

Impact of Massive Binary Star and Cosmic Evolution on Gravitational Wave Observations II: Double Compact Object Mergers

Floor S. Broekgaarden,^{1*} Edo Berger,¹ Simon Stevenson,^{2,3} Stephen Justham,^{4,5,6,7} Ilya Mandel,^{3,8,9} Martyna Chruslinska,^{7,10} Lieke A. C. van Son,^{1,6,7} Tom Wagg,^{11,1,7} Alejandro Vigna-Gómez,^{12,13} Selma E. de Mink,^{7,6,1} Debatri Chattopadhyay,^{2,3} Coenraad J. Neijssel⁹

¹Center for Astrophysics | Harvard & Smithsonian, 60 Garden St., Cambridge, MA 02138, USA

²Center for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn VIC 3122, Australia

³The ARC Center of Excellence for Gravitational Wave Discovery – OzGrav, Hawthorn VIC 3122, Australia

⁴School of Astronomy & Space Science, University of the Chinese Academy of Sciences, Beijing 100012, China

⁵National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China

⁶Anton Pannekoek Institute for Astronomy and GRAPPA, University of Amsterdam, Postbus 94249, 1090 GE Amsterdam, The Netherlands

⁷Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, 85741 Garching, Germany

⁸Monash Centre for Astrophysics, School of Physics and Astronomy, Monash University, Clayton, Victoria 3800, Australia

⁹Birmingham Institute for Gravitational Wave Astronomy and School of Physics and Astronomy, University of Birmingham, Birmingham, B15 2TT, United Kingdom

¹⁰Institute of Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen, PO Box 9010, 6500 GL Nijmegen

¹¹Department of Astronomy, University of Washington, Seattle, WA, 98195

¹²Niels Bohr International Academy, The Niels Bohr Institute, Blegdamsvej 17, DK-2100 Copenhagen, Denmark

¹³DARK, Niels Bohr Institute, University of Copenhagen, Jagtvej 128, DK-2200, Copenhagen, Denmark

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

Making the most of the rapidly increasing population of gravitational-wave detections of black hole (BH) and neutron star (NS) mergers requires comparing observations with population synthesis predictions. In this work we investigate the *combined* impact from the key uncertainties in population synthesis modelling: the physical processes in massive binary-star evolution, and the star formation history as a function of metallicity Z and redshift z , $\mathcal{S}(Z, z)$. Considering these uncertainties we create 560 different publicly available model realizations and calculate the rate and distribution characteristics of detectable BHBH, BHNS, and NSNS mergers. We find that our chosen stellar evolution and $\mathcal{S}(Z, z)$ variations can impact the predicted intrinsic and detectable merger rates by factors 10^2 – 10^4 . We find that BHBH rates are dominantly impacted by $\mathcal{S}(Z, z)$ variations, the NSNS rates by stellar evolution variations and BHNS rates by both. We then consider the impact from the combined uncertainty on the detectable mass distribution shapes (chirp mass, individual masses and mass ratio). We find that the BHBHs are dominantly impacted by $\mathcal{S}(Z, z)$ changes, whereas the BHNS mass distributions are predominantly impacted by massive binary-star evolution changes. For NSNS we find that both uncertainties are important. We also find that the shape of the delay time and birth metallicity distributions are typically dominated by the choice of $\mathcal{S}(Z, z)$ for BHBH, BHNS and NSNS. We identify several examples of robust features in the mass distributions predicted by all 560 models, such as that we expect more than 95 percent of BHBH detections to contain a BH $\gtrsim 8 M_{\odot}$ and have mass ratios $q_f \lesssim 4$. Our work, together with future studies that explore uncertainties beyond the variations considered here, will aid understanding what we can learn about the lives and deaths of massive stars throughout cosmic history from gravitational-wave sources.

Key words: (transients:) black hole - neutron star mergers – gravitational waves – stars: evolution

Cool? 

1 INTRODUCTION

The population of detected gravitational-wave (GW) events from binary black hole (BHBH), black hole–neutron star (BHNS) and binary neutron star (NSNS) mergers is rapidly increasing (Abbott et al. 2018, 2019, 2021b,d). These mergers carry unique information about the properties of BHs and NSs (such as their masses and spins), which in turn probes the formation, lives, and explosive deaths

of massive stars throughout cosmic history (e.g. Abbott et al. 2021c). To extract information from these double compact object (DCO) detections requires comparing their observed properties to theoretically simulated populations modelling their formation pathways.

A variety of formation channels for DCO mergers have been proposed (see Mandel & Farmer 2018; Mapelli 2021, for reviews), with one of the most commonly studied pathways being the *isolated binary evolution channel*, where the merging DCO systems are assumed to form from pairs of massive stars in wide, isolated, binary systems. This formation pathway can currently account for the observed DCO

* E-mail: floor.broekgaarden@cfa.harvard.edu

rates (Mandel & Broekgaarden 2021) and many of the observed DCO source properties (e.g., Vigna-Gómez et al. 2018; Belczynski et al. 2020, Broekgaarden & Berger 2021, but see §4.2 for a discussion).

However, modelling theoretical populations of DCO mergers from the isolated binary evolution channel is challenging as the simulations suffer from two key uncertainties: First, various physical processes in massive binary star evolution are uncertain; these include key evolutionary stages such as wind mass loss, stable mass transfer, common envelope (CE) episodes, and supernovae (SNe), which significantly impact the predicted rates and properties of DCO mergers. Second, critical uncertainties in the cosmic star formation and chemical evolution history that impact the metallicity-specific star formation rate density ($\mathcal{S}(Z, z)$), which is a function of birth metallicity (Z) and redshift (z). Uncertainties in the $\mathcal{S}(Z, z)$ also significantly impact the predicted rate and properties of DCO mergers as the birth metallicity significantly affects stellar evolution, including mass loss through line-driven winds and the (maximum) radial expansion.

Earlier works investigating the impact from stellar evolution and $\mathcal{S}(Z, z)$ uncertainties on the simulated DCO population typically focused on exploring only one of these two sets of uncertainties. For example, studies including Dominik et al. (2015), Giacobbo & Mapelli (2018), Kruckow et al. (2018), and Belczynski et al. (2020) explored population synthesis models with a large number of different massive binary stellar evolution assumptions, but only investigate one or a few $\mathcal{S}(Z, z)$ models. On the other hand, studies including Chruslinska et al. (2019), Neijssel et al. (2019), and Tang et al. (2020) focused on exploring a wide range of $\mathcal{S}(Z, z)$ models, but only considered one or a few different stellar evolution realizations. By only focusing on one of the two uncertainties it remains challenging to directly understand the combined or relative impact from the massive (binary) stellar evolution and $\mathcal{S}(Z, z)$ uncertainties on the detectable DCO population. This limits our capability to learn from GW observations and constrain the models.

Recently, Santoliquido et al. (2021) improved on this by presenting a large set of binary star evolution model variations (for mass transfer, CE phases and SNe) as well as $\mathcal{S}(Z, z)$ models, exploring the impact from each of the uncertainties on the predicted DCO merger *rates* (although the authors do not present the *combined* impact from both uncertainties except for their α_1 and α_5 model variation in their Figure 7). In addition, Chu et al. (2021) carried out a large study investigating the impact from variations in the CE phase and SN kicks in combination with four different $\mathcal{S}(Z, z)$ models focusing on their impact on the merger rates of NSNS mergers. However, besides impacting the merger rates, variations in stellar evolution models and $\mathcal{S}(Z, z)$ are also expected to impact the DCO merger *distribution shapes* of the properties of the detectable mergers. A large systematic study exploring this for GW events from all three types of DCO mergers (BHBH, BHNS and NSNS) is currently missing.

In Broekgaarden et al. (2021, from  paper I), we improved on this by investigating the  impact from uncertainties in both massive binary star evolution and the $\mathcal{S}(Z, z)$ on the predicted *rate* and *distribution shapes* focusing on, as a first step, the GW-detectable BHNS mergers. Here we continue this work by studying all three GW-detectable DCO merger types and by adding extra model realizations. We use rapid binary stellar evolution synthesis simulations coupled with analytical prescriptions for the $\mathcal{S}(Z, z)$ to present the DCO rate and properties for a total of 560 model realizations. Our method is described in §2. We present the impact from the physical processes in massive binary star evolution and the $\mathcal{S}(Z, z)$ on the predicted rate and shape of the distribution functions for DCO mergers in §3. We discuss these results in §4 and present our conclusions in §5.

2 METHOD

We use the simulations and methodology presented in paper I. Here we add new models F, G, J, S and T, which also leads to a minor shift in the model labels. The key assumptions in our simulations explore both massive (binary) star evolution and $\mathcal{S}(Z, z)$, leading to 560 (20 binary stellar evolution \times 28 $\mathcal{S}(Z, z)$) model realizations. We particularly choose our model variations to explore a broad span of uncertainty in the modelling. For that reason some of the variations are somewhat extreme (e.g., models S and T), but are chosen to explore their possible impact on the rate and distribution shapes of DCO mergers. We summarize our most important model assumptions below and summarize the model variations in Table 1 (massive binary star evolution) and Table A1 ($\mathcal{S}(Z, z)$). More details can be found in the Appendix and in paper I.

2.1 Massive binary-star population models

We simulate populations of GW sources with the rapid binary population synthesis code from the COMPAS¹ suite (Stevenson et al. 2017; Barrett et al. 2018; Vigna-Gómez et al. 2018; Broekgaarden et al. 2019; Neijssel et al. 2019; Team COMPAS: J. Riley et al. 2021). We model the formation of BH and NS mergers from the *isolated binary evolution* channel where the merging DCOs form from massive stars born in a binary system (Smarr & Blandford 1976; Srinivasan 1989). We explore uncertainties in our massive (binary) star evolution assumptions by studying 20 binary population synthesis model variations. We further use the efficient sampling algorithm STROOPWAFEL (Broekgaarden et al. 2021) to simulate $\gtrsim 10^6$ binary systems for 53 different Z values, resulting in typically $\sim 10^7$ DCO mergers in each simulation, making it one of the best-sampled population synthesis studies of its kind.

For our default model (A) settings we use the fiducial model summarized in paper I. For the remaining models we change one population parameter at a time (relative to the fiducial model) to explore the impact of the uncertain model assumptions. The only exception is model F (E+K), which combines the model changes from both models E and K. Our 20 models are summarized in Table 1 and we refer to them throughout the rest of the paper by the letters A to T and the abbreviated label names given in the second column of Table 1.

In models B, C, D, E and F we explore uncertainties in binary mass transfer prescriptions. Of these, models B, C and D vary the mass transfer efficiency assumption. This is defined in COMPAS by the parameter β , which determines for a given donated mass rate the fraction that is accreted by the companion star: $\beta \equiv (dM_{\text{acc}}/dt)/(dM_{\text{donor}}/dt)$, with M_{acc} and M_{donor} being the mass of the accretor and donor stars, respectively, and t the time. The excess mass is assumed to leave the binary from the vicinity of the accreting star through ‘isotropic re-emission’ (e.g., Massevitch & Yungelson 1975; Bhattacharya & van den Heuvel 1991; Soberman et al. 1997) and the angular momentum loss is calculated accordingly (cf., Equations 32 and 33 in Belczynski et al. 2008). Our fiducial model (A) assumes the accretion rate is limited by the star’s thermal timescale: $dM_{\text{acc}}/dt = 10M_{\text{acc}}/\tau_{\text{KH}}$, where τ_{KH} is the Kelvin-Helmholtz (thermal) timescale², and the factor of 10 is added to take into account the expansion of the accretor due to mass transfer

¹ Compact Object Mergers: Population Astrophysics and Statistics, <https://compas.science>

² Given in Equation 61 of Hurley et al. (2002), where we use a pre-factor of 30 Myr from Equation 2 in Kalogera & Webbink (1996).

μ	Label	Variation
A	fiducial	–
B	$\beta = 0.25$	fixed mass transfer efficiency of $\beta = 0.25$
C	$\beta = 0.5$	fixed mass transfer efficiency of $\beta = 0.5$
D	$\beta = 0.75$	fixed mass transfer efficiency of $\beta = 0.75$
E	unstable case BB	case BB mass transfer is always unstable
F	E + K	case BB mass transfer is always unstable & HG donor stars initiating a CE may survive
G	$\alpha = 0.1$	CE efficiency parameter $\alpha = 0.1$
H	$\alpha = 0.5$	CE efficiency parameter $\alpha = 0.5$
I	$\alpha = 2$	CE efficiency parameter $\alpha = 2$
J	$\alpha = 10$	CE efficiency parameter $\alpha = 10$
K	optimistic CE	HG donor stars initiating a CE may survive
L	rapid SN	Fryer rapid SN remnant mass model
M	$m_{\text{NS}} = 2 M_{\odot}$	maximum NS mass is fixed to $2 M_{\odot}$
N	$m_{\text{NS}} = 3 M_{\odot}$	maximum NS mass is fixed to $3 M_{\odot}$
O	no PISN	no PISN and pulsational-PISN
P	$\sigma_{\text{rms}}^{1D} = 100 \text{ km s}^{-1}$	$\sigma_{\text{rms}}^{1D} = 100 \text{ km s}^{-1}$ for core-collapse SNe
Q	$\sigma_{\text{rms}}^{1D} = 30 \text{ km s}^{-1}$	$\sigma_{\text{rms}}^{1D} = 30 \text{ km s}^{-1}$ for core-collapse SNe
R	$v_{k,BH} = 0$	we assume BHs receive no natal kick
S	$f_{\text{WR}} = 0.1$	Wolf-Rayet wind factor $f_{\text{WR}} = 0.1$
T	$f_{\text{WR}} = 5$	Wolf-Rayet wind factor $f_{\text{WR}} = 5$

Table 1. List of the 20 binary population synthesis models studied in this work. μ and ‘Label’ denote the alphabetical letter and abbreviated name used to label each model, ‘Variation’ denotes what we changed compared to the fiducial settings (model A). Models B, C, D, E and F vary mass transfer assumptions, models G, H, I, J and K vary common-envelope assumptions, models L, M, N, O, P, Q and R vary supernova assumptions and models S and T vary our Wolf-Rayet wind assumptions. Acronyms used are: common-envelope (CE), Hertzsprung Gap (HG), supernova (SN) neutron star (NS), black hole (BH) and pair-instability SN (PISN). We additionally use the subscripts root-mean-square (rms), one-dimensional (1D) and Wolf-Rayet (WR). Each model varies one assumption compared to the fiducial model, except for model F where we vary two assumptions (namely those from model E and K; unstable case BB and optimistic CE)

(cf., Paczyński & Sienkiewicz 1972, Hurley et al. 2002 and Schneider et al. 2015). Models B, C and D assume a different accretion rate limit for stars by fixing β to 0.25, 0.5, and 0.75, respectively. All of our models assume an Eddington-limited accretion rate for compact objects. In our fiducial model we assume that case BB mass transfer, which is mass transfer from a stripped post-helium-burning star onto the accretor (Delgado & Thomas 1981), is always stable. Model E explores a variation where case BB mass transfer is assumed to always be unstable, leading to a CE phase. This assumption causes case BB mass transferring systems to merge as stars in model E as our fiducial model assumes the now unstable CE phase initiated by a (helium) Hertzsprung Gap star is unsuccessful and leads to a merger (as described further in the ‘pessimistic CE assumption’ in the next paragraph). We therefore also add model F, which explores the effect of assuming unstable case BB mass transfer, but allowing Hertzsprung Gap donor stars to survive a CE phase (the ‘optimistic CE assumption’).

In models G, H, I, J and K we explore the effect of changing the CE prescription assumptions, which in COMPAS are parameterised using the ‘ $\alpha-\lambda$ ’ formalism from Webbink (1984) and de Kool (1990). Our fiducial model assumes a CE efficiency of $\alpha_{\text{CE}} = 1$ and uses for λ the “Nanjing lambda” prescription (cf., Dominik et al. 2012),

which is based on models from Xu & Li (2010a,b). In models G, H, I and J we change the CE efficiency to fixed values of 0.1, 0.5, 2 and 10, respectively. Compared to the fiducial model, lower and higher values of α_{CE} reduce and increase the efficiency with which the CE is ejected, respectively. In model K (and model F) we allow Hertzsprung gap stars that initiate a CE to possibly survive the CE (also known as the ‘optimistic’ CE assumption), whereas in our fiducial model these are assumed to always lead to an unsuccessful CE ejection (and merger), the ‘pessimistic’ CE scenario (cf., Dominik et al. 2012).

In models L, M, N, O, P, Q and R we vary the SN prescription assumptions. In model L we use the ‘rapid’ SN remnant mass model of Fryer et al. (2012) instead of their ‘delayed’ model, which is implemented in our fiducial model. The rapid model creates a mass gap at $\approx 2 M_{\odot}$, where no BHs form, whereas in the delayed model such BHs can form. Models M and N change our assumption for the maximum NS mass, by default $2.5 M_{\odot}$, to $2 M_{\odot}$ and $3 M_{\odot}$, respectively and we adopt the remnant mass prescription from Fryer et al. (2012) accordingly. In model O we do not use the prescription for pair-instability SNe and pulsational pair-instability SNe, therefore allowing the formation of BHs in the mass range of $\approx 40-100 M_{\odot}$. In models P and Q we change the root mean square velocity dispersion (σ_{rms}^{1D}) for the Maxwellian SN natal kick distribution for both BHs and NSs, to 100 km s^{-1} and 30 km s^{-1} , respectively (the fiducial model uses $\sigma_{\text{rms}}^{1D} = 265 \text{ km s}^{-1}$). For all our models we assume that a fraction of the ejected material (f_{fb}) falls back onto the compact object and we adjust the remnant mass and re-scale the SN kick magnitude accordingly (cf., Fryer et al. 2012). For ultra-stripped SN and electron-capture SNe we always draw the SN kick using a one-dimensional root-mean-square velocity dispersion of $\sigma_{\text{rms}}^{1D} = 30 \text{ km s}^{-1}$ following Pfahl et al. (2002) and Podsiadlowski et al. (2004). In model R we assume instead that only all BHs receive zero natal kicks, $v_k = 0 \text{ km s}^{-1}$.

Finally, models S and T explore the assumptions for the mass loss rate in Wolf-Rayet winds. Our Wolf-Rayet wind prescription follows Belczynski et al. (2010), which is based on Hamann & Koesterke (1998) and Vink & de Koter (2005), by parameterizing the wind strength with a multiplicative parameter f_{WR} (cf., Barrett et al. 2018). By default we use $f_{\text{WR}} = 1$ and in models S and T we vary this to 0.1 and 5, respectively, which largely spans the possible range for Wolf-Rayet winds inferred from observations (e.g., Vink 2017; Hamann et al. 2019; Shenar et al. 2019; Sander & Vink 2020).

2.2 Metallicity-specific star formation rate density models

The time between formation and merger of a DCO can range up to many Gyr (e.g., Neijssel et al. 2019). As a result, the DCOs that are detected by current ground-based GW observatories can originate from stars that formed throughout a wide range of redshifts with a large variety of birth metallicities (e.g., Chruslinska & Nelemans 2019). Similar to paper I, we follow the method of Neijssel et al. (2019) to create a metallicity-specific star formation rate density $\mathcal{S}(Z, z)$, which describes the star formation history as a function of birth (initial) metallicity³ Z and redshift z . To explore uncertainties in $\mathcal{S}(Z, z)$, we use a total of 28 $\mathcal{S}(Z, z)$ models. Each model convolves a star formation rate density (SFRD)⁴ with a metallicity probabil-

³ Where Z is the fractional metallicity, such that $X + Y + Z = 1$ with X and Y the mass fractions of hydrogen and helium, respectively.

⁴ We use $\mathcal{S}(Z, z)$ for the metallicity-specific star formation rate density as a function of metallicity and redshift and SFRD for the total star formation rate density across all metallicities, which we model as only a function of redshift.

ity distribution function, dP/dZ . Our first model $xyz = 000$ (see Table A1 for the labeling) is the so-called ‘preferred’ phenomenological model from Neijssel et al. (2019), which they fit to match the BHBH observations from the first two observing runs by LIGO and Virgo. The other 27 ($3 \times 3 \times 3$) models are constructed by combining a SFRD with a galaxy stellar mass function, the number density of galaxies per logarithmic mass bin, (GSMF) and a mass-metallicity relation (MZR). The three variations for each SFRD, GSMF and MZR that we consider are given in Table A1. Although many of the prescriptions in these models oversimplify the complex $\mathcal{S}(Z, z)$ that is evident from observations, these prescriptions provide a convenient parametrization for binary population synthesis studies. For more details about caveats and uncertainties in the modelling of $\mathcal{S}(Z, z)$ within binary population synthesis we refer the reader to Chruslinska & Nelemans (2019), Neijssel et al. (2019), Boco et al. (2021) and references therein.

3 RESULTS

3.1 Metallicity specific merger yields

In Figure 1 we show the calculated formation yields ($\mathcal{R}_{\text{form}}$) for BHBH, BHNS and NSNS mergers per unit solar mass of stars formed (dM_{SFR}) as a function of Z for our 20 binary population synthesis models. This formation yield only includes DCO systems that have a GW driven inspiral time (t_{inspiral}) that is smaller than the Hubble time (≈ 14 Gyr) and is given by

$$\mathcal{R}_{\text{form}}(Z) = \frac{dN_{\text{form}}}{dM_{\text{SFR}}}(Z), \quad (1)$$

where we convert the number of mergers in our COMPAS realizations to formation yields by integrating our simulation range over the full initial parameter space of binary systems (e.g., the full initial mass range of stars) and assuming a corrected⁵ binary fraction of $f_{\text{bin}} = 1$ (consistent with e.g., Sana 2017). Details are given in paper I.

3.1.1 Impact from the binary massive star evolution assumptions and metallicity

Figure 1 demonstrates that the variation in the BHBH merger formation yield is typically dominated by Z , causing variations on the order of $O(10^2)$ between $Z \approx 0.001$ and $Z \approx Z_{\odot}$. Variations in our massive binary population synthesis assumptions impact the BHBH merger yield only with factors of up to ≈ 3 at $Z \lesssim 0.1Z_{\odot}$ and on the order of $O(10)$ for most stellar evolution models at higher Z . The exceptions to this include model T ($f_{\text{WR}} = 5$), which has a much lower BHBH yield (by $O(10^2)$) compared to other stellar evolution models for $Z \gtrsim 0.2Z_{\odot}$ where BH formation is drastically reduced due to the strong stellar wind mass loss. Other examples include the optimistic CE models (F and K) that have higher BHBH yields compared to other stellar evolution models for $Z \gtrsim 0.2Z_{\odot}$.

The scatter in the BHNS and NSNS formation yields, on the other hand, are typically dominated by variations in the stellar and binary evolution models over most of the Z range, leading to variations of $\gtrsim O(10^2)$ in the rates, compared to typical ranges of $\lesssim O(10)$ from variations in Z for most models.

⁵ Corrected to take into account the differences in parameter range between population synthesis and observations, see de Mink & Belczynski (2015).

3.1.2 Trends in the metallicity-specific formation yields

As can be seen in Figure 1, the formation yield of BHBH mergers is constant at low metallicities and then steeply declines at $Z \gtrsim 0.1Z_{\odot}$ for all model variations. This steep decline of merging BHBHs at higher metallicities has been pointed out previously (e.g., Belczynski et al. 2010; Dominik et al. 2012), and is consistent with various recent works including Giacobbo et al. (2018), Giacobbo & Mapelli (2018), Klencki et al. (2018), Chruslinska & Nelemans (2019), Neijssel et al. (2019) and Santoliquido et al. (2021), demonstrating that this trend is present among many different population synthesis studies and simulation settings. This decline is due to several metallicity dependent effects. First, stronger stellar winds at higher metallicities lead to increased mass loss and hence wider binaries with less massive DCOs, reducing the fraction of BHBHs that will merge in a Hubble time. In addition, the less massive BHs have smaller amounts of fallback and hence receive larger natal kicks in our simulations, increasing the fraction that is disrupted during the SN (cf., Fryer et al. 2012; Klencki et al. 2018; Renzo et al. 2019). Indeed, in the model variations where BHs receive lower or no SN kick (Q and R) or the Wolf-Rayet winds are weaker (model S) the decline in BHBH yield is less severe. Second, there is a metallicity dependence of the radius expansion of stars in our single star evolution tracks (Hurley et al. 2000) leading more stars to enter the CE phase during the Hertzsprung Gap instead of during the giant phase at higher (\sim Solar) metallicities (cf., Figure 2 in Belczynski et al. 2010, Figure 7 in Linden et al. 2010 and Section 3.1.1 in Bavera et al. 2021). In our fiducial simulation such Hertzsprung Gap donor stars do not survive the CE phase. In models F (E+K) and K (optimistic CE) we instead assume these systems do survive, leading to a visible boost in BBH yields at higher Z , as seen in Figure 1.

The behavior of the BHNS yield in most models, on the other hand, first modestly increases as a function of Z and then declines around Z_{\odot} (Figure 1; Klencki et al. cf., 2018; Chruslinska & Nelemans cf., 2019; Román-Garza et al. cf., 2021). This causes the yield to broadly peak in the range $0.2Z_{\odot} \lesssim Z \lesssim Z_{\odot}$, where the location of the peak varies between models as shown in Figure 1. The same metallicity dependent effects as described for BHBH mergers above are at play for BHNS mergers. However, since BHNS systems typically already form DCOs with lower mass BHs and form from lower mass stars compared to BHBH mergers, these systems do not experience the same boost in formation yields at low metallicities compared to the BHBH mergers, leading to lower yields at lower Z . In addition, the peak in BHNS yield is caused by the maximum radial extent of stars in the Hertzsprung Gap, which generally increases with Z but decreases between $10^{-3} \lesssim Z \lesssim 10^{-2}$ in our simulations for the zero-age main sequence mass range of BHNS progenitors (initial primary masses of $15 \lesssim m_{1,i}/M_{\odot} \lesssim 50$). This is based on the single star evolution prescriptions from Hurley et al. (2000) implemented in COMPAS. As a result of the smaller radii there are fewer stellar mergers (cf., Giacobbo & Mapelli 2018) and the average initial separations of binaries that form merging BHNSs shifts to smaller values, which are more commonly formed. See also the discussion in §4.1.2 in Klencki et al. (2018) for additional details.

We find that the NSNS merger yield remains roughly constant as a function of Z (cf., Chruslinska et al. 2018; Klencki et al. 2018; Giacobbo & Mapelli 2018; Neijssel et al. 2019) as the metallicity-dependent effects described above are less significant for lower-mass stars that form NSNS binaries. The strong increase in formation yield around $Z = 0.0105$ in Figure 1 that is particularly noticeable for NSNS bins is due to a bifurcation in our λ prescription for the envelope binding energy (relevant for the CE phase) and is an

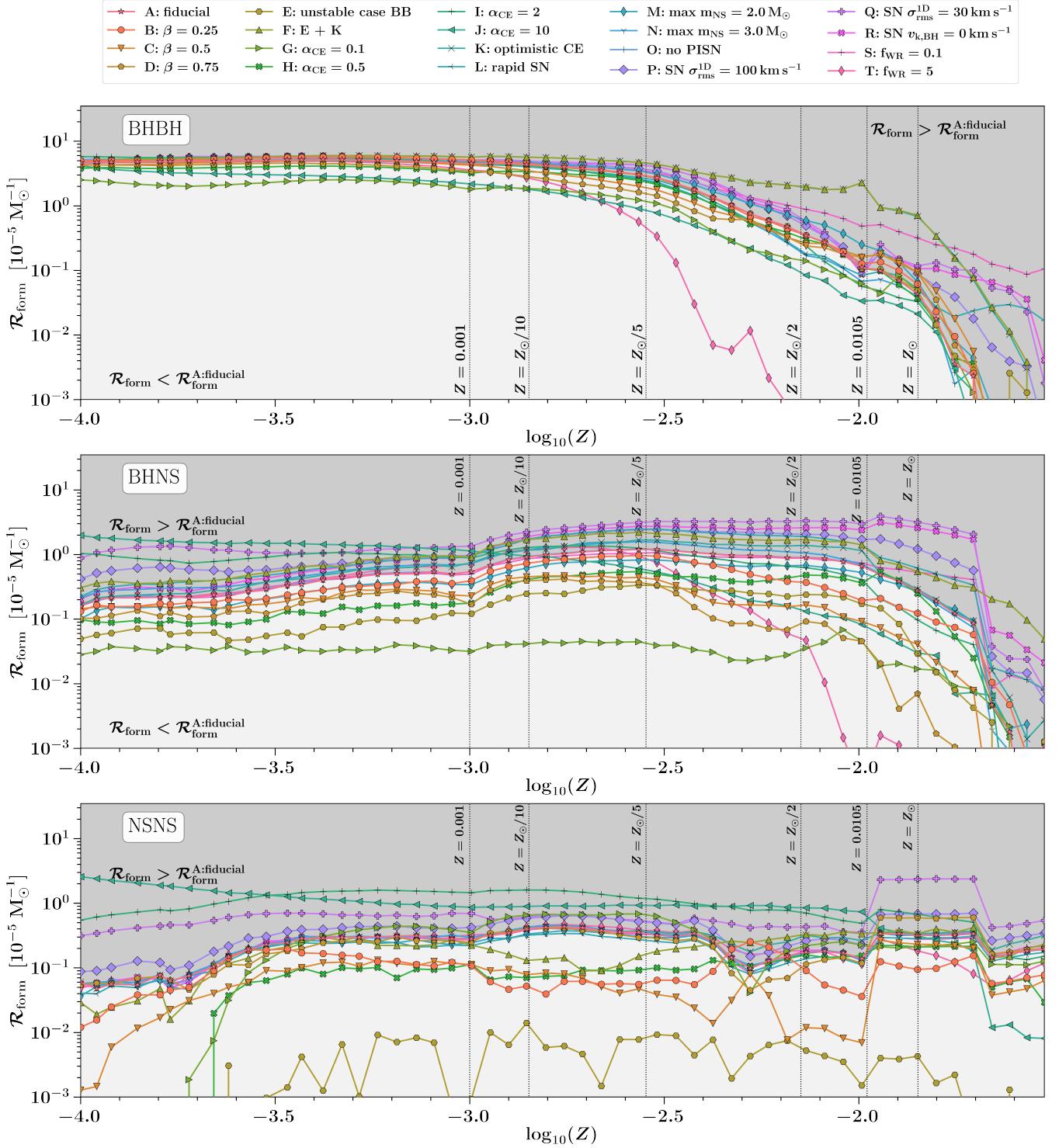


Figure 1. Formation rate of merging double compact objects per solar mass of star formation, $\mathcal{R}_{\text{form}} = dN_{\text{form}}/dM_{\text{SFR}}$ (Equation 1), as a function of birth metallicity Z . The yield only includes binaries that have GW-driven inspiral times $\lesssim 14$ Gyr. From top to bottom the panels show the formation yields for merging BHBH, BHNS and NSNS, respectively. Each color and marker type correspond to one of the 20 binary population synthesis models explored in this study (Table 1). The dark and light gray areas in the background mark where the formation yield is larger and smaller compared to our fiducial yield (A). Vertical dotted lines show fixed Z values to guide the reader. The marker points show the Z grid values that we simulated with COMPAS in each simulation. The sharp increase in yield around $Z \approx 0.0105$, particularly visible in the NSNS panel, is caused by an artificial bifurcation in our λ values for the common-envelope treatment in COMPAS at this Z . The scatter in the unstable case BB mass transfer model (E) in the NSNS panel is caused by sampling noise.

artifact in our modelling. More orbital energy is needed to successfully unbind the envelope for $Z > 0.0105$ in our models, leading to fewer NSNS binaries that survive the second SN and can merge in a Hubble time, thereby increasing the rate. Finally, we note that the scatter in model T (unstable case BB) is caused by sampling noise as a consequence of the very small number of NSNS mergers (only 371) in this simulation⁶. In the next Section we discuss this model and show that it under-predicts the inferred NSNS rate from GW observations by two orders of magnitude.

Comparing the panels in Figure 1 we find that at low Z the formation yield of merging BHBHs in a Hubble time exceeds that of merging BHNSs and NSNSs. The merging BHNS formation yield starts dominating over the BHBH yield at $Z \gtrsim 0.3 Z_{\odot}$, whereas for merging NSNS this occurs at $Z \gtrsim 0.5 Z_{\odot}$ for most models. We show these merging DCO ratios in Figure A1. Using our $\mathcal{S}(Z, z)$ models we can translate these transition metallicities into approximate typical redshifts. For example, Figure A2 shows that the average metallicity of star formation is $\langle Z \rangle \approx 0.5 Z_{\odot}$ at $z \approx 3$ and $z \approx 0.4$ for the $\mu = 312$ and $xyz = 231$ $\mathcal{S}(Z, z)$ models, respectively. These $\mathcal{S}(Z, z)$ models correspond to the highest and lowest typical metallicity star formation within our simulated $\mathcal{S}(Z, z)$ models and, as we will show in the next section, to one of the lowest and highest DCO merger rate density predictions, respectively.

3.2 Intrinsic merger rates

In Figure 2 we show the population synthesis calculated local (intrinsic) merger rate densities for BHBH, BHNS and NSNS binaries. There are three main findings that we describe in more detail below. First, we find that the combined variations in our massive binary star and $\mathcal{S}(Z, z)$ assumptions impact the intrinsic DCO rates with resulting ranges of ≈ 210 –1030. Second, only a subset of the 560 models matches the ranges for the inferred BHBH, BHNS and NSNS merger rate density from GW observations (Abbott et al. 2021a,c,d). Third, we find that the calculated merger rates for different types of DCO are sensitive to different uncertainties in the modelling. Most strikingly, we find that the BHBH merger rates are primarily sensitive to variations in $\mathcal{S}(Z, z)$ models, whereas the NSNS merger rates are predominantly impacted by binary stellar evolution model variations. This means that the observed BHBH and NSNS rates can provide a test-bed for $\mathcal{S}(Z, z)$ and stellar evolution uncertainties, respectively. More details are provided in the DCO specific subsections below.

The intrinsic rates are calculated using the formation yields from Equation 1 and by taking into account the $\mathcal{S}(Z, z)$ weighting discussed in §2.2. The local merger rate is then obtained using $t_m = 0$ in the equation

$$\mathcal{R}_m(t_m) = \frac{d^2 N_{\text{merger}}}{dt_s dV_c} (t_m) = \int dZ \int_0^{t_m} dt_{\text{delay}} \mathcal{S}(Z, t_{\text{form}}) \mathcal{R}_{\text{form}}(Z, t_{\text{delay}}), \quad (2)$$

where t_s is time in the source frame measured from the Big Bang⁷, t_m is the time of the DCO merger, t_{delay} is the delay time between the formation of the DCO and its merger, and V_c is the co-moving volume. The delay time is $t_{\text{delay}} = t_{\text{evolve}} + t_{\text{inspiral}}$, the total time from the onset of hydrogen burning at ZAMS to forming a DCO until the second SN (t_{evolve}), and the time it takes the DCO to

⁶ See this online table.

⁷ Where we use the WMAP9–cosmology from Astropy (Hinshaw et al. 2013).

coalesce from the moment of the second SN (t_{inspiral}); see Figure 1 in paper I for more details. The binary initially forms at t_{form} , which we set as $t_{\text{form}} = t_m - t_{\text{delay}}$ in this equation. For the $\mathcal{S}(Z, z)$ we use the 28 models from Table A1. In practice we estimate Equation 2 with a Riemann sum where we sum over our metallicity grid and redshift (time) bins.

To quantify the scatter in the predicted intrinsic rates we calculate the mean of the ratios between the maximum and minimum predicted rates given by

$$\langle \sigma_{\mu} \rangle = \frac{1}{28} \sum_{xyz=000}^{xyz=333} \frac{\max(\mathcal{R}_{m,Axyz}^0, \dots, \mathcal{R}_{m,Txyz}^0)}{\min(\mathcal{R}_{m,Axyz}^0, \dots, \mathcal{R}_{m,Txyz}^0)}, \quad (3)$$

and

$$\langle \sigma_{xyz} \rangle = \frac{1}{20} \sum_{\mu=A}^{\mu=T} \frac{\max(\mathcal{R}_{m,\mu 000}^0, \dots, \mathcal{R}_{m,\mu 333}^0)}{\min(\mathcal{R}_{m,\mu 000}^0, \dots, \mathcal{R}_{m,\mu 333}^0)}, \quad (4)$$

where we used the short-hand notation $\mathcal{R}_m^0 = \mathcal{R}_m(t_m(z \approx 0))$. Intuitively, Equation 3 represents the uncertainty range from stellar evolution as averaged over the $\mathcal{S}(Z, z)$ models, whereas Equation 4 represents the uncertainty range from $\mathcal{S}(Z, z)$ as averaged over the stellar evolution models. Large (small) values for σ_{μ} and σ_{xyz} correspond to large (small) impacts by binary star evolution and $\mathcal{S}(Z, z)$ variations, respectively.

3.2.1 BHBH merger rates

We find that the intrinsic merging BHBH rates are predicted to lie in the range $\mathcal{R}_m^{0,\text{BHBH}} \approx 3.8$ –810 Gpc $^{-3}$ yr $^{-1}$ as shown in the top panel in Figure 2. The combined uncertainty thus impacts the predicted rates with a factor of up to⁸ ≈ 212 . We find that the variations in stellar and binary evolution impact the BHBH rate on average for the $\mathcal{S}(Z, z)$ models with $\langle \sigma_{\mu} \rangle \approx 6$, while variations in our $\mathcal{S}(Z, z)$ models impact the predicted BHBH rate with uncertainties of $\langle \sigma_{xyz} \rangle \approx 20$. The $\mathcal{S}(Z, z)$ is in our explored model variations is thus dominant in impacting the uncertainty in the BHBH rate (cf. Santoliquido et al. 2021), which results from the formation yield of BHBHs being dominantly sensitive to metallicity (Figure 1).

We find that the majority of our stellar and binary evolution model variations do not impact the BHBH rate with more than a factor ≈ 2 , as can be seen in the top panel of Figure 2 when comparing the predicted rates with our fiducial model. Model T ($f_{\text{WR}} = 5$) has the highest impact on the rates. The other largest BHBH rate changes are from stellar evolution models that change the CE assumptions (e.g., models F, G and K).

The 90% credible intervals from the GW inferred rate of merging BHBHs is $\mathcal{R}_m^{\text{BHBH}} = 10.3$ –104 Gpc $^{-3}$ yr $^{-1}$ based on GW detections as reported in the GWTC-2 and GWTC-2.1 catalogues (Abbott et al. 2021a,c)⁹. Figure 2 shows that only a subset of our $\mathcal{S}(Z, z)$ models (lowest lines in the top panel) predict merging BHBH rates that are consistent with this range. The $\mathcal{S}(Z, z)$ realizations that typically

⁸ This is the ratio between the minimum and maximum predicted rates quoted with the error bar in the left of each panel in Figure 2. This is not the same as multiplying $\langle \sigma_{xyz} \rangle$ and $\langle \sigma_{\mu} \rangle$, as the latter are averages calculated using Equations 3 and 4, and because the impacts from stellar evolution and $\mathcal{S}(Z, z)$ variations are not (fully) independent.

⁹ This range spans the minimum and maximum over the several 90% credible intervals given in Abbott et al. (2021a) and Abbott et al. (2021c). We compile the observed rates for the BHNS and NSNS in a similar way.

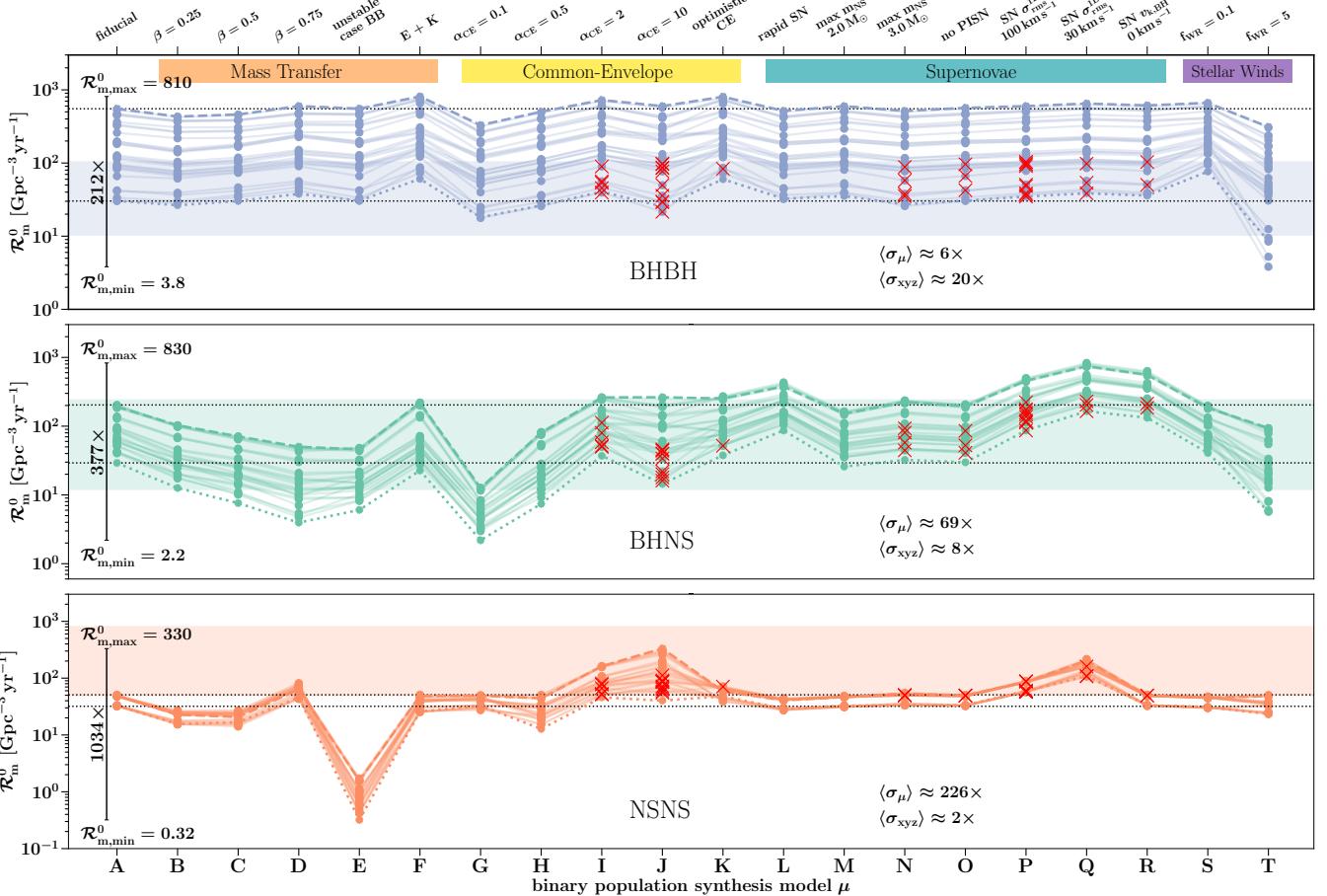


Figure 2. Intrinsic (local) merger rate densities for merging BHBHs (top panel), BHNSs (middle panel) and NSNSs (bottom panel) systems for our 560 model variations. The rates are for merging DCOs at $z \approx 0$ calculated using Equation 2 (i.e., without applying GW detector selection effects). We show for each of the 20 binary population synthesis models (Table 1) the merger rates for the 28 variations in $S(Z, z)$ (Table A1). We connect predictions that use the same $S(Z, z)$ model with a line for visual reasons only. Two $S(Z, z)$ variations, xyz=231 (dashed) and xyz=312 (dotted), are highlighted, corresponding to the $S(Z, z)$ models resulting in one of the highest and lowest merger rate predictions, respectively. The values $\langle \sigma_\mu \rangle$ and $\langle \sigma_{xyz} \rangle$ quote a proxy for the mean scatter (Equation 3 and 4) in the predicted rates due to variations in binary population synthesis and $S(Z, z)$, respectively. The minimum and maximum rates and their ratio are quoted with a black error bar on the left. We use the short-hand notation $\mathcal{R}_m^0 \equiv (dN_{\text{det}}^2/dt_s dV_c)(t_m(z \approx 0))$. Shaded horizontal bars indicate the ranges defined by the 90% credible intervals for the intrinsic merger rates inferred from GW observations in Abbott et al. (2021a,c,d): $\mathcal{R}_m^{0,\text{BHBH}} = 10.3-104 \text{ Gpc}^{-3} \text{ yr}^{-1}$ (BHBH), $\mathcal{R}_m^{0,\text{BHNS}} = 12-242 \text{ Gpc}^{-3} \text{ yr}^{-1}$ (BHNS) and $\mathcal{R}_m^{0,\text{NSNS}} = 49-810 \text{ Gpc}^{-3} \text{ yr}^{-1}$ (NSNS). Model realizations where all three of the BHBH, BHNS and NSNS simulated merger rate densities overlap with these observed ranges are marked with red crosses. Dashed horizontal lines indicate the range for the fiducial model (A). In the top we added colored labels to indicate what physics assumptions are varied compared to our fiducial assumptions in the models.

over-predict the observed BHBH rate for all stellar evolution models are the ones with an implementation of the Langer & Norman (2006) MZR ($z = 1$). This MZR model corresponds to the lowest age Z among our MZR variations (right panel, Figure A2), which significantly increases the BHBH yield. The over-prediction is even more significant if other formation channels than the isolated binary evolution channel further contribute to the observed BHBH rate (e.g., Abbott et al. 2021c; Zevin et al. 2021). GW observed BHBH rates thus potentially present a good testbed for constraining $S(Z, z)$ models.

3.2.2 BHNS merger rates

For the merging BHNS systems we find rates in the range $\mathcal{R}_m^{0,\text{BHNS}} \approx 2.2-830 \text{ Gpc}^{-3} \text{ yr}^{-1}$ as shown in the middle panel of Figure 2. Our combined model variations impact the predicted BHNS rates with

an uncertainty factor of up to ≈ 377 . We find that variations in the rate are typically $\langle \sigma_\mu \rangle \approx 69$ for binary evolution variations, while the impact from $S(Z, z)$ variations typically leads to a range of $\langle \sigma_{xyz} \rangle \approx 8$.

There are a few trends visible in the BHNS rates in Figure 2. First, the predicted rate declines with increasing mass transfer efficiency, β (models B, C and D). The increase in β leads the secondary to accrete more mass during the first stable mass transfer phase, resulting in more massive secondaries and typically larger separations. The larger masses and/or larger separations at higher β values results in more binaries eventually disrupting during a SN, merging during the CE phase, and/or forming a BHBH binary instead of a BHNS (cf., Kruckow et al. 2018). In addition, this also impacts t_{inspiral} .

Edo: this is a very long discussion of BBH and BNS in the middle of a BHNS section - can it be moved elsewhere? The BHBH and NSNS rates instead increase with increasing values for

This different behavior comes from a combination of ways in which the formation pathways of binaries leading to merging BHNSs are different from merging BHBHs and NSNSs and how changing β impacts this. An example is that the impact from β on the evolution of the binary is connected with the mass ratios of the binary (the mass ratio impacts the lobes, whether the mass transfer is stable or unstable and if the binary widens or shrinks). The merging BHBH and NSNS populations are DCOs with more equal (i.e. $q \sim 1$) mass ratio distributions compared to BHNS (see also §3.4). The binaries that form merging BHBHs and NSNSs are as a result binned to different mass ratio populations at the same evolutionary stages compared to merging BHNSs. This results in the μ parameter impacting the populations differently. Another difference is that the merging BHNS typically form from different formation channels within isolated binary evolution compared to BHBH and NSNS. For example, in our fiducial model (A) a substantial fraction ($\gtrsim 25\%$ of all merging BHBHs in our simulation) of the BHBHs form through only stable mass transfer episodes without engaging a CE, whereas for BHNSs this is much smaller ($\lesssim 5\%$). The parameter β has a different impact on each of these formation channels.

Second, the BHNS rate increases with increasing values for α (models G, H, I and J) as a result of more efficient ejection of the CE. This causes typically less systems to merge during the CE (cf., Kruckow et al. 2018, Figure 17). On the other hand, the rates slightly decline for the highest α_{CE} values as in this case binaries do not shrink enough to merge in a Hubble time (see also the discussion on α in §3.1.1 from Bavera et al. 2021). The highest predicted BHNS rates are from the models P, Q and R, where the majority of the $\mathcal{S}(Z, z)$ variations predict BHNS merger rates $\mathcal{R}_m^{0, BHNS} \gtrsim 100 \text{ Gpc}^{-3} \text{ yr}^{-1}$. These are the models with reduced SNe kicks, which increases the fraction of systems that stay bound during the SN compared to our fiducial model (see paper I for a more detailed discussion).

Almost all predicted BHNS rates are consistent with the range $\mathcal{R}_m^{0, BHNS} 12\text{--}242 \text{ Gpc}^{-3} \text{ yr}^{-1}$ spanned by the 90% credible intervals from observations (Abbott et al. 2021d). The exceptions are a subset of the $\mathcal{S}(Z, z)$ models in combination with models C, D, E and G that under-predict the BHNS rate and a few of the $\mathcal{S}(Z, z)$ models with $z = 1$ in combination with stellar evolution model Q ($\sigma_{rms}^{1D} = 30 \text{ km s}^{-1}$), which slightly over-predict the observed upper limit. Future improved constraints might enable ruling out models.

3.2.3 NSNS merger rates

The bottom panel in Figure 2 shows our predicted merging NSNS rates. We find values in the range $\mathcal{R}_m^{0, NSNS} \approx 0.32\text{--}330 \text{ Gpc}^{-3} \text{ yr}^{-1}$. Combined, our 560 model realizations impact the estimated NSNS rates up to a factor of ≈ 1034 . The uncertainty from stellar and binary evolution dominates the scatter in the calculated NSNS rates (cf., Santoliquido et al. 2021), impacting the rate by $\langle \sigma_\mu \rangle \approx 226$ compared to $\langle \sigma_{xyz} \rangle \approx 2$ when varying the $\mathcal{S}(Z, z)$.

Model E in particular is an outlier in the predicted NSNSs merger rate. In this channel the formation yield of merging NSNSs is extremely low, as also shown in Figure 1. This is because the majority of systems leading to merging NSNSs experience a case BB mass transfer phase that ultra-strips the NS progenitor (Tauris et al. 2017; Vigna-Gómez et al. 2018). In our model variation E we assume this phase of mass transfer to be unstable, leading to a stellar merger in combination with our pessimistic CE assumption. Indeed, in model F where we have the same settings as in model E but add the ‘optimistic’ CE assumptions, we obtain NSNS rates consistent with our

fiducial model. If we exclude model E we find instead $\langle \sigma_\mu \rangle \approx 10 \times$ (while $\langle \sigma_{xyz} \rangle$ remains unchanged).

A large fraction of our predicted NSNS merger rates fall below the observed NSNS merger rate density range of $\mathcal{R}_m^{0, NSNS} = 49\text{--}810 \text{ Gpc}^{-3} \text{ yr}^{-1}$ from Abbott et al. (2021a,c). However, models D, I, J, K, N, O, P and Q are consistent with the observed merger rate density for some $\mathcal{S}(Z, z)$ model variations. In particular, models I, J, P and Q lie well above the lower observed range of $49 \text{ Gpc}^{-3} \text{ yr}^{-1}$. These are models in which the NSNS rate is boosted. For example, in models I and J fewer binaries merge during the CE phase compared to the fiducial model. In models P and Q the smaller value for the root-mean-square velocity dispersion reduces the kick velocities that decreases the number of binaries that disrupts during the SN. More discussion is provided in §4.

3.3 GW detectable merger rates

The top panel in Figure 3 shows the predicted detectable merger rates for our 560 model variations. These are calculated by volume integrating Equation 2 and taking into account the sensitivity of a GW detector, quantified by the probability P_{det} of observing a merging binary of specified masses at a given redshift and corresponding distance. Here we assume a detector sensitivity comparable to Advanced LIGO in its design configuration (hereafter, design sensitivity; LIGO Scientific Collaboration et al. 2015; Abbott et al. 2016, 2018). We follow Barrett et al. (2018) and choose a detector signal-to-noise ratio threshold of 8 as a proxy for detectability by a GW detector network (e.g., LVK). The detectable merger rate is given by:

$$\mathcal{R}_{det}(t_{det}, m_1, m_2) = \frac{d^3 N_{det}}{dt_{det} dm_1 dm_2} = \int dV_C \frac{dt_s}{dt_{det}} \mathcal{R}_m(t_m) P_{det}(m_1, m_2, z), \quad (5)$$

where t_{det} is the time in the detector (i.e., the observer) frame, and m_1 and m_2 are the component masses of the DCO in the source frame (see paper I for further details).

Considering the 560 model variations we find predicted detected rates for an LVK detector network at design sensitivity in the range $\mathcal{R}_{det}^{BHBH} \approx 17\text{--}12000 \text{ yr}^{-1}$, $\mathcal{R}_{det}^{BHNS} \approx 0.36\text{--}180 \text{ yr}^{-1}$ and $\mathcal{R}_{det}^{NSNS} \approx 0\text{--}12 \text{ yr}^{-1}$, as shown in the top panel of Figure 3. We find that the stellar, binary, and cosmic evolution combined impact the predicted detectable merger rates by factors of up to $\approx 500 - 1270$.

3.3.1 Relative merger rates

The bottom panel of Figure 3 shows the relative merger rates between the different DCO channels for LVK at design sensitivity. We find that almost all of our models predict a higher BHNS detection rate compared to the NSNS detection rate, except for model D (which has a mass transfer efficiency of $\beta = 0.75$) and model G (which assumes $\alpha_{CE} = 0.1$) in which a subset of the $\mathcal{S}(Z, z)$ models lead to higher NSNS detection rates. The higher BHNS rate is a result from both the high intrinsic yield of BHNS mergers compared to NSNS mergers (Figure 1) and the larger detection volume for BHNS compared to NSNS mergers as a result from their larger masses (§3.4).

3.4 GW detectable mass distribution functions

The uncertainties in stellar and binary evolution and $\mathcal{S}(Z, z)$ also impact the shapes of the mass distributions of detectable DCO mergers, in addition to the overall merger rate densities. We show

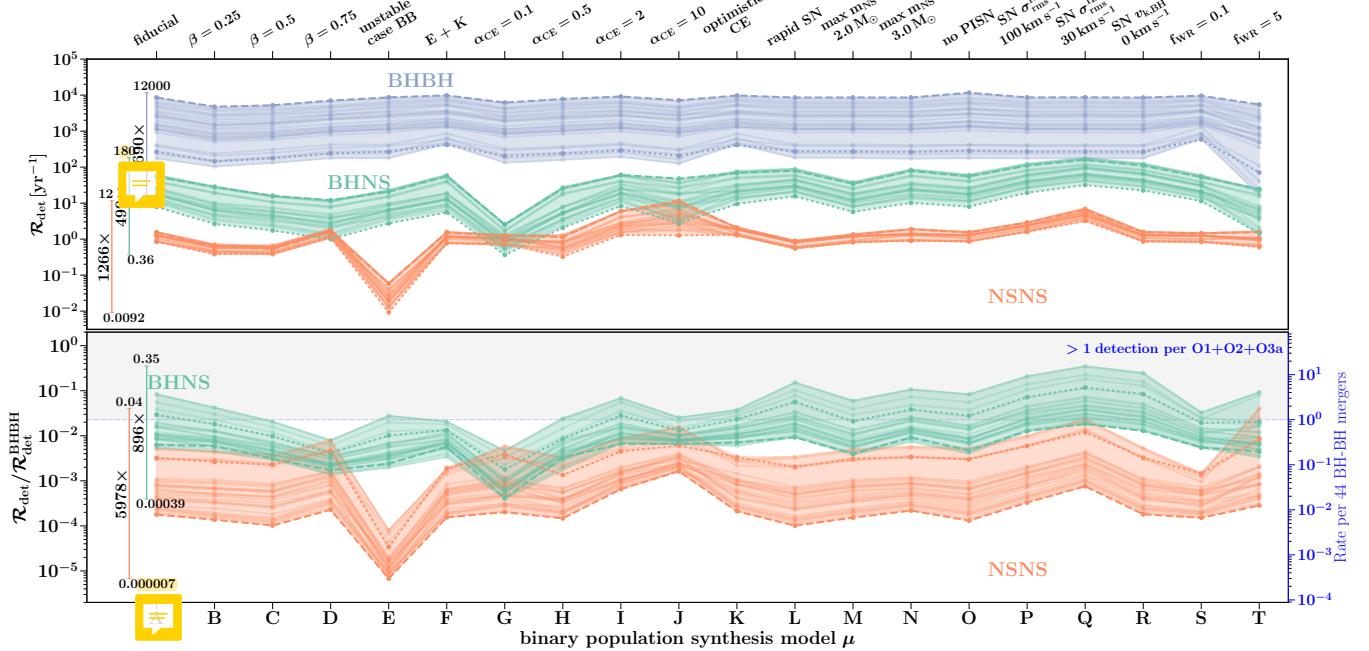


Figure 3. Top panel: The calculated detectable merger rates for merging BHBH (blue), BHNS (green) and NSNS (orange) for a GW network at design sensitivity for our 560 models. **Bottom panel:** The predicted merging BHNS and NSNS detectable rate ratios over merging BHBH. On the right axes we scale the ratios to the 44 BHBH detections with false-alarm rate $\lesssim 1 \text{ yr}^{-1}$ from the combined GWTC-1 and GWTC-2 catalogs (Abbott et al. 2019, 2021b,c). **Both panels:** The layout, lines and symbols are as in Figure 2.

this in Figures 4 (BHBH), 5 (BHNS) and 6 (NSNS) for our 560 model realizations. Each Figure shows the normalized distributions for the DCO component masses m_1 and m_2 , chirp mass $M_c = (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$, and mass ratio $q = m_1/m_2$, where we use subscripts ‘1’ and ‘2’ to indicate the more massive and less massive component in the binary, respectively. To compare the shapes we show kernel density distributions for the chirp mass and summary statistics (i.e., median, 50, 90 and 99 distribution percentiles) for the individual masses, chirp mass and mass ratio for the BHBH, BHNS and NSNS mergers. All distributions are weighted for the detection volume and sensitivity of LVK at design sensitivity using Equation 5. Each distribution is given by $d^2N_{\text{det}}/dt_{\text{det}} dx$, with x being one of the mass parameters mentioned above.

3.4.1 Impact from binary star and cosmic evolution on the distribution shapes

We qualitatively compare the impact on the shape of the detectable mass distributions from variations in stellar and binary evolution and variations in $\mathcal{S}(Z, z)$ by analyzing how the distributions (and distribution statistics) vary; see Figures 4, 5 and 6. Distributions that are similar between vertical panels (different colors) indicate that changes in stellar and binary evolution assumptions do not significantly impact the distribution shape. On the other hand, distributions that are similar between variations plotted within one sub-panel indicate these $\mathcal{S}(Z, z)$ realizations do not significantly impact the shape of the distributions. By comparing these two effects directly we can qualitatively analyze which of the two uncertainties dominates the shape of the detectable mass distributions.

For merging BHBH systems we find that the uncertainty in the shape of the mass distribution is dominated by variations in the

$\mathcal{S}(Z, z)$, reflecting the metallicity dependence of BHBH formation (Figure 4; §3.1), which also strongly impacts the resulting BH masses (cf., Belczynski et al. 2010; Giacobbo & Mapelli 2018). For example, the peak in the detectable BHBH chirp mass distribution around $30\text{--}35 M_{\odot}$ that is visible in the majority of models in Figure 4 disappears in a small subset of the $\mathcal{S}(Z, z)$ models. These are the $\mathcal{S}(Z, z)$ models with the Panter et al. (2004) GSMF and Ma et al. (2016) MZR ($xyz = 113, 213$ and 313), which correspond to realizations with high average metallicities compared to the other $\mathcal{S}(Z, z)$ models (see Figure B1 in paper I) and lead to fewer massive BHBH mergers. The stellar and binary evolution variation that impacts the mass distributions most drastically is the model in which we assume pair-instability SNe do not occur (model O), leading to the formation of BHs with masses $\gtrsim 40 M_{\odot}$. This model assumption may be unrealistic (Wesley (e.g., 2017); Farmer et al. (e.g., 2019)). Other significant changes, best visible in the kernel density functions (left column in Figure 4), are present for models V (optimistic CE), L (rapid SN remnant model) and T (fwr = 0.1). where, for example, double chirp mass peak shifts or disappears compared to the fiducial model A. Overall, we find that the BHBH mass distribution shapes can be significantly impacted by both our explored stellar evolution and $\mathcal{S}(Z, z)$ variations.

For merging BHNS systems, on the other hand, the distribution shapes are predominantly impacted by the stellar and binary evolution model variations compared to the $\mathcal{S}(Z, z)$ variations, (Figure 5). Primary effects include models E (unstable case BB), L (rapid SN), M (max $m_{\text{NS}} = 2 M_{\odot}$) and N (max $m_{\text{NS}} = 3 M_{\odot}$), where the median and distribution percentiles are visibly different compared to the fiducial model A. This indicates that the shape of the BHNS distribution shapes can present a good testbed for constraining stellar and

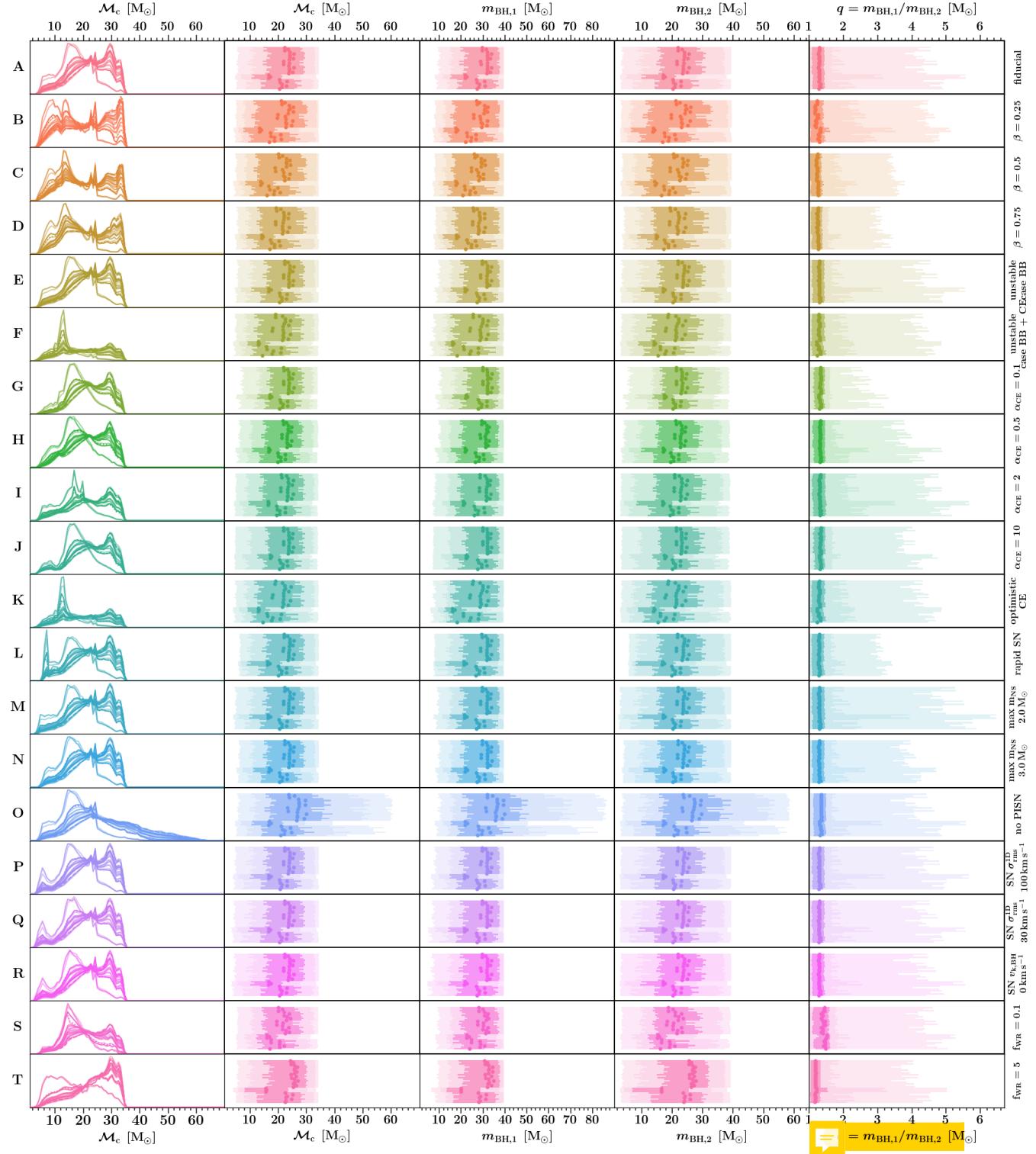


Figure 4. Shape of the distribution functions for detectable BHBH mergers from our 560 model realizations. **First column:** the probability distribution functions of the chirp mass \mathcal{M}_c . The distributions are normalized and given by $\mathcal{R}_{\text{det}}(x) = (d^2N_{\text{det}}/dt_{\text{det}} dx)$ for parameter x (Equation 5). **Other columns:** From left to right we show the chirp mass, the primary (most massive) and secondary (least massive) BH mass, and the mass ratio. In each sub-panel we show 28 individual horizontal bars that visualize the median (scatter points) and the 50, 90 and 99 (three shades) distribution percentiles. A zoom-in that shows the order of the $S(Z, z)$ models is given in Figure A3. **Rows:** From top to bottom, the panels/colors show the 20 different stellar evolution simulations from Table 1. Within each sub-panel we show the 28 $S(Z, z)$ models with different lines (first column) or horizontal bars (other columns).

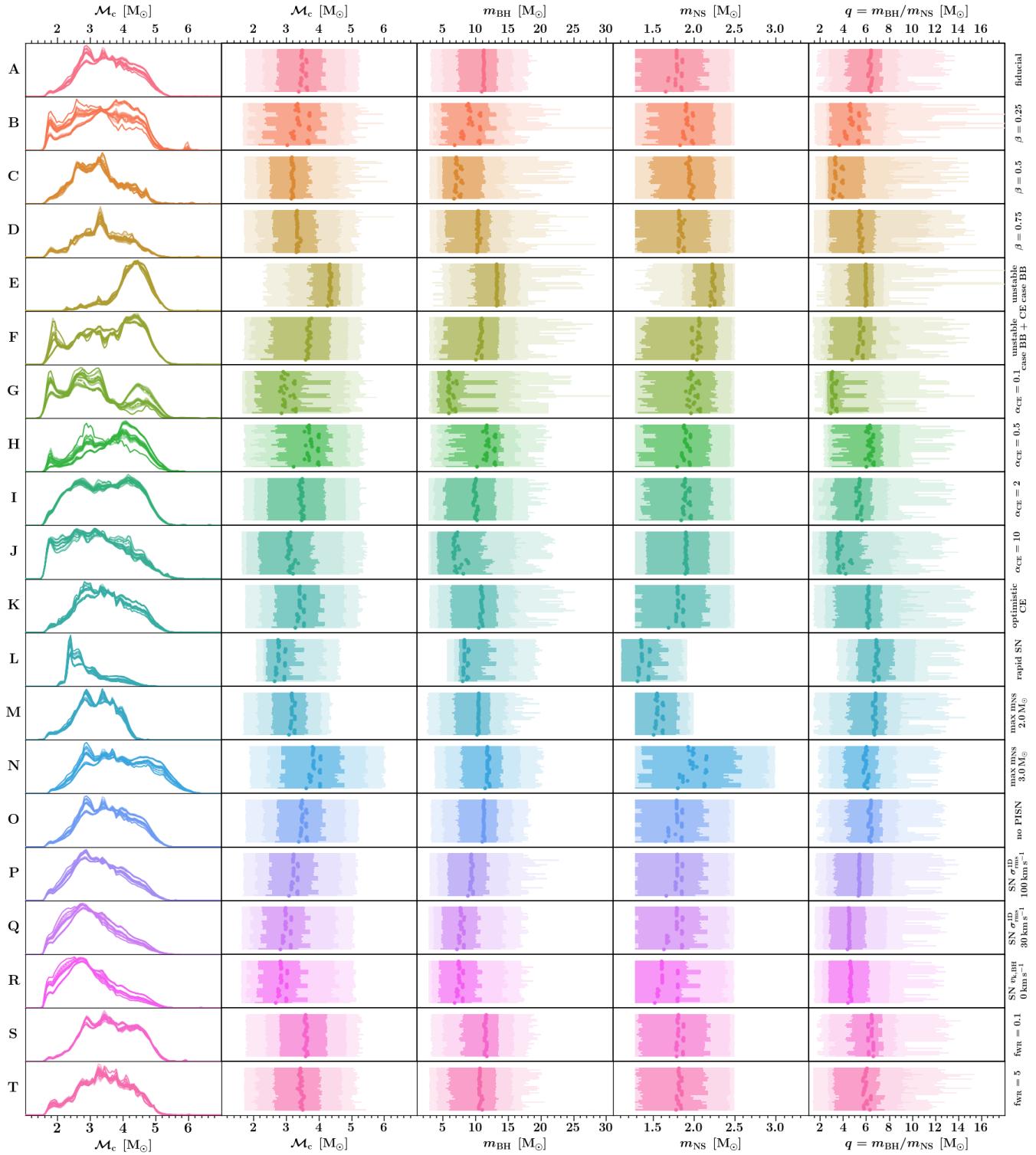


Figure 5. Same as Figure 4 for BHNS mergers. For the individual masses we show the BH (m_{BH}) and NS (m_{NS}) masses.

Binary evolution models. Edo: need to say something about SFRD variations - are they negligible?

For NSNS systems the shape of the normalized mass distributions are impacted by both sets of variations (Figure 6). The main reason the shape of the NSNS distributions is dependent on the $\mathcal{S}(Z, z)$ is

because the relative contributions from different formation channels leading to NSNS mergers are metallicity dependent. **Edo: is this not true for BHNS as well???** NSNS primarily form through two channels: the classic CE channel and the double-core common-envelope channel (e.g., Vigna-Gómez et al. 2018). Both channels result in

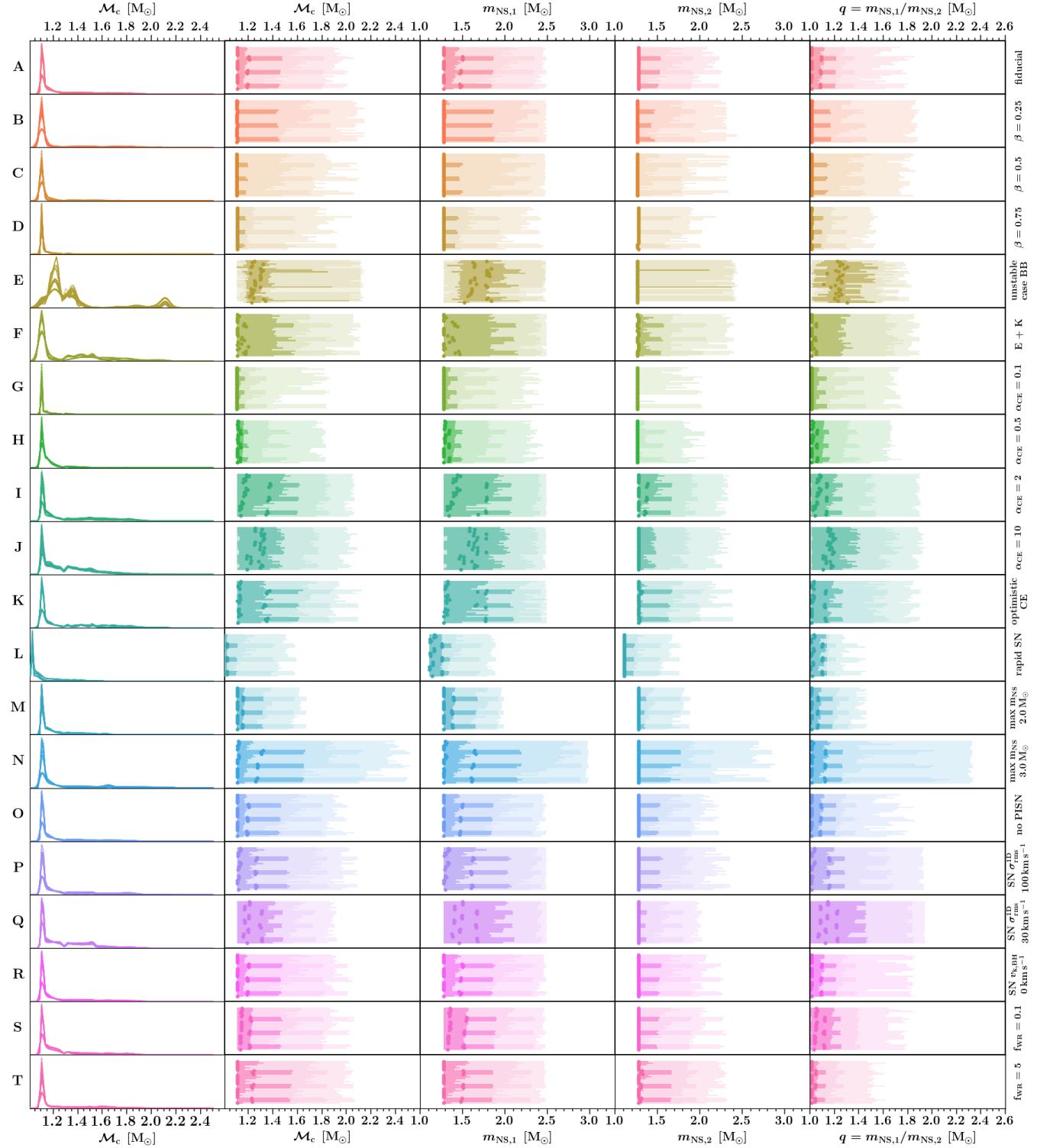


Figure 6. Same as Figure 4 for NSNS mergers. For the individual masses we show the primary (most massive, $m_{\text{NS},1}$) and secondary (least massive, $m_{\text{NS},2}$) NS mass. The shape of model E (unstable case BB) suffers significantly from sampling noise due to the low number of NSNS systems in this variation.

different NSNS mass properties; for example, the double core CE channel typically leads to more massive NSNS systems. Since the fractional contribution of each channel is metallicity dependent, the shape is impacted by different choices of $S(Z, z)$ models (this has a lesser impact on the merger rates; §3.1). Among the binary stellar

evolution variations, the models that impact the shape of the distribution significantly are realizations that change the SN remnant mass or kick velocity, and/or the CE efficiency (models G, H, I, J, K, L, M, N, P and Q). The distributions from model E (unstable case BB

mass transfer) are dominated by sampling noise as a result of the low NSNS yield in this model (§3.1).

In summary, we find that for BHNSs, the variations in the distribution shapes are typically dominated by stellar and binary evolution assumptions (see also paper I), suggesting that observations of BHNS systems can aid in constraining stellar evolution models. For merging BHBH and NSNS the distribution shapes are impacted by both sets of variations (Figures 4 and 6). We therefore argue that constraining stellar evolution or $\mathcal{S}(Z, z)$ models solely from the distribution shapes of merging BHBHs or NSNSs may be challenging. We discuss this further in §4.

3.4.2 Distribution properties considering all model realizations

In the previous section we showed that the model uncertainties in binary and stellar evolution, and $\mathcal{S}(Z, z)$ can significantly impact the shapes of the DCO mass distributions. Here we instead discuss several specific examples of features in the mass distributions in Figure 4, 5 and 6 that are robust across all 560 model variations explored in this work.

First, in all 560 model variations 95% (99%) of the detectable BHBH mergers have mass ratios $q_f \lesssim 4$ ($q_f \lesssim 6$) (last column in Figure 4). The BHBH distribution medians and percentiles furthermore indicate that in all model variations BHBHs from isolated binary evolution prefer order unity mass ratios. Detecting a significant fraction ($\gtrsim 5\%$) of BHBH mergers with large mass ratios ($q_f \gtrsim 5$) would point to other formation pathways or missing physics in our simulations. This is consistent with the majority of the 44 BHBH detections announced by LVK, which are inferred to have mass ratios consistent with unity (Abbott et al. 2021b). However, events including the BHBH merger GW190412 ($1/q_f = 0.28^{+0.12}_{-0.07}$; Abbott et al. 2020a) and the BHBH merger candidate GW190814 ($1/q_f = 0.112^{+0.008}_{-0.009}$; Abbott et al. 2020c), if common, hint to the existence of a population with more extreme mass ratios.

Second, in all 560 model variations $\lesssim 5\%$ of the detectable merging BHNS systems have BH masses $m_{BH} \gtrsim 18 M_\odot$ (third column of Figure 5). As discussed in paper I, this is due to the fact that more equal mass stars more readily survive important stellar evolution phases in the formation pathways to BHNS mergers; to form a BHNS one of the stars needs to be of sufficiently low mass to form a NS, leading to a preference for the other star to also be of lower mass than is typical in BHBH. Because the black hole mass dominates the chirp mass, we find that in all our 560 model realizations $\lesssim 5\%$ of BHNS mergers have chirp masses of $\gtrsim 5.5 M_\odot$. This suggests that detecting a BHNS merger with a BH mass $m_{BH} \gtrsim 18 M_\odot$ could indicate that the system did not form from isolated binary evolution, but instead from chemically homogeneous evolution and/or dynamical formation where such high BH masses are more commonly expected (e.g. Marchant et al. 2017; McKernan et al. 2020; Rastello et al. 2020). The two detected BHNS systems, GW200105 and GW200115, have $\mathcal{M}_c \lesssim 3.5 M_\odot$ (at the 90% credible interval) (Abbott et al. 2021d) consistent with our models (Broekgaarden & Berger 2021; Broekgaarden et al. 2021).

Third, we find that in all 560 model variations the NSs in detectable NSNS mergers typically have lower NS masses compared to those in detectable BHNS mergers. This is true for both NS components in the NSNS mergers (Figures 5 and 6). This is again due to the preference for equal mass binaries to survive important evolutionary phases leading to BHNS formation, thereby favoring more massive pre-NS stars to form a BHNS system (see paper I for more details). To date two confident NSNS and BHNS detections have been announced by

LVK. For the NSNS detections the NS masses are $m_{NS,1} = 1.46^{+0.12}_{-0.10}$ and $m_{NS,2} = 1.27^{+0.09}_{-0.09}$ (GW170817; Abbott et al. 2017) and $m_{NS,1} = 2.0^{+0.6}_{-0.3}$ and $m_{NS,2} = 1.4^{+0.3}_{-0.3}$ (GW190425; Abbott et al. 2020b). For the BHNS detections the NS masses were found to be $m_{NS} = 1.9^{+0.3}_{-0.2} M_\odot$ (GW200105) and $m_{NS} = 1.5^{+0.7}_{-0.3} M_\odot$ (GW200115) (Abbott et al. 2021d). More detections are needed to calculate robust median NS masses in NSNS and BHNS detections.

Fourth, $\gtrsim 95\%$ of all BHBH mergers are expected to contain a BH with $\gtrsim 8 M_\odot$ in all 560 model variations (third column of Figure 4). This is not the case for BHs in BHNS systems where typically $\gtrsim 25\%$ of the detectable mergers are expected to contain a BH of $\lesssim 8 M_\odot$ (Figure 5). Our models predict the secondary BH in the population of detected BHBH to also commonly be massive ($\gtrsim 8 M_\odot$), although many of our models do allow for $\gtrsim 25\%$ of the detectable BHBH mergers to contain a secondary BH with $\lesssim 10 M_\odot$. For the currently reported significant BHBH detections by LVK all of the inferred medians of the primary BH mass have BH masses $m_{BH,1} \gtrsim 8 M_\odot$ (Abbott et al. 2021a,b).

Besides the four specific points above, more common trends throughout our 560 model variations are visible in Figures 4, 5 and 6, particularly in the lower and upper bounds for the BH and NS masses. Several of these trends, however, are direct results from the remnant mass prescriptions used in our models. For example, all models with the exception of model O (no PISN) predict that 90% of the detected BHBH mergers have chirp masses in the range $\mathcal{M}_c \approx 5 - 35 M_\odot$. The upper limit is set by the maximum possible BH masses of $\sim 40 M_\odot$ allowed due to the implementation of pair-instability SN in our models (e.g., Marchant et al. 2019; Stevenson et al. 2019); in model O we do not implement pair-instability SNe, and find BH masses up to about $80 M_\odot$ and BHBH chirp masses up to about $60 M_\odot$. In addition, in most of our model realizations the majority of NSs in NSNS mergers have masses of $\approx 1.3 M_\odot$ (consistent with e.g., Tauris et al. 2017; Vigna-Gómez et al. 2018), due to at least one of the stars experiencing an electron-capture SN (ECSN), which in COMPAS leads to a mass of $1.26 M_\odot$ (paper I, Team COMPAS: J. Riley et al. 2021). Similarly, many NSNS mergers in our models originate from relatively low mass stars with carbon-oxygen cores of $\lesssim 2.5 M_\odot$ at the time of the SN, which in the delayed SN remnant mass prescription are mapped to a fixed value of $m_{NS} \approx 1.28 M_\odot$ (Fryer et al. 2012). Combined, this results in a peak of NSs masses at $m_{NS} \approx 1.3 M_\odot$. Exceptions to this are the rapid SN remnant mass model (model I) where these low mass stars are instead mapped to NS masses of $\approx 1.1 M_\odot$ leading to a broader peak at $\approx 1.1 - 1.3 M_\odot$. The upper and lower NS limits are also artificially set in our models to $\approx 1 M_\odot$ and $\approx 2.5 M_\odot$, respectively (cf., Fryer et al. 2012).

3.5 Delay time and birth metallicity distributions

The uncertainties in the stellar evolution and $\mathcal{S}(Z, z)$ prescriptions in population synthesis modelling can also impact the expected delay time (t_{delay}) and birth metallicity (Z) distributions of the detectable DCO mergers. We show the impact from the 560 model realizations explored in this study in Figures 7, 8 and 9. We focus on the t_{delay} and Z distributions as these properties can be (indirectly) constrained from observations (e.g., Im et al. 2017; Safarzadeh et al. 2019; Fishbach & Kalogera 2021). In addition, these distributions provide insights into the range of birth metallicities and star formation redshifts that current GW observations probe, which can be informative for future modelling (e.g., help understand where to focus computational time). We note that impacts on the t_{delay} distribution also impact (indirectly) the distribution of Z , because the

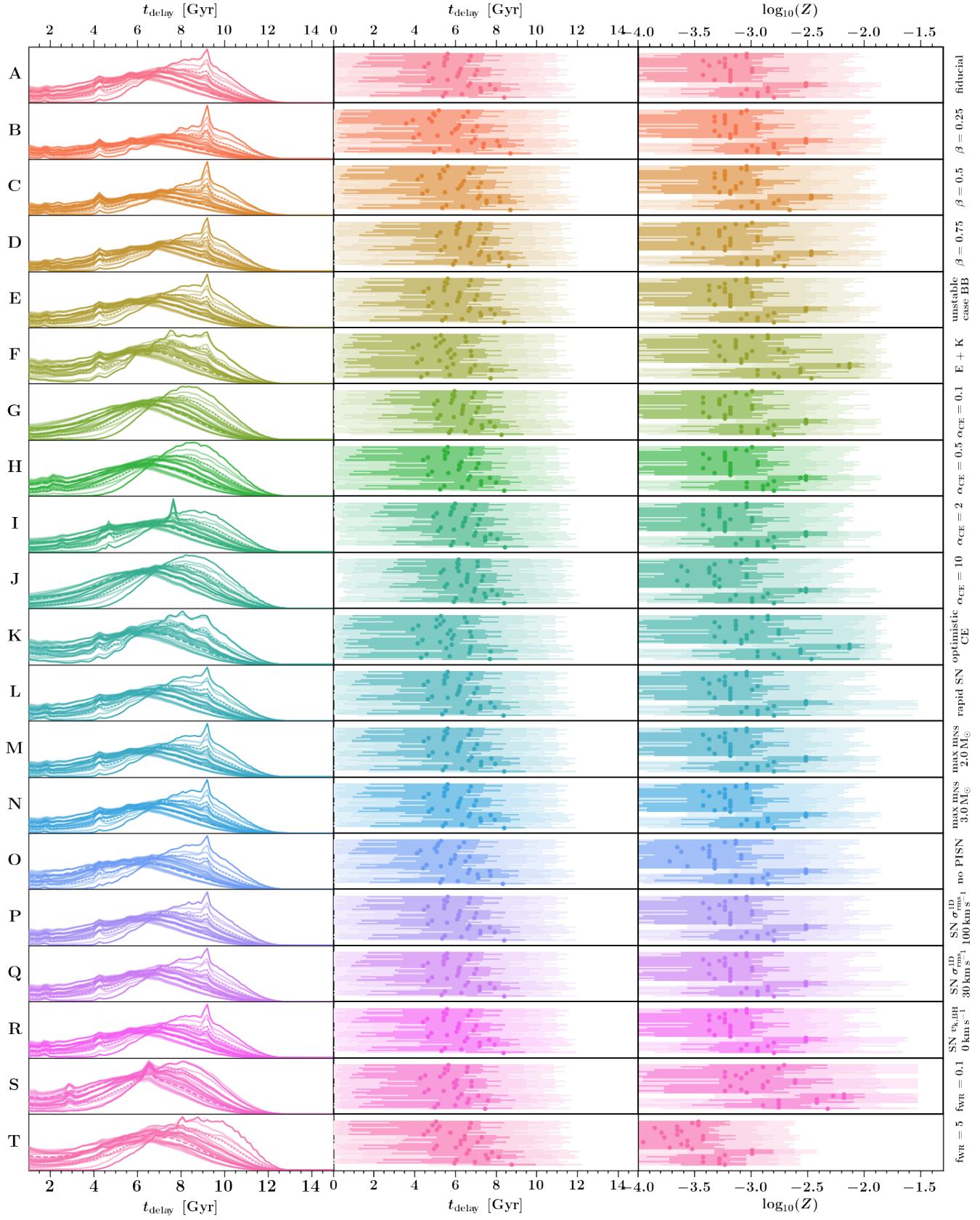


Figure 7. The delay time t_{delay} and birth metallicity Z distributions for BHBB mergers for our 560 model realizations. The distributions and percentiles are plotted as in Figure 4. The delay time is the time since formation of the binary system at the zero-age main sequence until the moment of merger.

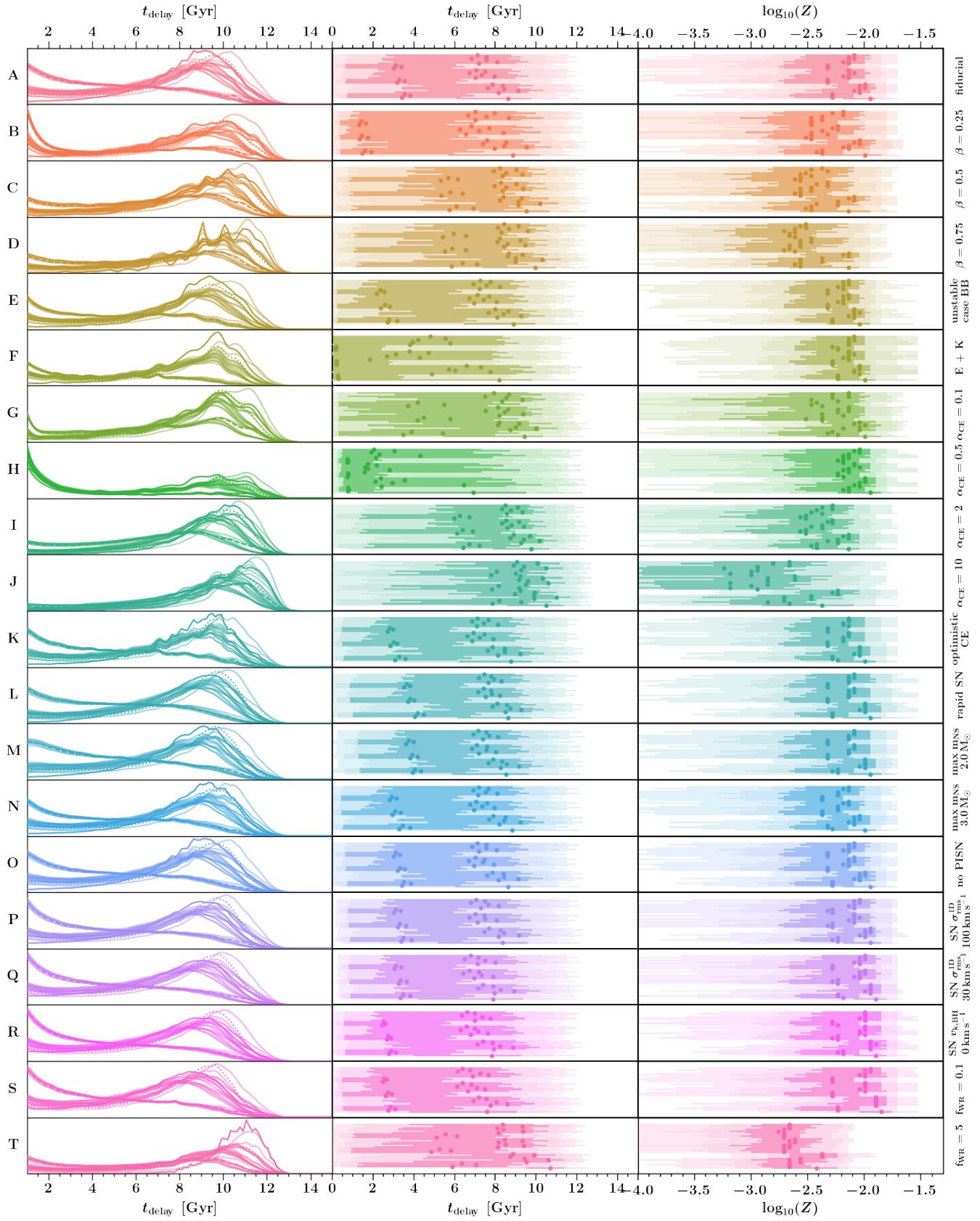


Figure 8. Same as Figure 7 for BHNS mergers.

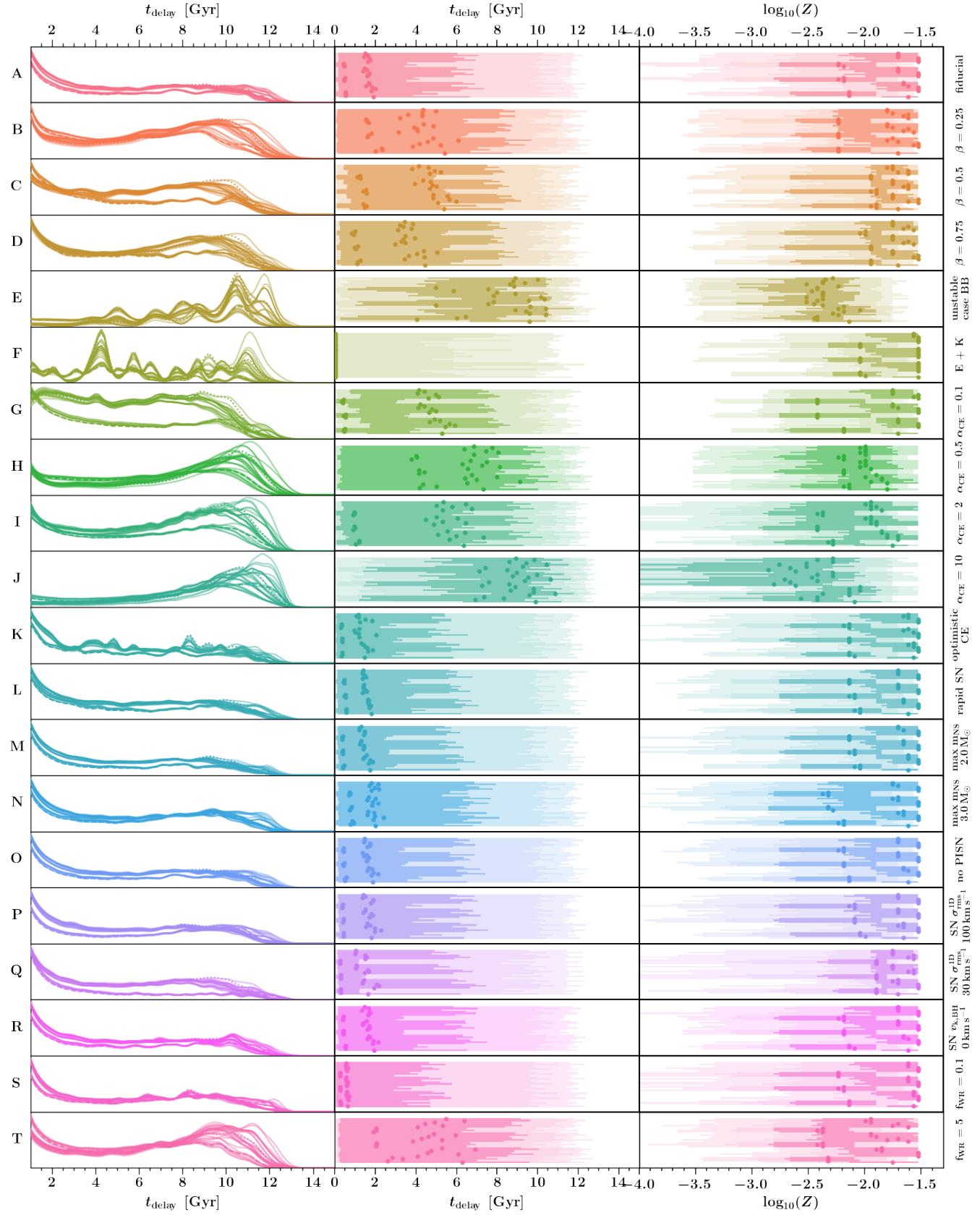


Figure 9. Same as Figure 7 for NSNS mergers. The shape of model E (unstable case BB) suffers significantly from sampling noise due to the low number of NSNS systems in this variation.

birth metallicity distribution (dP/dZ ; §2.2) evolves as a function of redshift.

The most striking feature of the t_{delay} and Z distributions is that they are significantly impacted by the $\mathcal{S}(Z, z)$ variations for all three DCO merger types. This is particularly discernible in the large scatter of the distribution percentiles (see each sub-panel). The impact from $\mathcal{S}(Z, z)$ models is particularly significant for the choice of the metallicity probability function (dP/dZ), which describes the distribution of birth metallicities of stars at a given redshift. This is because the birth metallicity strongly impacts the properties and rate of the DCO mergers, including through the metallicity-dependent formation yields, through mass loss that impact the DCO masses and can widen the orbit of the binary, and through the metallicity-dependent radial expansion of stars, which impacts mass transfer phases in our models (§3.1). In the convolution to obtain the detectable DCO mergers in Equation 5 all of these properties impact the resulting population. Thus, the t_{delay} and Z distributions of the detectable DCO mergers are significantly impacted by the choice of (dP/dZ) and $\mathcal{S}(Z, z)$.

Variations in the stellar and binary evolution assumptions can also significantly impact the t_{delay} and Z distributions. For example, model T ($f_{\text{WR}} = 5$), which increases the mass loss through stellar winds, drastically suppresses the number of BHBH and BHNS events that can form at higher metallicities ($\log(Z) \gtrsim -2.5$), thereby leading to fewer of the detectable BHBH and BHNS mergers forming from these metallicities (Figures 7 and 8). Model J ($\alpha_{\text{CE}} = 10$) resulting in longer t_{delay} and lower Z compared to our fiducial model in Figure 8 and 9. In model J orbital angular momentum of the binary can much more efficiently be transformed into ejecting the CE. This leads to less orbital shrinking during the CE phase and longer inspiral times. As a result only BHNSs and NSNSs that formed early in the Universe with low Z have had long enough the time to inspiral and be detectable today. Model J does not impact the BHBH distributions as significantly because most detectable BHBH mergers in our simulations only go through stable mass transfer phases.

Overall, we find that the delay time distributions span a broad range between a few Myr and the Hubble time. We find that our model variations indicate that the detectable BHBH mergers (Figure 7) have systems with the longest median delay times, compared to BHNS and NSNS mergers (Figures 8 and 9). However, exceptions exist, including model J ($\alpha_{\text{CE}} = 10$) where the median delay times of detected BHNS and NSNS mergers are larger than those of BHBH mergers for the reasons discussed above.

For the birth metallicity distribution shapes the right panels of Figures 7, 8 and 9 show that the detectable DCO mergers arise from a broad range of birth metallicities. As expected from the formation yield efficiencies, BHBH mergers typically originate from the lowest birth metallicities, with most of our models having a median $\log_{10}(Z)$ of $\lesssim -2.5$, whereas the detectable BHNS and NSNS populations originate in our models from higher Z with most models having median $\log_{10}(Z)$ of $\gtrsim -2.5$.

4 DISCUSSION

4.1 The reality of population synthesis models

In this paper we demonstrate the importance of considering the uncertainties in both massive star and binary evolution, as well as the $\mathcal{S}(Z, z)$ when aiming to learn about the formation, evolution, and mergers of binaries through a comparison of population synthesis models to GW observations. Here we discuss the implications of our results for interpreting population synthesis studies.

In §3.2 and §3.3 we showed that the calculated intrinsic and detectable merger rate densities (\mathcal{R}_{m}^0 and \mathcal{R}_{det} , respectively) can be impacted by factors of up to 200–1300 by combined uncertainties in stellar evolution and $\mathcal{S}(Z, z)$. Our results expand on and confirm earlier studies that focused on one source of uncertainty or the other (e.g., Dominik et al. 2015; Giacobbo & Mapelli 2018; Kruckow et al. 2018; Vigna-Gómez et al. 2018; Chruslinska et al. 2019; Neijssel et al. 2019; Belczynski et al. 2020; Tang et al. 2020; Santoliquido et al. 2021). In addition, we showed that the shape of the DCO detectable mass distributions (chirp mass, individual masses and mass ratio) can be significantly impacted by both stellar evolution and $\mathcal{S}(Z, z)$ uncertainties. These findings also expand on and confirm earlier exploratory work (Chruslinska et al. 2019, Figure 4) and (Neijssel et al. 2019, Figure 9), where the authors discuss the impact from a few model variations on the GW detectable BHBH chirp mass distribution. Similarly, we showed that the stellar evolution and $\mathcal{S}(Z, z)$ uncertainties can also significantly impact the shape of the expected DCO merger delay time and metallicity distributions. These findings provide a cautionary note for using GW observations to uncover stellar and binary evolution properties.

First, studies that draw conclusions by comparing population synthesis results with GW observations without considering the wide range of uncertainties in both stellar/binary evolution and $\mathcal{S}(Z, z)$ assumptions could be drastically biased by the specific model realization that is chosen to derive the results (cf. Belczynski et al. 2021). For example, studies including Zevin et al. (2017), Bouffanais et al. (2019), Zevin et al. (2021) and Franciolini et al. (2021) aim to determine the contributions of different formation channels to the observed BHBH population, using predicted merger rates and/or distribution shapes from only one or at most a few population synthesis model variations. Another example are studies including Fragione (2021), which estimate the number of BHNS mergers with an electromagnetic counterpart based on only a few binary population synthesis models (changing for example only SN kicks and α_{CE}). Our work shows that uncertainties in both stellar and binary evolution and $\mathcal{S}(Z, z)$ modelling can drastically impact the rate and distribution shapes of the BHBH and BHNS populations, challenging the ability to draw strong conclusions from population synthesis models at this stage.

The situation is even more complex when considering that even our large suite of models (560 realizations) still only represents a subset of the overall uncertainties in population synthesis studies. For example, even our models do not account for uncertainties in the stellar evolution tracks (e.g., Laplace et al. 2020; Agrawal et al. 2020), internal mixing (e.g., Schootemeijer et al. 2019), stellar rotation (e.g., de Mink & Mandel 2016; Mapelli et al. 2020), the CE phase (e.g., Klencki et al. 2021; Ivanova et al. 2020; Marchant et al. 2021), SNe natal kicks and physics (e.g., Belczyński & Bulik 1999; Shao & Li 2018; Giacobbo & Mapelli 2019), and initial distribution functions (e.g. de Mink & Belczynski 2015; Moe & Di Stefano 2017; Klencki et al. 2018). In addition, there are $\mathcal{S}(Z, z)$ model variations that we did not explore, including alternative analytical prescriptions (e.g. Chruslinska et al. 2019; Tang et al. 2020), prescriptions derived from cosmological (zoom-in) simulations (e.g., FIRE, Illustris, EAGLE, Millennium; Lamberts et al. 2016; Mapelli et al. 2017; du Buisson et al. 2020; Chu et al. 2021; van Son et al. 2021), and prescriptions derived more directly from observations (e.g. Chruslinska & Nelemans 2019; Chrušlińska et al. 2021). In addition, our work focuses on the formation of DCO mergers from the isolated binary evolution channel, but a significant fraction of mergers might originate from alternative formation channels, which typically suffer from

similar or additional modelling uncertainties (see [Gerosa & Fishbach 2021](#); [Mandel & Broekgaarden 2021](#); [Mapelli 2021](#), for reviews).

Second, the large impact on the DCO merger rates and distribution shapes indicate that comparing population synthesis simulations to GW detections can aid in learning about the lives, evolution and deaths of massive stars throughout cosmic history when the uncertainties are sufficiently understood. In particular, we have shown that simultaneously modelling uncertainties from stellar evolution and $\mathcal{S}(Z, z)$ on the rates and distribution shapes of all three DCO merger types helps reveal areas within binary population synthesis simulations that have the potential to aid in learning about massive stars throughout cosmic history from gravitational-wave observations. For example, in §3.2 we found that the BHBH merger rate densities predicted from binary population synthesis are relatively sensitive to our explored $\mathcal{S}(Z, z)$ uncertainties and might thereby present a good test bed to constrain $\mathcal{S}(Z, z)$ prescriptions. Similarly, we found that the NSNS merger rate densities might be a good test bed to constrain stellar evolution models. These findings are in agreement with earlier work by [Chruslinska et al. \(2019\)](#); [Neijssel et al. \(2019\)](#); [Tang et al. \(2020\)](#); [Santoliquido et al. \(2021\)](#). Simultaneously comparing the predictions for all three DCO flavors can further aid in constraining models, as shown in Figure 2 for the merger rate densities. Another example is that we showed in §3.4 that the BHNS distribution shapes might particularly be a good test bed for stellar evolution models (Figure 5), whilst the shape of the mass distributions of BHBH detections have the potential to rule out certain $\mathcal{S}(Z, z)$ models (Figure 4). We emphasize that understanding which uncertainty these different DCO observable properties are sensitive to also aids in understanding where to best spend computational time and work in the simulations. We showed in §3.4.2 that some population synthesis results are robust under our 560 model variations. These features are important when identifying distinguishable characteristics of the isolated binary evolution channel compared to other formation channels. Constraining and learning from population synthesis models can be further aided in the future by adding additional constraints including those from redshift dependent rates and redshift dependent distribution shapes (cf. [Santoliquido et al. 2021](#); [Chu et al. 2021](#)) and additional observational constraints including those from X-ray binaries ([Belczynski et al. 2020](#); [Vinciguerra et al. 2020](#)), galactic pulsar binaries (e.g. [Kruckow et al. 2018](#); [Vigna-Gómez et al. 2018](#); [Chattopadhyay et al. 2020, 2021](#)), short gamma-ray bursts (e.g. [Shao & Li 2018](#); [Zevin et al. 2020](#)), or GW detectors beyond LVK (e.g. [Ng et al. 2021](#); [Wagg et al. 2021](#)).

4.2 Comparison with GW mergers

At the time of writing the LIGO and Virgo collaboration published two GW catalogs (GWTC-1 and GWTC-2) containing a total of 47 events with false alarm rate $\lesssim 1 \text{ yr}^{-1}$, consisting of 44 confident BHBH mergers, two confident NSNS mergers and one source that could be a BHNS candidate or a BHBH ([Abbott et al. 2019, 2021b,c](#)). If we scale the numbers from the predicted merger ratios in Figure 3 to a BHBH rate of 44 detections to obtain a very rough estimate, we find that for this number of BHBH mergers we predict between 0–15 BHNS and 0–2 NSNS detections. The BHNS and NSNS rates on the higher end are from the models with no BH natal kicks and/or low SN natal kicks (models P, Q and R). Although the O3a run did not contain confident BHNS detections, the LIGO-Virgo collaboration recently announced the first two observations of BHNS mergers from the O3b catalog ([Abbott et al. 2021d](#)), consistent with our broad range of expected BHNS detections.

Furthermore, in Figure 2 we showed the subset of model realiza-

tions that result in merger rate densities that match the ranges for the intrinsic merger rate densities based on the detected GW events. We found that the majority of our 560 model realizations overestimate the BHBH merger rate and underestimate the NSNS merger rate. This is consistent with findings by other studies on isolated binary evolution populations. For example, [Giacobbo & Mapelli \(2018\)](#); [Chruslinska et al. \(2018\)](#) and [Santoliquido et al. \(2021\)](#) also find that their predictions typically underestimate the NSNS merger rate (and overestimates the BHBH rate), and indicate that matching the observed NSNS rate required higher values for α_{CE} or smaller SN natal kicks. Further improved simulations and tighter constraints from observations will provide a test bed for these findings. Alternatively, the underestimated NSNS rate might be explained if a large fraction of events formed through different channels compared to the isolated binary evolution channel, examples include formation in dynamical clusters, active galactic nucleus disks, or higher multiple systems see for a review of the predicted rates from these channels [Mandel & Broekgaarden \(2021\)](#).

Besides the merger rates, the current GW catalogs also provide the masses of detected merging compact objects (e.g. [Abbott et al. 2021c](#)). We mentioned several examples of how the mass distributions of our 560 model variations compare to observed GW systems in (§3.4). When we compare the detected masses from [Abbott et al. \(2019, 2021b\)](#) and [Abbott et al. \(2021c\)](#) with our model predictions (§3.4) we find that many of the GW detected systems fall within our predicted distribution percentiles. A full comparison with the LVK data is out of scope of this paper, but we will mention an example of GW source properties that our models cannot explain below. First, we find that most of our model realizations do not match the several detected BHBH mergers with individual masses $\gtrsim 42 \text{ M}_\odot$. Examples include GW190521 with $m_{\text{BH},1} = 95.3^{+28.7}_{-18.9}$ and $m_{\text{BH},2} = 69.0^{+22.7}_{-23.1}$ and GW190929 with $m_{\text{BH},1} = 80.8^{+33.0}_{-33.2}$, which exceed the maximum possible BH mass of $\approx 42 \text{ M}_\odot$ in all our models except O, where we do not assume a pair-instability SN prescription. However, even in model O we find that 99% of the primary BHs are expected to be below $\lesssim 88 \text{ M}_\odot$ (Figure 4), challenging the inferred BH mass in GW190521. The origin of these massive BHs systems is still under debate. One possibility is that our assumed upper limit of BH masses around 42 M_\odot due to pair-instability SNe is higher (e.g. [Spera & Mapelli 2017](#); [Farmer et al. 2019, 2020](#); [Costa et al. 2021](#); [Woosley & Heger 2021](#)). Another possibility is that BHBH mergers with massive BHs formed through other channels than the isolated binary evolution channel such as via hierarchical dynamical mergers (e.g. [Anagnostou et al. 2020](#)), stellar mergers (e.g. [Spera et al. 2019](#); [Di Carlo et al. 2020](#); [Kremer et al. 2020](#)), triples (e.g. [Vigna-Gómez et al. 2021](#)) or in AGN disks (e.g. [Secunda et al. 2020](#)). See for a more detailed discussion, for example, [Abbott et al. \(2020d\)](#) and [Kimball et al. \(2021\)](#) and references therein.

5 CONCLUSIONS

In this study we have predicted the merger rates and properties of BHBH, BHNS and NSNS mergers detectable with current gravitational-wave detectors. We examined the impact of two key modelling uncertainties in population synthesis studies: uncertainties arising from massive binary star evolution and from not knowing the metallicity-specific star formation history, $\mathcal{S}(Z, z)$, which is a function of birth metallicity Z and redshift z . We accomplish this by simulating populations of binaries over a grid of 53 birth metallicities and taking into account the $\mathcal{S}(Z, z)$ and detection probability for observations by a ground-based GW detectors at design sensitivity.

We present 560 publicly available model realizations (the largest of its kind), and investigate how the two key uncertainties impact the predicted merger rates and distribution shapes for the masses, delay times and birth metallicities. Our main findings within the context of the variations we considered are summarized below.

1. Merging DCO formation yields: We find for all our stellar evolution variations that the merging BHBH (BHNS) formation yield is typically a rapidly decreasing function of metallicity at $Z \gtrsim Z_{\odot}/10$ ($\gtrsim Z_{\odot}/2$) as a result of [line-driven stellar winds](#). The NSNS yield, on the other hand, is [relatively independent of metallicity](#). We find that the formation yield of BHBH mergers for $Z \lesssim Z_{\odot}/10$ is remarkably constant over massive binary-star models and Z . The formation yields of BHNS and NSNS mergers are impacted over the full metallicity range by our massive binary-star assumptions (see Figure 1).

2. Merging DCO merger rates: We find that the calculated intrinsic and detectable merger rate densities (\mathcal{R}_m^0 and \mathcal{R}_{det} , respectively) can be impacted by factors $200\times$ – $1300\times$ by combined uncertainties in stellar evolution and $S(Z, z)$. The 560 model variations lead to estimated merger rate densities in the ranges $\mathcal{R}_m^{0,\text{BHBH}} = 3.8\text{--}810 \text{ Gpc}^{-3} \text{ yr}^{-1}$ and $\mathcal{R}_{\text{det}}^{\text{BHBH}} = 17\text{--}12000 \text{ yr}^{-1}$ for BHBHs, $\mathcal{R}_m^{0,\text{BHNS}} = 2.2\text{--}830 \text{ Gpc}^{-3} \text{ yr}^{-1}$ and $\mathcal{R}_{\text{det}}^{\text{BHNS}} = 0.36\text{--}180 \text{ yr}^{-1}$ for BHNSs, and $\mathcal{R}_m^{0,\text{NSNS}} = 0.32\text{--}330 \text{ Gpc}^{-3} \text{ yr}^{-1}$ and $\mathcal{R}_{\text{det}}^{\text{NSNS}} = 0\text{--}12 \text{ yr}^{-1}$ for NSNSs. In particular, we found that the estimated BHBH merger rate densities are relatively sensitive to our explored $S(Z, z)$ uncertainties and might thereby present a good test bed to constrain $S(Z, z)$ prescriptions. On the other hand, we found that the NSNS merger rate densities are sensitive to stellar evolution models and therefore might be a good test bed to constrain stellar evolution models. For BHNSs we find that the calculated rates are significantly impacted by both uncertainties, (see Figure 2 and 3).

3. Merging DCO mass distribution shapes: We show in Figure 4, 5 and 6 the impact from the massive binary-star and $S(Z, z)$ variations on the (normalized) shape of the detectable DCO mass distributions (chirp mass, individual component masses and mass ratios). We find that the shape of the detectable BHBH mass distributions are dominated by the $S(Z, z)$ variations. On the other hand, we find that the shape of the BHNS mass distributions are dominated by variations in binary stellar evolution. For NSNS mergers we find that the mass distribution shapes are impacted by both variations in stellar evolution and $S(Z, z)$. This implies that the BHBH and NSNS mass distributions can provide a test bed for $S(Z, z)$ models, whilst BHNS and NSNS provide a test bed for stellar evolution assumptions.

4. Merging DCO t_{delay} and Z distribution shapes: We show in Figure 7, 8 and 9 the impact from binary stellar evolution and $S(Z, z)$ variations on the delay time and birth metallicity distributions calculated in our models. We find that the $S(Z, z)$ has a significant impact for BHBH, BHNS and NSNS on the shape of the delay time and metallicity distributions. Several stellar evolution models, including those affecting stellar winds can also significantly impact the delay time and Z distribution shapes.

5. Consistent features among our 560 model variations: We find several features of DCO mass distributions that are consistent among all 560 model variations. This includes that 95 percent of the detectable BHBH mergers have mass ratios $q_f \lesssim 4$ in all 560 model variations, as well as that 99 percent of BHBH mergers have mass ratios $q_f \lesssim 6$ (Figure 4). We find that across all 560 model variations,

we estimate that less than 5 percent of the detectable merging BHNS systems have BH masses $m_{\text{BH}} \gtrsim 18 M_{\odot}$ (Figure 5). We also find that all 560 model variations predict that the NSs in detectable NSNS mergers to typically have lower NS masses compared to the NSs in detectable BHNS mergers (Figure 5 and 6). In addition, more than 95 percent of BHBH mergers are expected to contain a BH with a mass $\gtrsim 8 M_{\odot}$ (Figure 4).

All in all, our results highlight the importance of considering the uncertainty from both the stellar evolution and $S(Z, z)$ uncertainties when studying and evaluating population synthesis simulations of BHNS, BHBH and NSNS mergers.

ACKNOWLEDGEMENTS

Simulations in this paper made use of the COMPAS rapid binary population synthesis code, which is freely available at [http://github.com/TeamCOMPAS/COMPAS](https://github.com/TeamCOMPAS/COMPAS) (Team COMPAS: J. Riley et al. 2021) including work based on (Stevenson et al. 2017; Barrett et al. 2018; Vigna-Gómez et al. 2018; Broekgaarden et al. 2019). The simulations performed in this work were simulated with a COMPAS version that predates the publicly available code. Our version of the code is most similar to version 02.13.01 of the publicly available COMPAS code. Requests for the original code can be made to the lead author. The authors used STROOPWAFEL from (Broekgaarden et al. 2019), publicly available at <https://github.com/FloorBroekgaarden/STROOPWAFEL>¹⁰.

The authors thank everyone in the COMPAS collaboration and Berger Time-Domain Group for help. In addition, the authors thank the Harvard FAS research computing group for technical support on the simulations and high performance computing part of the research. Lastly, FSB wants to acknowledge the amount of serendipity and privilege that was involved to end up pursuing this astronomy research. The Berger Time-Domain Group is supported in part by NSF grant AST-1714498 and NASA grant NNX15AE50G. Some of the authors are supported by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (OzGrav), through project number CE170100004. FSB is supported in part by the Prins Bernard Cultuurfonds studiebeurs 2021. IM is a recipient of the Australian Research Council Future Fellowship FT190100574. A.V.-G. acknowledges funding support by the Danish National Research Foundation (DNRF132). SJ, LvS and SdM acknowledge funding from the Netherlands Organisation for Scientific Research (NWO), as part of the Vidi research program BinWaves (project number 639.042.728) and the European Union’s Horizon 2020 research and innovation program from the European Research Council (ERC, Grant agreement No. 715063). LvS, TW and SdM acknowledges support by the National Science Foundation under Grant No. (NSF 2009131). This research has made use of NASA’s Astrophysics Data System Bibliographic Services.

DATA AVAILABILITY

All data used in this work is publicly available on Zenodo at Broekgaarden (2021a, BHBH) Broekgaarden (2021b, BHNS) and Broekgaarden (2021c, NSNS). All code to reproduce the results

¹⁰ For the latest pip installable version of STROOPWAFEL please contact Floor Broekgaarden.

and figures in this paper (and additional figures) are publicly available on Github at <https://github.com/FloorBroekgaarden/Double-Compact-Object-Mergers>.

SOFTWARE

The authors made use of Python from the Python Software Foundation. Python Language Reference, version 3.6. Available at <http://www.python.org> (van Rossum 1995). In addition the following Python packages were used: matplotlib (Hunter 2007), NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), ipython/jupyter (Perez & Granger 2007; Kluyver et al. 2016), pandas (Wes McKinney 2010), Seaborn (Waskom & the seaborn development team 2020), Astropy (Astropy Collaboration et al. 2018) and hdf5 (Collette 2013).

REFERENCES

- Abbott B. P., et al., 2016, *Living Reviews in Relativity*, **19**, 1
 Abbott B. P., et al., 2017, *Phys. Rev. Lett.*, **119**, 161101
 Abbott B. P., et al., 2018, *Living Reviews in Relativity*, **21**, 3
 Abbott B. P., et al., 2019, *Physical Review X*, **9**, 031040
 Abbott R., et al., 2020a, arXiv e-prints, p. arXiv:2004.08342
 Abbott B. P., et al., 2020b, *ApJ*, **892**, L3
 Abbott R., et al., 2020c, *ApJ*, **896**, L44
 Abbott R., et al., 2020d, *ApJ*, **900**, L13
 Abbott R., Abbott T. D., Abraham S., Acernese F., et al., 2021a, arXiv e-prints, p. arXiv:2108.01045
 Abbott R., et al., 2021b, *Physical Review X*, **11**, 021053
 Abbott R., et al., 2021c, *ApJ*, **913**, L7
 Abbott R., et al., 2021d, *ApJ*, **915**, L5
 Agrawal P., Hurley J., Stevenson S., Szécsi D., Flynn C., 2020, *MNRAS*, **497**, 4549
 Anagnostou O., Trenti M., Melatos A., 2020, arXiv e-prints, p. arXiv:2010.06161
 Astropy Collaboration et al., 2018, *AJ*, **156**, 123
 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., 2021, *A&A*, **647**, A153
 Belczynski K., Bulik T., 1999, *A&A*, **346**, 91
 Belczynski K., Kalogera V., Rasio F. A., Taam R. E., Zezas A., Bulik T., Maccarone T. J., Ivanova N., 2008, *ApJS*, **174**, 223
 Belczynski K., Dominik M., Bulik T., O’Shaughnessy R., Fryer C., Holz D. E., 2010, *ApJ*, **715**, L138
 Belczynski K., et al., 2020, *A&A*, **636**, A104
 Belczynski K., et al., 2021, arXiv e-prints, p. arXiv:2108.10885
 Bhattacharya D., van den Heuvel E. P. J., 1991, *Phys. Rep.*, **203**, 1
 Boco L., Lapi A., Chruslinska M., Donevski D., Sicilia A., Danese L., 2021, *ApJ*, **907**, 110
 Bouffanais Y., Mapelli M., Gerosa D., Di Carlo U. N., Giacobbo N., Berti E., Baibhav V., 2019, *ApJ*, **886**, 25
 Broekgaarden F. S., 2021a, BHBB simulations from: Impact of Massive Binary Star and Cosmic Evolution on Gravitational Wave Observations II: Double Compact Object Mergers, doi:10.5281/zenodo.5651073, <https://doi.org/10.5281/zenodo.5651073>
 Broekgaarden F. S., 2021b, BHNS simulations from: Impact of Massive Binary Star and Cosmic Evolution on Gravitational Wave Observations II: Double Compact Object Mergers, doi:10.5281/zenodo.5178777, <https://doi.org/10.5281/zenodo.5178777>
 Broekgaarden F. S., 2021c, NSNS simulations from: Impact of Massive Binary Star and Cosmic Evolution on Gravitational Wave Observations II: Double Compact Object Mergers, doi:10.5281/zenodo.5189849, <https://doi.org/10.5281/zenodo.5189849>
 Broekgaarden F. S., Berger E., 2021, *ApJ*, **920**, L13
 Broekgaarden F. S., et al., 2019, *MNRAS*, **490**, 5228
 Broekgaarden F. S., et al., 2021, *MNRAS*,
 Chatopadhyay D., Stevenson S., Hurley J. R., Rossi L. J., Flynn C., 2020, *MNRAS*, **494**, 1587
 Chatopadhyay D., Stevenson S., Hurley J. R., Bailes M., Broekgaarden F., 2021, *MNRAS*, **504**, 3682
 Chruslinska M., Nelemans G., 2019, *MNRAS*, **488**, 5300
 Chruslinska M., Belczynski K., Klencki J., Benacquista M., 2018, *MNRAS*, **474**, 2937
 Chruslinska M., Nelemans G., Belczynski K., 2019, *MNRAS*, **482**, 5012
 Chruślińska M., Nelemans G., Boco L., Lapi A., 2021, *MNRAS*,
 Chu Q., Yu S., Lu Y., 2021, *MNRAS*,
 Collette A., 2013, Python and HDF5. O'Reilly
 Costa G., Bressan A., Mapelli M., Marigo P., Iorio G., Spera M., 2021, *MNRAS*, **501**, 4514
 de Kool M., 1990, *ApJ*, **358**, 189
 de Mink S. E., Belczynski K., 2015, *ApJ*, **814**, 58
 de Mink S. E., Mandel I., 2016, *MNRAS*, **460**, 3545
 Delgado A. J., Thomas H. C., 1981, *A&A*, **96**, 142
 Di Carlo U. N., Mapelli M., Bouffanais Y., Giacobbo N., Santoliquido F., Bressan A., Spera M., Haardt F., 2020, *MNRAS*, **497**, 1043
 Dominik M., Belczynski K., Fryer C., Holz D. E., Berti E., Bulik T., Mandel I., O’Shaughnessy R., 2012, *ApJ*, **759**, 52
 Dominik M., et al., 2015, *ApJ*, **806**, 263
 du Buisson L., et al., 2020, *MNRAS*, **499**, 5941
 Farmer R., Renzo M., de Mink S. E., Marchant P., Justham S., 2019, *ApJ*, **887**, 53
 Farmer R., Renzo M., de Mink S. E., Fishbach M., Justham S., 2020, *ApJ*, **902**, L36
 Fishbach M., Kalogera V., 2021, *ApJ*, **914**, L30
 Fragione G., 2021, arXiv e-prints, p. arXiv:2110.09604
 Franciolini G., et al., 2021, arXiv e-prints, p. arXiv:2105.03349
 Fryer C. L., Belczynski K., Wiktorowicz G., Dominik M., Kalogera V., Holz D. E., 2012, *ApJ*, **749**, 91
 Furlong M., et al., 2015, *MNRAS*, **450**, 4486
 Gerosa D., Fishbach M., 2021, *Nature Astronomy*, **5**, 749
 Giacobbo N., Mapelli M., 2018, *MNRAS*, **480**, 2011
 Giacobbo N., Mapelli M., 2019, *MNRAS*, **482**, 2234
 Giacobbo N., Mapelli M., Spera M., 2018, *MNRAS*, **474**, 2959
 Hamann W. R., Koesterke L., 1998, *A&A*, **335**, 1003
 Hamann W. R., et al., 2019, *A&A*, **625**, A57
 Harris C. R., et al., 2020, *Nature*, **585**, 357–362
 Hinshaw G., et al., 2013, *ApJS*, **208**, 19
 Hunter J. D., 2007, *Computing in Science and Engineering*, **9**, 90
 Hurley J. R., Pols O. R., Tout C. A., 2000, *MNRAS*, **315**, 543
 Hurley J. R., Tout C. A., Pols O. R., 2002, *MNRAS*, **329**, 897
 Im M., et al., 2017, *ApJ*, **849**, L16
 Ivanova N., Justham S., Ricker P., 2020, Common Envelope Evolution, doi:10.1088/2514-3433/abb6f0.
 Kalogera V., Webbink R. F., 1996, *ApJ*, **458**, 301
 Kimball C., et al., 2021, *ApJ*, **915**, L35
 Klencki J., Moe M., Gladysz W., Chruslinska M., Holz D. E., Belczynski K., 2018, *A&A*, **619**, A77
 Klencki J., Nelemans G., Istrate A. G., Chruslinska M., 2021, *A&A*, **645**, A54
 Kluyver T., et al., 2016, in ELPUB. pp 87–90
 Kremer K., et al., 2020, *ApJ*, **903**, 45
 Kruckow M. U., Tauris T. M., Langer N., Kramer M., Izzard R. G., 2018, *MNRAS*, **481**, 1908
 LIGO Scientific Collaboration et al., 2015, *Classical and Quantum Gravity*, **32**, 074001
 Lamberts A., Garrison-Kimmel S., Clausen D. R., Hopkins P. F., 2016, *MNRAS*, **463**, L31
 Langer N., Norman C. A., 2006, *ApJ*, **638**, L63
 Laplace E., Götzberg Y., de Mink S. E., Justham S., Farmer R., 2020, *A&A*, **637**, A6
 Linden T., Kalogera V., Sepinsky J. F., Prestwich A., Zezas A., Gallagher J. S., 2010, *ApJ*, **725**, 1984

- Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, *MNRAS*, **456**, 2140
- Madau P., Dickinson M., 2014, *ARA&A*, **52**, 415
- Madau P., Fragos T., 2017, *ApJ*, **840**, 39
- Mandel I., Broekgaarden F. S., 2021, arXiv e-prints, p. [arXiv:2107.14239](https://arxiv.org/abs/2107.14239)
- Mandel I., Farmer A., 2018, arXiv e-prints, p. [arXiv:1806.05820](https://arxiv.org/abs/1806.05820)
- Mapelli M., 2021, Formation Channels of Single and Binary Stellar-Mass Black Holes. p. 4, doi:[10.1007/978-981-15-4702-7_16-1](https://doi.org/10.1007/978-981-15-4702-7_16-1)
- Mapelli M., Giacobbo N., Ripamonti E., Spera M., 2017, *MNRAS*, **472**, 2422
- Mapelli M., Spera M., Montanari E., Limongi M., Chieffi A., Giacobbo N., Bressan A., Bouffanais Y., 2020, *ApJ*, **888**, 76
- Marchant P., Langer N., Podsiadlowski P., Tauris T. M., de Mink S., Mandel I., Moriya T. J., 2017, *A&A*, **604**, A55
- Marchant P., Renzo M., Farmer R., Pappas K. M. W., Taam R. E., de Mink S. E., Kalogera V., 2019, *ApJ*, **882**, 36
- Marchant P., Pappas K. M. W., Gallegos-Garcia M., Berry C. P. L., Taam R. E., Kalogera V., Podsiadlowski P., 2021, *A&A*, **650**, A107
- Massevitch A., Yungelson L., 1975, Mem. Soc. Astron. Italiana, **46**, 217
- McKernan B., Ford K. E. S., O'Shaughnessy R., 2020, *MNRAS*, **498**, 4088
- Moe M., Di Stefano R., 2017, *The Astrophysical Journal Supplement Series*, **230**, 15
- Neijssel C. J., et al., 2019, *MNRAS*, **490**, 3740
- Ng K. K. Y., Vitale S., Farr W. M., Rodriguez C. L., 2021, *ApJ*, **913**, L5
- Paczyński B., Sienkiewicz R., 1972, *Acta Astron.*, **22**, 73
- Panter B., Heavens A. F., Jimenez R., 2004, *MNRAS*, **355**, 764
- Perez F., Granger B. E., 2007, *Computing in Science and Engineering*, **9**, 21
- Pfahl E., Rappaport S., Podsiadlowski P., 2002, *ApJ*, **571**, L37
- Podsiadlowski P., Langer N., Poelarends A. J. T., Rappaport S., Heger A., Pfahl E., 2004, *ApJ*, **612**, 1044
- Rastello S., Mapelli M., Di Carlo U. N., Giacobbo N., Santoliquido F., Spera M., Ballone A., Iorio G., 2020, *MNRAS*, **497**, 1563
- Renzo M., et al., 2019, *A&A*, **624**, A66
- Román-Garza J., et al., 2021, *ApJ*, **912**, L23
- Safarzadeh M., Berger E., Ng K. K. Y., Chen H.-Y., Vitale S., Whittle C., Scannapieco E., 2019, *ApJ*, **878**, L13
- Sana H., 2017, in Eldridge J. J., Bray J. C., McClelland L. A. S., Xiao L., eds, IAU Symposium Vol. 329, The Lives and Death-Throes of Massive Stars. pp 110–117 ([arXiv:1703.01608](https://arxiv.org/abs/1703.01608)), doi:[10.1017/S1743921317003209](https://doi.org/10.1017/S1743921317003209)
- Sander A. A. C., Vink J. S., 2020, *MNRAS*, **499**, 873
- Santoliquido F., Mapelli M., Giacobbo N., Bouffanais Y., Artale M. C., 2021, *MNRAS*, **502**, 4877
- Schneider F. R. N., Izzard R. G., Langer N., de Mink S. E., 2015, *ApJ*, **805**, 20
- Schootemeijer A., Langer N., Grin N. J., Wang C., 2019, *A&A*, **625**, A132
- Secunda A., et al., 2020, *ApJ*, **903**, 133
- Shao Y., Li X.-D., 2018, *ApJ*, **867**, 124
- Shenar T., et al., 2019, *A&A*, **627**, A151
- Smarr L. L., Blandford R., 1976, *ApJ*, **207**, 574
- 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., Giacobbo N., Trani A. A., Bressan A., Costa G., 2019, *MNRAS*, **485**, 889
- Srinivasan G., 1989, *A&ARv*, **1**, 209
- Stevenson S., Vigna-Gómez A., Mandel I., Barrett J. W., Neijssel C. J., Perkins D., de Mink S. E., 2017, *Nature Communications*, **8**, 14906
- Stevenson S., Sampson M., Powell J., Vigna-Gómez A., Neijssel C. J., Szécsi D., Mandel I., 2019, *ApJ*, **882**, 121
- Strolger L.-G., et al., 2004, *ApJ*, **613**, 200
- Tang P. N., Eldridge J. J., Stanway E. R., Bray J. C., 2020, *MNRAS*, **493**, L6
- Tauris T. M., et al., 2017, *ApJ*, **846**, 170
- Team COMPAS: J. Riley et al., 2021, arXiv e-prints, p. [arXiv:2109.10352](https://arxiv.org/abs/2109.10352)
- van Rossum G., 1995, Technical Report CS-R9526, Python tutorial. Centrum voor Wiskunde en Informatica (CWI), Amsterdam
- van Son L. A. C., et al., 2021, arXiv e-prints, p. [arXiv:2110.01634](https://arxiv.org/abs/2110.01634)
- Vigna-Gómez A., et al., 2018, *MNRAS*, **481**, 4009
- Vigna-Gómez A., Toonen S., Ramirez-Ruiz E., Leigh N. W. C., Riley J., Haster C.-J., 2021, *ApJ*, **907**, L19
- Vinciguerra S., et al., 2020, *MNRAS*, **498**, 4705
- Vink J. S., 2017, *A&A*, **607**, L8
- Vink J. S., de Koter A., 2005, *A&A*, **442**, 587
- Virtanen P., et al., 2020, *Nature Methods*, **17**, 261
- Wagg T., et al., 2021, *ApJ*, **913**, L5
- Waskom M., the seaborn development team 2020, mwaskom/seaborn, doi:[10.5281/zenodo.592845](https://doi.org/10.5281/zenodo.592845), <https://doi.org/10.5281/zenodo.592845>
- Webbink R. F., 1984, *ApJ*, **277**, 355
- Wes McKinney 2010, in Stéfan van der Walt Jarrod Millman eds, Proceedings of the 9th Python in Science Conference. pp 56 – 61, doi:[10.25080/Majora-92bf1922-00a](https://doi.org/10.25080/Majora-92bf1922-00a)
- Wootsey S. E., 2017, *ApJ*, **836**, 244
- Wootsey S. E., Heger A., 2021, *ApJ*, **912**, L31
- Xu X.-J., Li X.-D., 2010a, *ApJ*, **716**, 114
- Xu X.-J., Li X.-D., 2010b, *ApJ*, **722**, 1985
- Zevin M., Pankow C., Rodriguez C. L., Sampson L., Chase E., Kalogera V., Rasio F. A., 2017, *ApJ*, **846**, 82
- Zevin M., Kelley L. Z., Nugent A., Fong W.-f., Berry C. P. L., Kalogera V., 2020, *ApJ*, **904**, 190
- Zevin M., et al., 2021, *ApJ*, **910**, 152

APPENDIX A: ADDITIONAL TABLES AND FIGURES

In this appendix we show a Table and three additional figures that are mentioned throughout the main text.

First, Table A1 summarizes the 28 $\mathcal{S}(Z, z)$ models that are explored in this study. More details about the assumptions in each model can be found in Broekgaarden et al. (2021) and Neijssel et al. (2019).

Second, in Figure A1 we show similar to Figure 1 the formation yield of DCO systems that merge in a Hubble time as a function of the birth metallicity. However, in this figure we show the relative formation yields between BHNS to BHBH (left panel) and between NSNS and BHBH (right panel). It can be seen that in most models the formation yield of BHNS mergers exceeds that of BHBH around birth metallicities $Z \gtrsim Z/2$. The yield of NSNS mergers surpasses the BHBH yield for most models around Z/Z_\odot .

Third, in Figure A2 we visualize two of our $\mathcal{S}(Z, z)$ model variations, xyz = 312 and xyz = 231. The xyz = 312 (xyz = 231) model is the $\mathcal{S}(Z, z)$ distribution that correspond to one of the highest (lowest) average Z and results in one of the lowest (highest) BHBH merger rates in Figure 2.

Last, Figure A3 shows a zoom in on one of the percentile distribution panels from Figure 4 (the top row, second column panel). We highlight in this figure the labels of the $\mathcal{S}(Z, z)$ models that correspond to the order of the horizontal percentile bars.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.

xyz index	SFRD [x]	GSMF [y]	MZR [z]
000	preferred phenomenological model from Neijssel et al. (2019)		
1	Madau & Dickinson (2014)	Panter et al. (2004)	Langer & Norman (2006)
2	Strolger et al. (2004)	Furlong et al. (2015) single Schechter	Langer & Norman (2006) + offset
3	Madau & Fragos (2017)	Furlong et al. (2015) double Schechter	Ma et al. (2016)

Table A1. List of the assumptions for the metallicity-specific star formation rate, $\mathcal{S}(Z, z)$, models that we explore in this study. 27 $\mathcal{S}(Z, z)$ models are obtained by combining a star formation rate density (SFRD) with a galaxy stellar mass function (GSMF) and mass-metallicity relation (MZR). See §2.2 and paper I for more details. The models are named in the convention xyz, where x, y, z $\in [1, 2, 3]$ are the index numbers for the models used for the SFRD, GSMF and MZR, respectively. For example, the combination of using the Madau & Dickinson (2014) SFRD with the Panter et al. (2004) GSMF and the Ma et al. (2016) MZR is labeled 113. The preferred phenomenological model from Neijssel et al. (2019) has the label 000 and is not a specific combination but a parameterized model that is built to be flexible and is fitted to match the observed BHBH merger rate and chirp mass distribution from the first two runs of LIGO and Virgo.

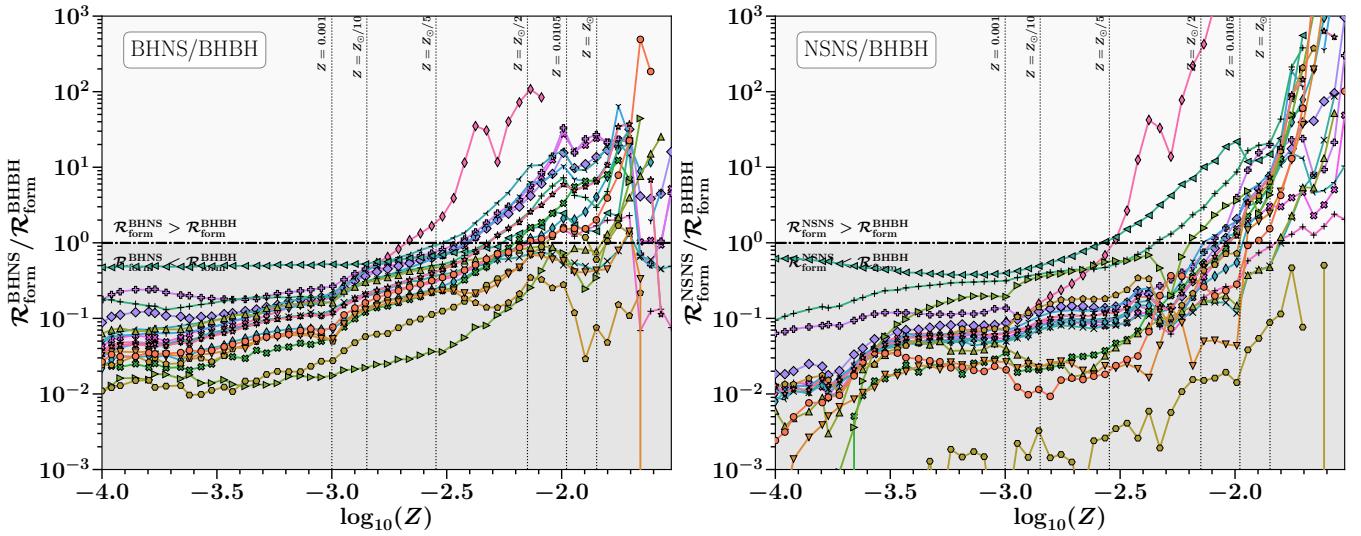
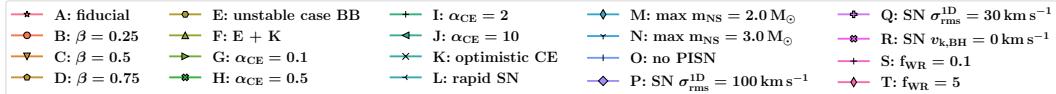


Figure A1. The ratio of formation yields of merging double compact objects per solar mass of star formation ($dN_{\text{form}}/dM_{\text{SFR}}$) as a function of birth metallicity (Z). The left panel shows the ratio of BHNS to BHBH formation yields and the right panel shows the ratio of NSNS to BHBH formation yields. Each color and marker type corresponds to one of the 20 binary population synthesis models explored in this study (Table 1). Vertical dotted lines show fixed Z values. The marker points show the Z grid points that we simulated with COMPAS . Scatter points are missing where we divide by zero.

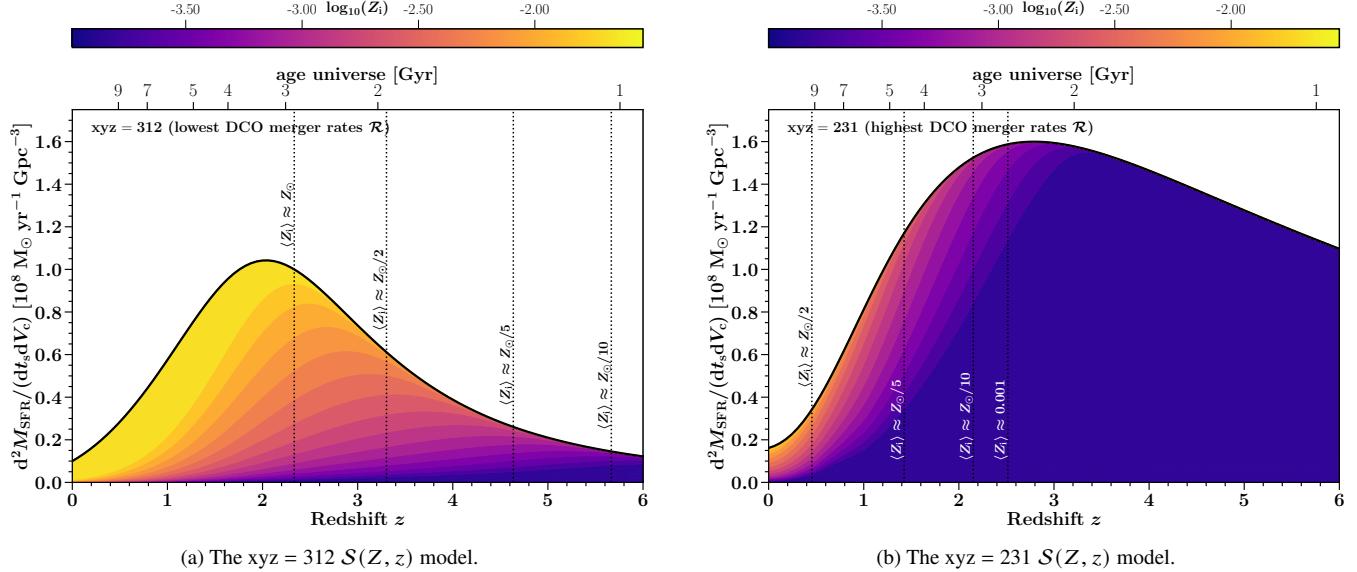


Figure A2. The SFRD as a function of redshift and metallicity for two $\mathcal{S}(Z, z)$ models from Table A1 used in this study. The left panel shows the $xyz = 312$ model whilst the right panel shows the $xyz = 231$ $\mathcal{S}(Z, z)$ model. In color we show the build up of the SFRD from the contribution by different Z , where for visual purposes we divided the SFRD in a few metallicity bins.

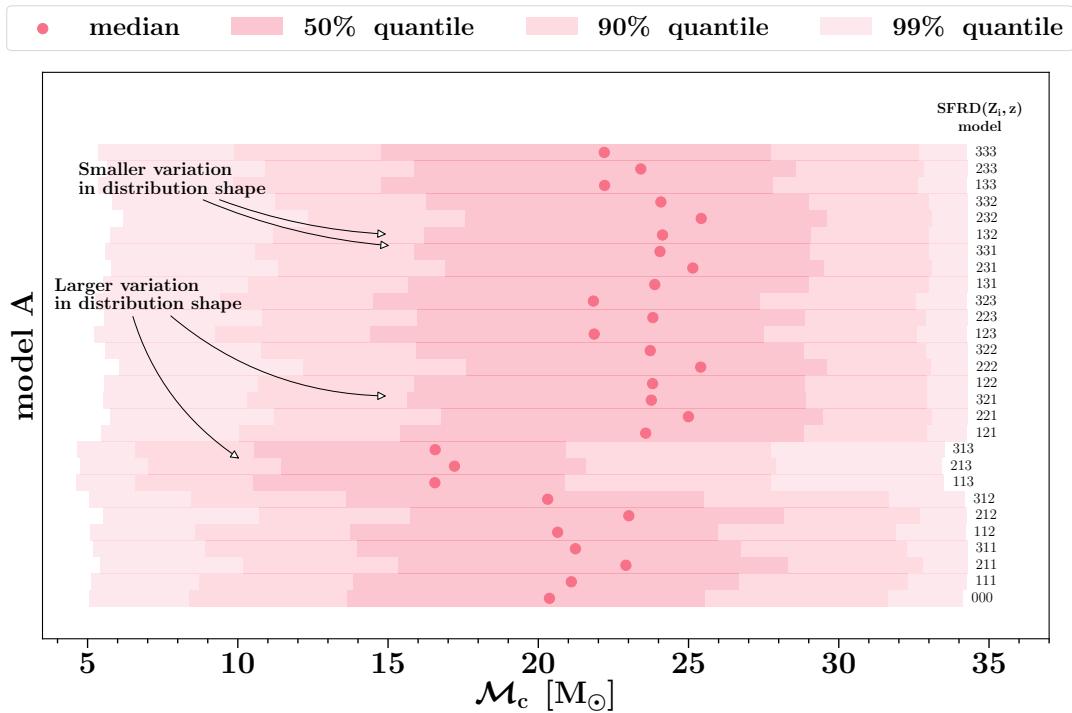


Figure A3. Zoom in on the chirp mass percentile panel for model A for detectable BHBH mergers at design sensitivity from Figure 4. Each of the 28 individual horizontal bars visualizes the median (scatter points), 50 , 90 and 99 (three shades) distribution percentiles. On the right of the horizontal bar we show the $\mathcal{S}(Z, z)$ model name as given in Table A1.