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


DOI
10.1093/mnras/stz2840

Publication date
2019

Document Version
Final published version

Published in
Monthly Notices of the Royal Astronomical Society

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)
The effect of the metallicity-specific star formation history on double compact object mergers

Coenraad J. Neijssel, Alejandro Vigna-Gómez, Simon Stevenson, Jim W. Barrett, Sebastian M. Gaebel, Floor S. Broekgaarden, Selma E. de Mink, Dorottya Szécsi, Serena Vinciguerra, and Ilya Mandel

1 Birmingham Institute for Gravitational Wave Astronomy and School of Physics and Astronomy, University of Birmingham, Birmingham B15 2TT, UK
2 Monash Centre for Astrophysics, School of Physics and Astronomy, Monash University, Clayton, Victoria 3800, Australia
3 The ARC Center of Excellence for Gravitational Wave Discovery – OzGrav
4 Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik, D-30167 Hannover, Germany
5 DARK, Niels Bohr Institute, University of Copenhagen, Blegdamsvej 17, DK-2100 Copenhagen, Denmark
6 Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn VIC 3122, Australia
7 Klarna Bank AB (publ). Sveavägen 46, SE-111 34 Stockholm, Sweden
8 Anton Pannekoek Institute for Astronomy and GRAPPA, University of Amsterdam, Postbus 94249, NL-1090 GE Amsterdam, the Netherlands
9 Center for Astrophysics, Harvard, Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
10 I. Physikalisches Institut, Universität zu Köln, Zülpicher-Straße 77, D-50937 Cologne, Germany
11 Leibniz Universität Hannover, D-30167 Hannover, Germany

Accepted 2019 October 3. Received 2019 September 12; in original form 2019 July 22

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 catalogue 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 that scales with redshift as $\langle Z \rangle = 0.035 \times 10^{-0.23z}$, assuming a star formation rate of $0.01 \times (1+z)^{2.77}/(1+(1+z)/2.9)^{4.7} \text{M}_\odot \text{Mpc}^{-3} \text{yr}^{-1}$. Intriguingly, this default model predicts that 80 per cent 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 – (stars:) binaries: general – stars: massive – galaxies: star formation.

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 Gpc$^{-3}$ yr$^{-1}$, whereas for BNSs it is 110–3840 Gpc$^{-3}$ 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 holes (see Miller 2016; Giacobbo & Mapelli 2018; Mandel & Farmer 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 ($M > 8 \text{M}_\odot$) in the field are born in binaries (Kiminki & Kobulnicky 2012; Sana et al. 2012; Moe & D&I 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, Nelemans & Belczynski 2019).

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, Kalogera & Bulik 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 (Strolger et al. 2004; Madau & Dickinson 2014). 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, Mapelli & Spera 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 catalogues, 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 SFR (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 ($t_{\text{delay}}$). The formation time $t_i$ is related to the merger time $t_m$ by $t_i = t_m - t_{\text{delay}}$. We can calculate the rate of mergers at any given time as a function of chirp mass as

$$\frac{d^3 N_{\text{merge}}}{dt \, dV \, d\text{chirp}}(t_m) = \int dZ \int dt_{\text{delay}} \frac{d^3 N_{\text{form}}}{dt_{\text{delay}} d\text{MSSFR} d\text{chirp}}(Z) \times \frac{d^3 M_{\text{SFR}}}{dt \, dV \, dZ}(t_i = t_m - t_{\text{delay}}),$$

where $Z$ is the redshift, $t_i$ is the time in the source frame of the merger, and $V_i$ is the comoving volume. The first term in the integrand is the number of DCOs per unit chirp mass $M_{\text{chirp}}$ per unit delay time and per unit chirp mass $M_{\text{chirp}} = M_1^{s/3} M_2^{s/3} (M_1 + M_2)^{-s/3}$, 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 star-forming mass $M_{\text{SFR}}$ at the birth of the binary per unit time, volume, and metallicity (MSSFR), 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 ($P_{\text{det}}$) given its chirp-mass ($M_{\text{chirp}}$) and luminosity distance ($D_L(z)$). The total number of observable mergers ($N_{\text{obs}}$) per unit chirp mass per unit observing time ($t_{\text{obs}}$) is then

$$\frac{d^2 N_{\text{obs}}}{dt_{\text{obs}} d\text{chirp}} = \int_0^{z_{\text{max}}} dz \frac{dV_i}{dz} \frac{dV_i}{dz} P_{\text{det}}(M_{\text{chirp}}, D_L(z)), \quad (2)$$

where $dV_i/dz$ is the differential comoving volume as a function of redshift and $dV_i/dz = 1/(1+z)$ translates the rate to the observer frame (e.g. Hogg 1999). We assume a flat cosmology with $\Omega_M = 0.308$ and a Hubble constant of $H_0 = 67.8 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (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), Madar & Frongia (2017), and Chruslinska & Nelemans (2019).

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

(i) 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.

(ii) Section 3: $M_{\text{SFR}}$ -- 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.

(iii) Section 4: $d^3 M_{\text{SFR}}/dt \, dV \, d\text{chirp}$ -- MSSFR

We combine observations and simulations of galaxy stellar mass distributions with mass--metallicity relations to construct an MSSFR. These different prescriptions introduce an uncertainty into our DCO merger rate distributions. We propose a parametrized, smooth metallicity distribution, which facilitates the exploration of the MSSFR parameter space.

(iv) Section 5: $d^3 N_{\text{merge}}/dt \, dV \, d\text{chirp}$ -- 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.

(v) Section 6: $d^3 M_{\text{SFR}}/dt \, dV \, d\text{chirp}$ -- 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 F-wave events from BBH mergers.

(vi) Section 7: Discussion and Conclusion

We review our findings and discuss future prospects.
The five main initial conditions that describe a stellar binary are: the primary mass $M_1$ 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

$$m_2 = m_1 \times q,$$

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 ($0.1 < a/\text{AU} < 1000$) (Öpik 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 $0.0001 \leq Z \leq 0.03$. We evolve three million binaries with a total star forming mass of the order of $6.5 \times 10^5 \, M_\odot$ per grid-point.

To optimize 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 $5 \, M_\odot$ (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 percent 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 $\sim 3.1 \times 10^8 \, M_\odot$. It is this star forming mass that we use as our normalization $dM_{\text{SF}}$. Exact binary fractions and hence our normalization is still uncertain (Sana 2017; Moe & Di Stefano 2017). For example, adopting a 50 per cent binary fraction would increase the SFR in single stars, increasing our normalization by 30 per cent.

### 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, Pols & Tout (2000), Hurley, Tout & Pols (2002) to rapidly evolve binaries. Our wind mass-loss rates for stars with temperatures below $12,500 \, K$ are prescribed by Hurley et al. (2000) and references therein. For hot massive stars ($T > 12,500 \, K$) we use the wind mass-loss rates by Vink, de Koter & Lamers (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 overall wind mass-loss rate of $1.5 \times 10^{-4} \, M_\odot \, \text{yr}^{-1}$ (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 on to 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 $d\log (R_\odot)/d\log (m)$ against the response of the stellar radius to mass transfer $d\log (R_\odot)/d\log (m)$ (Paczynski & Sienkiewicz 1972; Hjellming & Webbink 1987; Soberman, Phinney & van den Heuvel 1997). We approximate the radial response of the star depending on its stellar type. The stellar types are defined in Hurley et al. (2000).

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

(ii) Hertzsprung gap (HG): We use $d\log (R_\odot)/d\log (m) = 6.5$ for the 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. (2017).

(iii) 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).

(iv) 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 ultrastripped stars based on Tauris, Langer & Podsiedlowski (2015) and Tauris et al. (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 $\beta$ of the mass lost by the donor. In our model, this mass transfer efficiency $\beta$ depends on the ratio of the thermal time-scales $t_{th}$ of the stars $\beta = \min \left(1, C \times t_{th1}/t_{th2}\right)$, 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 (Paczyński 1976). This is a complex phase and we parametrize 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 $\alpha$ 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 parametrized by $\lambda$ (de Kool 1990). Our choices of $\lambda$ 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-supernovae (SNe) 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 1D 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 ultrastrippped by an NS companion, we lower the 1D root mean square kick 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 $f_\text{fb}$ of mass that falls back on to the newly born compact object is prescribed by Fryer et al. (2012). All of the ejecta falls back ($f_\text{fb} = 1$) for carbon–oxygen core masses above $11$ M$_\odot$. This natal kick is proportionally reduced based on the fallback fraction according to

$$V_{\text{kick}} = (1 - f_\text{fb})V_{\text{kick,drawn}},$$

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 binaries (BHNSs) are left for another study. We also 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 likely to happen for massive stars in binaries as shown by observations (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. 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 overpredicts 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 per cent of all the BBHs form through one of three distinct channels. Here we briefly summarize 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 Fig. 1). We focus
on the systems that evolved at a metallicity of $Z = 0.1 Z_\odot$, 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, Postnov & Prokhorov (1997), Belczynski, Bulik & Kalogera (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 per cent). The majority of these first mass transfer episodes will happen between a post-MS primary star and an MS companion (49.26 per cent). 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 per cent).

  - **First supernova**- The exposed core is both massive enough to collapse into a BH and the binary survives the supernova (2.66 per cent).

  - **Unstable mass transfer**- The secondary star evolves and starts an episode of dynamically unstable mass transfer resulting in a common envelope (0.87 per cent). The system is able to expel the envelope leaving a tighter binary (0.50 per cent).

  - **Second supernova**- The secondary also collapses into a BH and the binary system survives the second supernova (0.43 per cent).

  - **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 per cent 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 per cent).

- **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 per cent).

  - **Second supernova**- The secondary collapses into a BH without disrupting the binary (1.02 per cent).

  - **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 per cent).

  Compared to Stevenson et al. (2017) we changed the radial response of HG donors to mass-loss (see Section 2.3). In combination with our prescription for the angular momentum lost during non-conservative mass transfer on to a compact-object primary (see Section 2.3.1), the mass transfer is on average now stable for mass ratios up to $m_{\text{donor}}/m_{\text{accretor}} = 4.5$. Mass transfer from such donors that are significantly more massive than accretors can substantially harden the binary (van den Heuvel, Portegies Zwart & de Mink 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. 2017; 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, Podsiałowski & Sana (2006). **Unstable mass transfer**. In this scenario both stars evolved beyond the HG before engaging in an episode of mass transfer (1.40 per cent). This mass transfer is dynamically unstable (1.27 per cent) and the binary survives the common-envelope ejection (0.71 per cent). 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 per cent 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 per cent 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 Fig. 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 solar metallicities.

At higher metallicities, higher wind mass-loss rates prevent the growth of the carbon–oxygen core (Belczynski et al. 2010; Spera, Mapelli & Bressan 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 equation 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

---

1In this study, we define the solar metallicity mass fraction as $Z_\odot = 0.0142$ and the solar oxygen abundance as $\log_{10}[O/H]_\odot + 12 = 8.69$ based on Asplund et al. (2009); see appendix A5 for details.
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 $t_{\text{delay}} < 14$ Gyr. The integral under the curve is the yield plotted in Fig. 1. Higher metallicities yield lower total DCO masses, particularly for BBHs.

3.3 Total mass distribution

Fig. 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 and 35 $M_\odot$. More massive BBHs are suppressed by the IMF and wind-driven mass-loss. The most massive BBH 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 supernovae (PISNe) – 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 $Z = 0.5 Z_\odot$ and $Z = 0.2 Z_\odot$, 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 $M_\odot$ in our simulations, although these are relatively rare.

The presence of spikes in BBH masses, particularly in the highest mass bin at $Z = 0.5 Z_\odot$ and $Z = 0.2 Z_\odot$, 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).
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 density function $dP/dZ$:

$$\frac{d^3M_{\text{SFR}}}{dT_dV_c}(z) = \frac{d^2M_{\text{SFR}}}{dT_dV_c}(z) \times \frac{dP}{dZ}(z).$$

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 summarize 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 assumption that yields sufficient degrees of freedom given current observational constraints.

Fig. 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 IMF (Madau & 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:

$$\frac{d^2M_{\text{SFR}}}{dT_dV_c} = a \frac{(1 + z)^b}{1 + [(1 + z)/c]^d} M_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}.$$  

The entire parameter space would be 4D, 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 an 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. 2016) in Appendix A. We show the metallicity distribution at several redshifts from a combination of some of these predictions in Fig. 6. This figure also shows our fiducial model—a lognormal distribution in metallicity

$$\frac{dP}{dZ}(z) = \frac{1}{Z_\sigma \sqrt{2\pi}} e^{-[(\ln Z - \mu)/\sigma]^2},$$

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

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

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

$$Z(z) = Z_0 10^{\alpha z},$$

where $Z_0$ is the mean metallicity at $z = 0$ and the parameter $\alpha$ has negative values, yielding lower mean metallicity at higher redshifts. Therefore the free parameters for the metallicity distribution are $Z_0$, 

Figure 4. Mass ratio distributions for BBHs (blue), BHNSs (mint), and BNSs (red) merging in $t_{\text{delay}} < 14 \text{ Gyr}$ 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 0.1 $Z_\odot$ has a prominent feature around $q \approx 0.65$.

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 normalization. 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, Heavens & Jimenez (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, $\alpha$, and $\sigma$. 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-$\alpha$ galaxies.

5 DCO MERGER DISTRIBUTIONS

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, 2018a,b) for inference on gravitational-wave signals.

5.1 Rate and redshift of cosmic DCOs mergers

Fig. 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 Fig. 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, Holz & Farr 2018; Vitale & Farr 2018). This is consistent with results from other MSSFR models: the variation with the Ma et al. (2016) 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

Fig. 8 shows the normalized total mass distribution of BBH mergers at several redshifts for our preferred model MSSFR model. They are 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 Fig. 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 Fig. 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) yield lower mass tails relative to MSSFR models with higher metallicity (our preferred model) or reduced high-redshift, low-metallicity SFR (the Ma et al. 2016 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 Figs 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. 2018a,b). We show the mass priors assumed by Abbott et al. (2016, 2018b) in Fig. 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 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 equation (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, 2018a). We do not include in our analysis the six 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, Taracchini & Buonanno 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 $P_{\text{det}} (M_{\text{chirp}}, D_L)$.

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 results in a peak rate at a redshift between 0.1 and 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 equation 2 for the additional factors of $dV_z/(1 + z)$). Because mergers involving less-massive NSs cannot be observed as far as BBH mergers, detection rates of BHNSs, and BNS mergers peak at $z \approx 0.03$ and $z \approx 0.015$, respectively. As discussed in Section 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 year. The combined observing time of the first two observing runs is about 166 d: 48 d of coincident data for the first and 118 d for the second observing run (Abbott et al. 2016, 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.

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 here since these are typically better observationally constrained than the total masses.

Mergers of more massive DCOs emit louder gravitational-wave signals 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 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 equation 2 for the additional factors of $dV_z/(1 + z)$). Because mergers involving less-massive NSs cannot be observed as far as BBH mergers, detection rates of BHNSs, and BNS mergers peak at $z \approx 0.03$ and $z \approx 0.015$, respectively. As discussed in Section 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 year. The combined observing time of the first two observing runs is about 166 d: 48 d of coincident data for the first and 118 d for the second observing run (Abbott et al. 2016, 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 one in four years of observing time, suggesting that GW170817 was a fortuitous event if we believe our models. 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.

**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, 2018a). The total log likelihood $L_{\text{tot}}$ is the sum of the log likelihoods of the chirp-mass distribution $L_{\text{chirp}}$ and the rate $L_R$.

<table>
<thead>
<tr>
<th>SFR</th>
<th>Variation MSSFR</th>
<th>MZ</th>
<th>GSMF</th>
<th>Detection rate $(\text{yr}^{-1})$</th>
<th>Likelihoods $(\log_{10})$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BBH</td>
<td>BHNS</td>
</tr>
<tr>
<td>Preferred model</td>
<td></td>
<td></td>
<td></td>
<td>22.15</td>
<td>0.23</td>
</tr>
<tr>
<td>Ma et al.</td>
<td>1</td>
<td></td>
<td></td>
<td>18.43</td>
<td>0.4</td>
</tr>
<tr>
<td>2</td>
<td>94.35</td>
<td>0.5</td>
<td>0.12</td>
<td>-32.42</td>
<td>-8.86</td>
</tr>
<tr>
<td>3</td>
<td>113.92</td>
<td>0.52</td>
<td>0.13</td>
<td>-32.48</td>
<td>-11.9</td>
</tr>
<tr>
<td>Langer et al.</td>
<td>1</td>
<td></td>
<td></td>
<td>247.27</td>
<td>1.28</td>
</tr>
<tr>
<td>2</td>
<td>441.08</td>
<td>1.19</td>
<td>0.22</td>
<td>-32.61</td>
<td>-70.6</td>
</tr>
<tr>
<td>3</td>
<td>492.27</td>
<td>1.25</td>
<td>0.23</td>
<td>-32.77</td>
<td>-80.23</td>
</tr>
<tr>
<td>Langer et al., offset</td>
<td></td>
<td></td>
<td></td>
<td>28.72</td>
<td>0.23</td>
</tr>
<tr>
<td>2</td>
<td>120.3</td>
<td>0.35</td>
<td>0.11</td>
<td>-32.68</td>
<td>-12.93</td>
</tr>
<tr>
<td>3</td>
<td>148.74</td>
<td>0.35</td>
<td>0.11</td>
<td>-32.87</td>
<td>-17.62</td>
</tr>
<tr>
<td>Strolger et al.</td>
<td></td>
<td></td>
<td></td>
<td>32.93</td>
<td>0.52</td>
</tr>
<tr>
<td>Ma et al.</td>
<td>1</td>
<td></td>
<td></td>
<td>203.93</td>
<td>0.6</td>
</tr>
<tr>
<td>2</td>
<td>208.21</td>
<td>0.61</td>
<td>0.14</td>
<td>-32.65</td>
<td>-27.9</td>
</tr>
<tr>
<td>Langer et al.</td>
<td>1</td>
<td></td>
<td></td>
<td>406.39</td>
<td>1.28</td>
</tr>
<tr>
<td>2</td>
<td>659.25</td>
<td>1.19</td>
<td>0.24</td>
<td>-32.98</td>
<td>-111.92</td>
</tr>
<tr>
<td>3</td>
<td>710.91</td>
<td>1.25</td>
<td>0.24</td>
<td>-33.09</td>
<td>-121.79</td>
</tr>
<tr>
<td>Langer et al., offset</td>
<td></td>
<td></td>
<td></td>
<td>89.79</td>
<td>0.33</td>
</tr>
<tr>
<td>2</td>
<td>267.34</td>
<td>0.43</td>
<td>0.12</td>
<td>-33.2</td>
<td>-38.48</td>
</tr>
<tr>
<td>3</td>
<td>292.76</td>
<td>0.43</td>
<td>0.12</td>
<td>-33.22</td>
<td>-43.1</td>
</tr>
</tbody>
</table>
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 overlap between the black and cyan regions indicates that the observations are consistent with the model within the statistical uncertainty of this limited data set, although the CDF of the data does not perfectly match the model prediction.

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 \( L_{\text{tot}} \) is the sum of the rate log likelihood \( L_R \) and the likelihood of the normalized chirp-mass distribution \( L_{M_{\text{chirp}}} \):

\[
L_{\text{tot}} = L_{M_{\text{chirp}}} + L_R.
\]

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 chirp-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 and 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 five-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 parametrization that is useful for inference. We also introduced a particular choice of these five 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 versus 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 an 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 ‘Fingerspitzengefuhl’. 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 and affect the ratio between BBH, BHNS, and BNS detection rates. This is comparable to the impact of uncertainties on evolutionary physics such as wind mass-loss rates, mass transfer efficiency, of common-envelope evolution, and BH natal kicks (Dominik et al. 2012; Giacobbo & Mapelli 2018; Kruckow et al. 2018).

The sensitivity to MSSFR is predominantly driven by the impact of metallicity on the yield of BHs per unit star forming mass. This is consistent with earlier findings (e.g. Dominik et al. 2015; Chruslinska et al. 2019). In particular, Chruslinska et al. (2019) also find that a higher average metallicity is required in order to not overpredict 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. 2019). 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⊙, i.e. with chirp masses between 45 and 115 M⊙ for equal-mass binaries. Abbott et al. (2018b) find that existing gravitational-wave detections show evidence for a maximum black hole mass of around ~45 M⊙, consistent with population synthesis studies such as Belczynski et al. (2016b), Spera & Mapelli (2017), and Stevenson et al. (2019). However, as seen 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 (Langer & Norman 2006; Madau & Dickinson 2014; Ma et al. 2016). 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 reionization (Stanway, Eldridge & Becker 2016), SNe (Chruslinska & Nelemans 2019; Eldridge, Stanway & Tang 2019), and X-ray binaries (Madau & Fragos 2017) can further help to lift the degeneracy between binary physics and the MSSFR.

We introduced a phenomenological description of the MSSFR with five continuous parameters (Section 4) to facilitate the joint exploration of the MSSFR and parametrized evolutionary assumptions. We also proposed a particular choice of the MSSFR model parameters that represents a good match to the BBH gravitational-wave detections made during the first two observing runs of advanced LIGO and Virgo. We use constraints from BBH observations rather than BNS or BHNS because the precision of the latter observational constraints is limited by the small number of detections. Assuming our preferred model, the observed BNS detection GW170817 appears to be fortuitous, however our models are consistent with other observational constraints on BNS merger rates, such as observations of Galactic binary pulsars (Vigna-Gomez et al. 2018). Future detections will enable joint inference on both the MSSFR and the parametrized evolutionary model uncertainties and may shift the preferred model presented in this study. In the meantime, we can use this model in other population studies, such as Stevenson et al. (2019). 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 Fig. 12. We predict 380 BBH detections per year, or approximately one detection per day, within our default evolutionary model.

Our phenomenological model for the redshift-dependent MSSFR is adequate for exploring cosmologically averaged merger rates. Studies of specific galaxy hosts or host types (e.g. in the context of electromagnetic counterpart observations) require more detailed models, such as those considered by Lamberts et al. (2016), Chruslinska & Nelemans (2019), and Boco et al. (2019).

Fig. 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 per cent 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⊙ 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 Fig. 3) could be improved with more efficient importance sampling techniques (Broeckgaarden 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 per cent. Our phenomenological MSSFR model can be incorporated into this hierarchical modelling framework to enable joint inference on binary evolution and the cosmic history of star formation.

ACKNOWLEDGEMENTS

We thank the anonymous referee for helpful suggestions. CIN 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. AVG acknowledges support from Consejo Nacional de Ciencia y Tecnologia (CONACYT). SS is supported by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (OzGrav), through project number CE17010004. SMG is supported by STFC grant ST/M004090/1. SdM acknowledges funding by the European Union’s Horizon 2020 research and innovation programme from the European Research Council (ERC), Grant agreement No. 715063, and by the Netherlands Organisation for Scientific Research (NWO) as part of the Vidi research program BinWaves with project number 639.042.728. DSz accepts funding for Scientific Research (NWO) as part of the Vidi research program Wave Discovery (OzGrav), through project number CE17010004. This paper used the Astropy library (Astropy Collaboration 2013, 2018), matplotlib (Hunter 2007), and numpy (Oliphant 2006).
Effect of MSSFR on DCO mergers

APPENDIX A: METALLICITY-SPECIFIC STAR FORMATION RATE

A1 Cosmological star formation rate

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

\[
d_{21}M_{\text{SFR}} dV_{\nu} = 0.182 \times (t^{2.26} e^{-t/1.865} + 0.071 e^{0.071(t-0.1)/1.865}) M_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}.
\]

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 fit of Savaglio et al. (2005) beyond their upper limit of \( \log_{10}(M_\ast) = 11 \) 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).

Madau & Dickinson (2014):

\[
d_{21}M_{\text{SFR}} dV_{\nu} = 0.015 \frac{(1 + z)^{2.7}}{1 + [(1 + z)/2.9]^{3.6}} M_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}.
\]

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

\[
d_{21}M_{\text{SFR}} dV_{\nu} = 0.015 \frac{(1 + z)^{2.7}}{1 + [(1 + z)/2.9]^{3.6}} M_\odot \text{ yr}^{-1} \text{ Mpc}^{-3}.
\]
In this study we define the solar metallicity mass fraction as $Z$. Conversions to metallicity depend on the assumed solar abundances.

We describe the MZ relations considered in this work in this subsection, and the GSMFs in the next one. This shifts the distribution towards lower galaxy stellar masses, which translates to lower metallicities.

UV continuum slopes, redshift, and dust obscuration witnessed in lower redshift galaxies can be extrapolated to higher redshift observations (Bouwens et al. 2009). The sensitivity to the assumed IMF introduces additional uncertainties (Hopkins & Beacom 2006). These combined uncertainties lead to a broad range of high-redshift SFR estimates (Hopkins & Beacom 2006), between $10^7$ and $10^8 M_\odot \text{Gpc}^{-3} \text{yr}^{-1}$ at $z = 6$ (cf. Fig. 5). Gamma-ray bursts are not sensitive to dust, but the connection between their rate and the underlying SFR is itself model dependent. Yüksel et al. (2008) recover an SFR drop-off at higher redshifts from gamma-ray burst observations that is less steep than the drop-off recovered from UV observations. The various SFR models used in this study are generally consistent within uncertainties in the literature, despite being based on different observational evidence and different assumptions.

### 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 birth. Observations are typically given in terms of the ratio of the number density of oxygen and hydrogen in the gas, generally written as $\log_{10}(\text{O/H}) + 12$. 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 $\log_{10}(\text{O/H})_\odot + 12 = 8.69$ 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 resolution methodologies 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 an 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

$$\log_{10}(\text{O/H}) + 12 = 0.478 \log_{10}\left(\frac{M_*}{M_\odot}\right) + 4.062.$$  \hspace{1cm} (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:

$$M_* = \left(\frac{Z}{Z_\odot}\right)^2.$$  \hspace{1cm} (A4)

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

$$Z = Z_\odot 10^{-0.3z}.$$  \hspace{1cm} (A5)

When we translate this back into an MZ relation we find that there is a 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. (2016). 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 $\leq \log_{10}(M_*/M_\odot) \leq 11$ and redshifts between 0 and 6. Ma et al. (2016) give the MZ relationship as

$$\log_{10}\left(\frac{Z_{\text{gas}}}{Z_\odot}\right) = 0.35 \left[\log_{10}\left(\frac{M_*}{M_\odot}\right) - 10\right] + 0.93 e^{-0.43z} - 1.05.$$  \hspace{1cm} (A6)

### A3 Galaxy stellar mass density function

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, Glazebrook & Driver (2008) show that for galaxies within the mass range of 8.5 $\leq \log_{10}(M_*/M_\odot) \leq 12$ at redshift $z < 0.1$, there is a good agreement on the shape of the GSMF.

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

$$\Phi(M,z) dM = \phi_1(z) \left(\frac{M_*}{M(z)}\right)^{-\alpha(z)} e^{-M(z)/M_*} dM_*,$$  \hspace{1cm} (A7)

where $\alpha$ determines the slope of the GSMF at the low-mass end, $M_*$ is the turnover mass, and $\phi_1$ the overall normalization. However,
extrapolate the Furlong et al. (2015) GSMF over the full range [0, 6.5]. For comparison we also use a single Schechter function of Panter et al. (2004) as used in Langer & Norman (2006) and our previous work (Barrett et al. 2018). The same (pessimistic) evolutionary model is assumed. The white square denotes the point in parameter space we assumed for this study.

\begin{equation}
\Phi_{M_1}(z) dM = e^{-M_{\odot}} M_1^2 \left( \frac{M_1}{M_{\odot}} \right)^{-\alpha_1(z)} \left[ \phi_1(z) \left( \frac{M_1}{M_{\odot}} \right)^{-\alpha_1(z)} + \phi_2(z) \left( \frac{M_1}{M_{\odot}} \right)^{-\alpha_2(z)} \right] dM.
\end{equation}

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 \( \phi_2 \), 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 normalization which in principle carries information on the star formation history, although Furlong et al. (2015) note that the normalization 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 parametrize and test the SFR and the metallicity distribution. Consequently, the normalization coefficients \( \phi \) are relevant only for describing the ratio between the two Schechter functions in equation (A8).

Fig. 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 towards 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.

\textbf{A4 Metallicity-specific star formation rate - MSSFR}

The MZ relation allows us to convert the GSMF into a metallicity distribution \( d\Phi/dZ \) (the last term of equation 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 equation (A8) into a mass density by multiplying by \( M_\star \). The form of this equation makes it possible to carry out the mass integral analytically, with the amount of mass at \( M_\star \leq M_\star \) given through the incomplete gamma functions \( \hat{\Gamma} \):

\begin{equation}
\int_0^{M_\star} M_\star \Phi_{M_\star} dM_\star = \Phi_1 \hat{\Gamma}(\alpha_1 + 2, \frac{M_\star}{M_{\odot}}) + \Phi_2 \hat{\Gamma}(\alpha_2 + 2, \frac{M_\star}{M_{\odot}}).
\end{equation}

The fraction of mass in the range between \( M_\star \leq M_{\star_2} \leq M_\star_1 \) can be obtained from the above equation after normalization with the complete gamma function \( \Gamma \). Fig. 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 × 3 MZ × 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 extrasolar and subsolar metallicities at \( z = 0 \). The GSMF variants include a static redshift-independent fit and two redshift-dependent fits, which evolve towards higher galaxy stellar masses at lower redshifts (see Tables C1, C2).

\textbf{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 \( \log_{10}[O/H]_\odot + 12 = 8.69 \) 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. (2016) assume a mass fraction of \( Z_\odot = 0.02 \) and a specific iron mass fraction of 0.00173 to obtain an oxygen abundance of \( \log_{10}[O/H]_\odot + 12 = 9.0 \). 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. Fig. A3 shows the predicted number of BBH detections per year at the sensitivity of the first observing run over a 2D grid of solar metallicities and oxygen abundances, while keeping
Figure B1. Single stellar tracks of stars with initial masses of 90 $M_{\odot}$, 100 $M_{\odot}$, and 110 $M_{\odot}$. 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 on to the Hertzsprung gap. The two tracks evolve identically from point 1, resulting in the same remnant mass.

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 $L_{\text{tot}}(d|M)$ of observing the data set $d$, which consists of $N_{\text{obs}}$ 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, Farr & Gair 2013)

$$L_{\text{tot}}(d|M) = \prod_{i=1}^{N_{\text{obs}}} P_M(d_i|M)^{d_i}.$$
Here, we focus on the chirp mass $M_c$, 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

$$
\log_{10}(L_{\text{tot}}(d|M)) = \log_{10}(L(N_{\text{obs}}|N_M)) + \sum_{i=1}^{N_{\text{sim}}} \log_{10}(L(M_{c,i}|p_M(M_c))) ,
$$

(C2)

where $M_{c,i}$ is the measured chirp-mass of the $i$th gravitational wave observation and $p_M(M_c)$ is chirp-mass distribution characterizing the MSSFR model $M$. The first term is abbreviated as $L_R$ in Table C2. The second term, $L_{N_{\text{sim}}} = L(M_{c,i}|p_M(M_c))$ is the probability of detecting a chirp mass $M_{c,i}$ 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 $N_{\text{sim}}$ chirp masses produced by the COMPAS simulation as a 1D Gaussian centred on the simulated chirp-mass value $M_i$. All Gaussians have the same bandwidth $\sigma$, determined using the default ‘Scott’s rule’ (Scott 2015) of the Gaussian kernel density estimator in the scipy package (Oliphant 2007, Perez, Granger & Hunter 2011). Each simulated data point $j$ contributes to the overall probability density function proportionally to its observing rate $R_j$, estimated in equation (2). Therefore we re-weight each data point by $R_j$ and normalize by the total rate $R_{\text{tot}}$.

$$
p_M(M_c) = \frac{1}{R_{\text{tot}}} \sum_{j=1}^{N_{\text{sim}}} R_j \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(M_{c,i} - M_c)^2}{2\sigma^2}} .
$$

(C3)

For a single perfect detection, the likelihood of observing a chirp-mass $M_c$ would be given by equation (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, 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 reweight 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 $i$, characterized by the approximated Gaussian posterior of the chirp mass $p_M(M_c)$, given an MSSFR model $M$, becomes

$$
L(M_{c,i}|p_M(M_c)) = \int_0^\infty p_M(M_c) p_M(M_c) \, dM_c.
$$

(C4)

Fig. 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. (2016), and the GSMF of Furlong et al. (2015) (blue); and the convolution between the two (black). The unnormalized integral of this convolution is our estimate of the likelihood $L(M_{c,i}|p_M(M_c))$.

### 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 resample 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.

<table>
<thead>
<tr>
<th>MSSFR Variation</th>
<th>BBH Rates</th>
<th>BHNS Rates</th>
<th>BNS Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFR</td>
<td>O1 det.</td>
<td>O1 det.</td>
<td>O1 det.</td>
</tr>
<tr>
<td>Favour model</td>
<td>$z = 0$</td>
<td>$z = 0$</td>
<td>$z = 0$</td>
</tr>
<tr>
<td>Ma et al.</td>
<td>$Gpc^{-3}$ yr$^{-1}$</td>
<td>$Gpc^{-3}$ yr$^{-1}$</td>
<td>$Gpc^{-3}$ yr$^{-1}$</td>
</tr>
<tr>
<td>Madau et al.</td>
<td>21.80$^{+0.47}_{-0.60}$</td>
<td>32.22$^{+1.7}_{-1.1}$</td>
<td>85.97$^{+2.47}_{-2.41}$</td>
</tr>
<tr>
<td>Langer et al.</td>
<td>113.92$^{+1.12}_{-1.22}$</td>
<td>22.94$^{+1.22}_{-1.12}$</td>
<td>91.42$^{+2.37}_{-2.24}$</td>
</tr>
</tbody>
</table>

Table C1. Table showing the merger and detection rates per DCO type. The columns labelled ‘$z = 0$ merg.’ are the merger rate per year per cubic gigaparsec at zero redshift without selection effects; columns labelled ‘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.
Table C1 – continued

<table>
<thead>
<tr>
<th>SFR</th>
<th>MZ</th>
<th>GSMF</th>
<th>BBH Rates</th>
<th>BHNS Rates</th>
<th>BNS Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Langer et al., offset</td>
<td>1  2  3</td>
<td>10.39 9.43 9.09</td>
<td>28.74 8.40 6.94</td>
<td>18.96 7.10 5.77</td>
<td>0.23 ±0.01 0.35 ±0.01 0.52 ±0.03</td>
</tr>
<tr>
<td>Strolger et al.</td>
<td>Ma et al.</td>
<td>1  2  3</td>
<td>10.98 9.72 9.36</td>
<td>32.94 8.93 6.63</td>
<td>42.11 2.71 1.81</td>
</tr>
<tr>
<td>Langer et al.</td>
<td>1  2  3</td>
<td>9.19 8.99 8.69</td>
<td>59.25 15.27 14.58</td>
<td>95.98 6.85 6.2</td>
<td>1.19 ±0.03 1.06 ±0.03 0.43 ±0.02</td>
</tr>
<tr>
<td>Langer et al., offset</td>
<td>1  2  3</td>
<td>13.52 12.55 12.55</td>
<td>89.70 14.22 14.22</td>
<td>27.88 1.96 1.96</td>
<td>0.33 ±0.02 0.25 ±0.02 0.43 ±0.02</td>
</tr>
<tr>
<td>Preferred model</td>
<td>1  2  3</td>
<td>19.05 18.05 17.05</td>
<td>36.80 16.65 16.65</td>
<td>15.88 13.13 13.13</td>
<td>0.23 ±0.00 0.25 ±0.00 0.33 ±0.00</td>
</tr>
</tbody>
</table>

**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.**, $L_{\text{M}}$ is the total likelihood, $L_{\text{R}}$ is the Poisson likelihood of observing 10 BBH events over 166.6 yr of coincident observation, and $L_{\text{tot}}$ 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 $1 = \text{Panter et al. (2004)}$, $2 = \text{Furlong et al. (2015)}$ (single Schechter function), $3 = \text{Furlong et al. (2015)}$ (double Schechter function). Optimistic and pessimistic variants relate to the ability to eject the common envelope when the donor is a Hertzsprung-gap star.

<table>
<thead>
<tr>
<th>MSSFR variation</th>
<th>Likelihoods ($\log_{10}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFR</td>
<td>MZ</td>
</tr>
<tr>
<td>Pessimistic</td>
<td></td>
</tr>
<tr>
<td>Preferred model</td>
<td>1  2  3</td>
</tr>
<tr>
<td>Madau et al.</td>
<td>1  2  3</td>
</tr>
<tr>
<td>Langer et al.</td>
<td>1  2  3</td>
</tr>
<tr>
<td>Langer et al., offset</td>
<td>1  2  3</td>
</tr>
<tr>
<td>Preferred model</td>
<td>1  2  3</td>
</tr>
<tr>
<td>Optimistic</td>
<td>1  2  3</td>
</tr>
<tr>
<td>Madau et al.</td>
<td>1  2  3</td>
</tr>
<tr>
<td>Langer et al.</td>
<td>1  2  3</td>
</tr>
</tbody>
</table>

### Table C2 – continued

<table>
<thead>
<tr>
<th>MSSFR variation</th>
<th>Likelihoods (log_{10})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mathcal{L}_{Mz}$</td>
</tr>
<tr>
<td>SFR</td>
<td>MZ</td>
</tr>
<tr>
<td>Langer et al.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$-32.24 \pm 0.05$</td>
</tr>
<tr>
<td>2</td>
<td>$-32.61 \pm 0.06$</td>
</tr>
<tr>
<td>3</td>
<td>$-32.77 \pm 0.07$</td>
</tr>
<tr>
<td>Langer et al., offset</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$-32.3 \pm 0.1$</td>
</tr>
<tr>
<td>2</td>
<td>$-32.68 \pm 0.08$</td>
</tr>
<tr>
<td>3</td>
<td>$-32.87 \pm 0.07$</td>
</tr>
<tr>
<td>Strolger et al.</td>
<td></td>
</tr>
<tr>
<td>Ma et al.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$-33.82 \pm 0.11$</td>
</tr>
<tr>
<td>2</td>
<td>$-32.81 \pm 0.08$</td>
</tr>
<tr>
<td>3</td>
<td>$-32.65 \pm 0.08$</td>
</tr>
<tr>
<td>Langer et al.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$-33.44 \pm 0.06$</td>
</tr>
<tr>
<td>2</td>
<td>$-32.98 \pm 0.1$</td>
</tr>
<tr>
<td>3</td>
<td>$-33.09 \pm 0.1$</td>
</tr>
<tr>
<td>Langer et al., offset</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$-32.46 \pm 0.11$</td>
</tr>
<tr>
<td>2</td>
<td>$-33.2 \pm 0.12$</td>
</tr>
<tr>
<td>3</td>
<td>$-33.22 \pm 0.11$</td>
</tr>
</tbody>
</table>

**Preferred model**

<table>
<thead>
<tr>
<th>Preferred model</th>
<th>Likelihoods (log_{10})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Madau et al.</td>
<td></td>
</tr>
<tr>
<td>Ma et al.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$-36.37 \pm 0.1$</td>
</tr>
<tr>
<td>2</td>
<td>$-32.61 \pm 0.05$</td>
</tr>
<tr>
<td>3</td>
<td>$-32.52 \pm 0.03$</td>
</tr>
<tr>
<td>Langer et al.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$-32.55 \pm 0.05$</td>
</tr>
<tr>
<td>2</td>
<td>$-32.55 \pm 0.05$</td>
</tr>
<tr>
<td>3</td>
<td>$-32.67 \pm 0.05$</td>
</tr>
<tr>
<td>Langer et al., offset</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$-32.85 \pm 0.07$</td>
</tr>
<tr>
<td>2</td>
<td>$-32.49 \pm 0.06$</td>
</tr>
<tr>
<td>3</td>
<td>$-32.59 \pm 0.06$</td>
</tr>
<tr>
<td>Strolger et al.</td>
<td></td>
</tr>
<tr>
<td>Ma et al.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$-35.37 \pm 0.11$</td>
</tr>
<tr>
<td>2</td>
<td>$-32.54 \pm 0.08$</td>
</tr>
<tr>
<td>3</td>
<td>$-32.39 \pm 0.05$</td>
</tr>
<tr>
<td>Langer et al.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$-32.38 \pm 0.05$</td>
</tr>
<tr>
<td>2</td>
<td>$-32.78 \pm 0.09$</td>
</tr>
<tr>
<td>3</td>
<td>$-32.86 \pm 0.08$</td>
</tr>
<tr>
<td>Langer et al., offset</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$-32.24 \pm 0.06$</td>
</tr>
<tr>
<td>2</td>
<td>$-32.83 \pm 0.1$</td>
</tr>
<tr>
<td>3</td>
<td>$-32.84 \pm 0.08$</td>
</tr>
</tbody>
</table>

This paper has been typeset from a TeX/LaTeX file prepared by the author.