+2.18}_{-1.91}$	$0.73^{+0.02}_{-0.02}$	$232.11^{+5.66}_{-5.96}$	$0.34^{+0.01}_{-0.01}$
		3	$335.02^{+3.3}_{-4.01}$	$170.2^{+2.17}_{-2.02}$	$75.38^{+2.2}_{-1.9}$	$0.73^{+0.02}_{-0.02}$	$231.52^{+5.68}_{-5.52}$	$0.33^{+0.01}_{-0.01}$
Strolger et al.	Ma et al. (2004)	1	$361.59^{+5.79}_{-6.7}$	$64.2^{+1.08}_{-1.12}$	$111.51^{+4.29}_{-3.31}$	$1.09^{+0.04}_{-0.03}$	$238.44^{+4.5}_{-4.66}$	$0.34^{+0.01}_{-0.01}$
		2	$523.11^{+5.83}_{-8.21}$	$239.42^{+5.7}_{-4.94}$	$116.48^{+3.27}_{-2.7}$	$1.18^{+0.03}_{-0.03}$	$235.98^{+4.6}_{-4.62}$	$0.34^{+0.01}_{-0.01}$
		3	$547.87^{+6.08}_{-7.09}$	$244.58^{+3.4}_{-3.41}$	$119.46^{+3.05}_{-2.75}$	$1.22^{+0.03}_{-0.03}$	$234.53^{+3.89}_{-4.63}$	$0.34^{+0.01}_{-0.01}$
	Langer et al.	1	$1061.19^{+8.04}_{-9.88}$	$486.78^{+5.82}_{-5.84}$	$204.37^{+3.37}_{-3.35}$	$2.33^{+0.05}_{-0.04}$	$220.3^{+3.44}_{-3.03}$	$0.34^{+0.01}_{-0.01}$
		2	$1118.71^{+8.74}_{-10.66}$	$730.55^{+14.82}_{-14.5}$	$189.09^{+2.96}_{-3.06}$	$2.14^{+0.04}_{-0.04}$	$221.11^{+4.03}_{-3.54}$	$0.34^{+0.01}_{-0.01}$
		3	$1157.22^{+9.06}_{-11.19}$	$785.44^{+14.83}_{-14.23}$	$194.98^{+3.02}_{-3.31}$	$2.22^{+0.04}_{-0.04}$	$219.59^{+3.93}_{-3.66}$	$0.34^{+0.01}_{-0.01}$
	Langer et al., offset	1	$291.8^{+5.57}_{-5.44}$	$107.66^{+1.37}_{-1.53}$	$76.78^{+3.4}_{-2.67}$	$0.7^{+0.02}_{-0.02}$	$244.96^{+6.31}_{-6.62}$	$0.35^{+0.01}_{-0.01}$
		2	$440.74^{+6.82}_{-7.95}$	$290.5^{+8.53}_{-7.21}$	$87.17^{+2.78}_{-2.77}$	$0.86^{+0.03}_{-0.02}$	$240.95^{+6.21}_{-6.05}$	$0.35^{+0.01}_{-0.01}$
		3	$460.27^{+6.42}_{-7.84}$	$315.88^{+7.94}_{-6.5}$	$87.82^{+2.79}_{-2.54}$	$0.86^{+0.02}_{-0.03}$	$240.19^{+6.13}_{-5.99}$	$0.34^{+0.01}_{-0.01}$

Table C1. Table showing the merger and detection rates per DCO type. The columns labeled '$z = 0$ merg.' are the merger rate per year per cubic gigaparsec at zero redshift without selection effects; columns labeled 'O1 det.' are the expected rate of detections per year at the sensitivity of the first observing run. The error bars show the 90 per cent confidence interval due to Monte Carlo sampling evaluated via bootstrapping. The numbers in the column GSMF refer to 1=Panter et al. (2004), 2=Furlong et al. (2015) (single Schechter function), 3=Furlong et al. (2015) (double Schechter function). Optimistic and pessimistic variants relate to the ability to eject the common envelope when the donor is a Hertzsprung-gap star.

MSSFR Variation			Likelihoods ($\log_{10}$)		
SFR	MZ	GSMF	$\mathcal{L}_{\rm M_c}$	$\mathcal{L}_{\rm R}$	$\mathcal{L}_{\rm tot}$
			Pessimistic		
	Preferred model		$-32.32^{+0.16}_{-0.18}$	$-0.90^{+0.00}_{-0.00}$	$-33.22^{+0.16}_{-0.18}$
Madau et al.	Ma et al. (2004)	1	$-33.9^{+0.14}_{-0.16}$	$-0.97^{+0.01}_{-0.02}$	$-34.87^{+0.14}_{-0.16}$
		2	$-32.42^{+0.07}_{-0.08}$	$-8.86^{+0.21}_{-0.21}$	$-41.28^{+0.24}_{-0.26}$
		3	$-32.48^{+0.07}_{-0.07}$	$-11.9^{+0.19}_{-0.18}$	$-44.38^{+0.23}_{-0.22}$
	Langer et al.	1	$-32.24^{+0.05}_{-0.05}$	$-34.85^{+0.5}_{-0.47}$	$-67.09^{+0.49}_{-0.47}$
		2	$-32.61^{+0.06}_{-0.06}$	$-70.6^{+1.23}_{-1.05}$	$-103.21^{+1.22}_{-1.05}$
		3	$-32.77^{+0.06}_{-0.07}$	$-80.23^{+1.32}_{-1.04}$	$-113.0^{+1.3}_{-1.09}$
	Langer et al., offset	1	$-32.3^{+0.09}_{-0.1}$	$-1.07^{+0.02}_{-0.02}$	$-33.38^{+0.09}_{-0.1}$
		2	$-32.68^{+0.08}_{-0.08}$	$-12.93^{+0.38}_{-0.33}$	$-45.61^{+0.42}_{-0.38}$
		3	$-32.87^{+0.07}_{-0.07}$	$-17.62^{+0.35}_{-0.33}$	$-50.49^{+0.38}_{-0.36}$
Strolger et al.	Ma et al. (2004)	1	$-33.82^{+0.14}_{-0.17}$	$-1.31^{+0.06}_{-0.05}$	$-35.13^{+0.18}_{-0.17}$
		2	$-32.81^{+0.11}_{-0.1}$	$-27.14^{+0.87}_{-0.91}$	$-59.95^{+0.93}_{-0.97}$
		3	$-32.65^{+0.08}_{-0.08}$	$-27.9^{+0.61}_{-0.58}$	$-60.54^{+0.63}_{-0.61}$
	Langer et al.	1	$-32.44^{+0.06}_{-0.06}$	$-64.11^{+1.08}_{-0.97}$	$-96.55^{+1.08}_{-1.02}$
		2	$-32.98^{+0.1}_{-0.1}$	$-111.92^{+2.83}_{-2.81}$	$-144.9^{+2.9}_{-2.83}$
		3	$-33.09^{+0.09}_{-0.1}$	$-121.79^{+2.8}_{-2.83}$	$-154.87^{+2.86}_{-2.9}$
	Langer et al., offset	1	$-32.46^{+0.08}_{-0.1}$	$-8.18^{+0.21}_{-0.21}$	$-40.63^{+0.23}_{-0.26}$
		2	$-33.2^{+0.12}_{-0.12}$	$-38.48^{+1.44}_{-1.48}$	$-71.68^{+1.51}_{-1.51}$
		3	$-33.22^{+0.11}_{-0.11}$	$-43.1^{+1.29}_{-1.27}$	$-76.33^{+1.35}_{-1.34}$
			Optimistic		
	Preferred model		$-33.1^{+0.08}_{-0.11}$	$-1.58^{+0.05}_{-0.05}$	$-34.73^{+0.12}_{-0.11}$
Madau et al.	Ma et al. (2004)	1	$-36.37^{+0.11}_{-0.1}$	$-2.46^{+0.07}_{-0.07}$	$-38.84^{+0.16}_{-0.14}$
		2	$-32.61^{+0.05}_{-0.06}$	$-14.24^{+0.25}_{-0.23}$	$-46.85^{+0.27}_{-0.26}$
		3	$-32.52^{+0.04}_{-0.05}$	$-17.59^{+0.24}_{-0.21}$	$-50.1^{+0.24}_{-0.22}$
	Langer et al.	1	$-32.55^{+0.05}_{-0.04}$	$-50.38^{+0.52}_{-0.52}$	$-82.93^{+0.5}_{-0.53}$
		2	$-32.55^{+0.04}_{-0.05}$	$-85.01^{+1.21}_{-0.98}$	$-117.56^{+1.23}_{-1.02}$
		3	$-32.67^{+0.04}_{-0.05}$	$-95.41^{+1.28}_{-1.07}$	$-128.08^{+1.29}_{-1.11}$
	Langer et al., offset	1	$-32.85^{+0.07}_{-0.06}$	$-2.18^{+0.05}_{-0.05}$	$-35.03^{+0.09}_{-0.08}$
		2	$-32.49^{+0.06}_{-0.06}$	$-16.48^{+0.39}_{-0.37}$	$-48.97^{+0.43}_{-0.39}$
		3	$-32.59^{+0.05}_{-0.06}$	$-21.27^{+0.35}_{-0.36}$	$-53.85^{+0.38}_{-0.36}$
Strolger et al.	Ma et al. (2004)	1	$-35.37^{+0.11}_{-0.12}$	$-4.58^{+0.15}_{-0.14}$	$-39.95^{+0.21}_{-0.19}$
		2	$-32.54^{+0.08}_{-0.08}$	$-33.45^{+0.89}_{-0.99}$	$-65.99^{+0.95}_{-0.98}$
		3	$-32.39^{+0.05}_{-0.05}$	$-34.37^{+0.62}_{-0.6}$	$-66.77^{+0.65}_{-0.62}$
	Langer et al.	1	$-32.38^{+0.05}_{-0.05}$	$-79.19^{+1.11}_{-1.08}$	$-111.57^{+1.14}_{-1.11}$
		2	$-32.78^{+0.08}_{-0.09}$	$-125.54^{+2.79}_{-2.83}$	$-158.32^{+2.82}_{-2.86}$
		3	$-32.86^{+0.08}_{-0.08}$	$-136.06^{+2.78}_{-2.82}$	$-168.92^{+2.79}_{-2.88}$
	Langer et al., offset	1	$-32.24^{+0.06}_{-0.07}$	$-10.91^{+0.24}_{-0.21}$	$-43.15^{+0.23}_{-0.24}$
		2	$-32.83^{+0.1}_{-0.1}$	$-42.69^{+1.46}_{-1.51}$	$-75.52^{+1.53}_{-1.51}$
		3	$-32.84^{+0.09}_{-0.08}$	$-47.34^{+1.31}_{-1.35}$	$-80.18^{+1.31}_{-1.36}$

Table C2. Table showing the log likelihoods of observing the rate and chirp mass distribution of BBH mergers detected during the first two observing runs, within our default binary evolution model and for a range of MSSFR variations. $\mathcal{L}_{\rm tot}$ is the total likelihood, $\mathcal{L}_{\rm R}$ is the Poisson likelihood of observing 10 BBH events over 166 days of coincident observation, and $\mathcal{L}_{\rm M_c}$ is the likelihood of observing the chirp-mass distribution. The error bars show the 90 per cent confidence interval due to Monte Carlo sampling evaluated via bootstrapping. The numbers in the column GSMF refer to 1=Panter et al. (2004), 2=Furlong et al. (2015) (single Schechter function), 3=Furlong et al. (2015) (double Schechter function). Optimistic and pessimistic variants relate to the ability to eject the common envelope when the donor is a Hertzsprung-gap star.

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