

## Gravitational wave sources in our Galactic backyard: Predictions for BHBH, BHNS and NSNS binaries in LISA

T. WAGG,<sup>1,2</sup> F.S. BROEKGAARDEN,<sup>1</sup> S.E. DE MINK,<sup>2,3,1</sup> N. FRANKEL,<sup>4</sup> L.A.C. VAN SON,<sup>1,3,2</sup> AND S. JUSTHAM<sup>5,3,2</sup>

<sup>1</sup>Center for Astrophysics — Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

<sup>2</sup>Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, 85741 Garching, Germany

<sup>3</sup>Anton Pannekoek Institute for Astronomy & Grappa, University of Amsterdam, Postbus 94249, 1090 GE Amsterdam, The Netherlands

<sup>4</sup>Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany

<sup>5</sup>School of Astronomy & Space Science, University of the Chinese Academy of Sciences, Beijing 100012, China

(Received October 10, 2021)

### ABSTRACT

We present predictions for the properties of the population of Galactic double black holes (BHBHs), black hole neutron stars (BHNSs) and double neutron stars (NSNSs) that will be detectable by the planned space-based gravitational wave detector LISA. We use rapid population synthesis to produce an extensive sample of double compact objects (DCOs) and combine this with an empirically-informed model to distribute them in a Milky Way-like galaxy based on their birth metallicity. We investigate the dependence of our results upon underlying physics assumptions by comparing the results of 20 physics variations that vary assumptions relating to mass transfer, common-envelope, supernova and wind mass loss physics. We predict that for a 4(10)-year mission, LISA will (typically) detect about 74(118) BHBHs, 42(71) BHNSs, 8(13) NSNSs. The predicted BHBH rate remains notably consistent under different physics assumptions, whilst in contrast the BHNS and NSNS rates each vary over 3 orders of magnitude. We discuss observable characteristics that could be used to distinguish the aforementioned DCOs from the more numerous double white dwarf population as well as for disentangling the BHBH, BHNS and NSNS populations from each other. We additionally assess the possibility of multi-messenger observations of pulsar populations by combining the capabilities of LISA and SKA.

*Keywords:* gravitational waves, LISA, black hole, neutron star, binary

### 1. INTRODUCTION

Since the first direct observation of gravitational waves (Abbott et al. 2016), the number of black hole (BH) and neutron star (NS) binaries observed by ground-based gravitational wave detectors has rapidly grown (Abbott et al. 2019, 2020b), offering exciting insights into the formation, lives and deaths of massive (binary) stars (e.g. Abbott et al. 2021).

The Laser Interferometer Space Antenna (LISA, Amaro-Seoane et al. 2017; Colpi et al. 2019) will provide observations in an entirely new regime of gravitational waves. LISA will observe at lower frequencies ( $10^{-5} \lesssim f/\text{Hz} \lesssim 10^{-1}$ ) than ground-based detectors and so will enable the study of sources that are imper-

ceptible by ground-based detectors, such as the mergers of supermassive black holes and extreme mass ratio inspirals (e.g. Begelman et al. 1980; Klein et al. 2016). Moreover, this frequency regime is also of interest for the detection of *local* stellar-mass double compact objects (DCO) millions of years before their merger. This presents an opportunity for both multimessenger detections to search for electromagnetic counterparts and multiband gravitational wave detections that can help to constrain binary characteristics (e.g. Sesana 2016; Gerosa et al. 2019). In addition, LISA will be able to measure the eccentricities of DCOs, which may yield further constraints on binary evolution, differentiate between formation channels and distinguish between DCO types (e.g. Nelemans et al. 2001; Breivik et al. 2016; Antonini et al. 2017; Rodriguez et al. 2018). Unlike ground-based detectors, LISA only detects stellar-mass sources in local galaxies, with the majority residing in

the Milky Way. These sources could be used as a probe for the structure of our galaxy (e.g. Korol et al. 2019).

Traditionally, predictions about the detection of stellar-mass sources with LISA focus on double white dwarf (WDWD) binaries, as they are abundantly present in our galaxy and are expected to be the dominant source of stellar-mass binaries that are detectable by LISA (Nelemans et al. 2001; Ruiter et al. 2010; Yu & Jeffery 2010; Nissanke et al. 2012; Korol et al. 2017; Lamberts et al. 2018). More recently, interest has grown in the detection of NS and BH binaries. Although they are more rare, LISA detections of these sources are potentially valuable for learning more about the evolution and endpoints of massive stars. In this paper we focus on making LISA predictions for double black hole binaries (BHBH), black hole neutron star binaries (BHNS) and double neutron star binaries (NSNS).

The detection of NSNSs in LISA could improve our understanding of many phenomena. Galactic NSNSs have been observed with electromagnetic signals for several decades (e.g. Hulse & Taylor 1975, see also refs. in Tauris et al. 2017; Vigna-Gómez et al. 2018) and more recently the mergers of NSNS binaries have been detected with ground-based gravitational wave detectors (e.g. Abbott et al. 2017a). A LISA detectable NSNS with a pulsar component close to merger would be ideal for connecting these population, as the binary could be observed from inspiral to merger. NSNS (and possible BHNS) binaries are useful sources for understanding the origin of r-process elements (e.g. Eichler et al. 1989) as well as the electromagnetic counterparts to gravitational wave signal such as kilonovae (e.g. Li & Paczyński 1998; Metzger 2017), short gamma-ray bursts (e.g. Berger 2014), radio emission (e.g. Hotokezaka et al. 2016) and neutrinos (e.g. Kyutoku et al. 2018).

BHBHs in the Milky Way present a greater observational challenge. To date, no BH has been observed in a binary with another compact object in the Milky Way and so LISA could provide the first detection of a Galactic BHBH. The only confirmed BHs in our galaxy have been discovered as components of X-ray binaries with companion stars (e.g. Bolton 1972; Webster & Murdin 1972). This sample of BHs has masses mainly constrained between  $5$  and  $10 M_{\odot}$  (Corral-Santana et al. 2016), a stark contrast to the more massive BHs observed with LIGO/Virgo that tend to have masses concentrated around  $30 M_{\odot}$  (Abbott et al. 2020b). These observations of X-ray binaries suggest the presence of a lower mass gap (from  $2-5 M_{\odot}$ ) in which there are no strong candidates for either black holes or neutron stars (Özel et al. 2010; Farr et al. 2011) but the gap's existence remains an open question (e.g. Kreidberg et al.

2012; Mandel & Müller 2020). Recently there has also been increased discussion over the maximum BH mass in our galaxy, with the claims of a  $70 M_{\odot}$  BH (Liu et al. 2019) which has subsequently been challenged (El-Badry & Quataert 2020; Abdul-Masih et al. 2020; Shenar et al. 2020; Eldridge et al. 2020, see also Liu et al. 2020) and revised measurements of the mass of Cygnus X-1 (Miller-Jones et al. 2021). A sample of BHBHs detected with LISA could possibly help to constrain the BH mass distribution.

One particularly interesting class of potential LISA sources is BHNSs. With the recent detection of two BHNSs by the LIGO scientific collaboration, the existence of these DCOs has been confirmed (The LIGO Scientific Collaboration et al. 2021). However, with only two detections (not including the low-confidence candidate GW190426 (Abbott et al. 2020b) or GW190425 and GW190814 which have not been ruled out as BHNSs (Abbott et al. 2020a,c)) and no electromagnetic counterparts, the formation rate and properties of BHNSs are still uncertain. Current predictions for the merger rate of BHNSs range across three orders of magnitude (e.g. Abadie et al. 2010; Broekgaarden et al. 2021) so the number of detections in LISA will be important in reducing this uncertainty, thereby refining our understanding of the remnants and evolution of massive stars. Similar to NSNSs, these binaries are also expected to have electromagnetic counterparts. A distinctly exciting possibility is the detection of a pulsar–BH system or millisecond pulsar–BH system (Narayan et al. 1991). These systems could be observed not only by LISA, but also radio telescopes such as MeerKAT and SKA, which would help to improve the measurement of individual system parameters and to constrain uncertain binary evolution processes (e.g. Pfahl et al. 2005; Chattopadhyay et al. 2020).

For the purposes of this investigation, we consider the ‘classical’ isolated binary evolution channel (e.g. Tutukov & Yungelson 1973, 1993; Smarr & Blandford 1976; Srinivasan 1989; Kalogera et al. 2007; Belczynski et al. 2016) in which double compact objects are formed following common-envelope ejection or a phase of highly non-conservative mass transfer (van den Heuvel 2011; van den Heuvel et al. 2017). We do not, however, account for several alternative proposed formation channels, which could affect the rate and distribution of detectable NS and BH binaries in LISA. These channels include: dynamical formation in dense star clusters (e.g. Sigurdsson & Hernquist 1993; Portegies Zwart & McMillan 2000; Miller & Lauburg 2009; Rodriguez et al. 2015), young/open star clusters (e.g. Ziosi et al. 2014; Di Carlo et al. 2020; Rastello et al. 2020, 2021) and (active)

galactic nuclei discs (e.g. Morris 1993; Antonini & Rasio 2016; McKernan et al. 2020), isolated (hierarchical) triple evolution involving Kozai-Lidov oscillations (e.g. Stephan et al. 2016; Silsbee & Tremaine 2017; Antonini et al. 2017; Toonen et al. 2020), and chemically homogeneous evolution through efficient rotational mixing (e.g. de Mink et al. 2009; Mandel & de Mink 2016; Marchant et al. 2016, 2017; du Buisson et al. 2020).

In this paper, we present predictions of the detection rate and distribution of binary properties (masses, frequency, eccentricity, distance, merger time) of BHBH, BHNS and NSNS binaries formed through isolated binary evolution in the Milky Way. We explore the effect of varying physical assumptions in our population synthesis model on our results as well as discuss the effect of extending the LISA mission length and the prospects for distinguishing DCO detections from the WDWD background.

Earlier work on BHBHs, BHNSs and NSNSs in LISA has used a variety of population synthesis codes, Milky Way models and LISA specifications, resulting in a wide range of predictions (Nelemans et al. 2001; Liu 2009; Belczynski et al. 2010; Liu & Zhang 2014; Lamberts et al. 2019; Lau et al. 2020; Breivik et al. 2020; Sesana et al. 2020). We build upon previous efforts but with several important improvements. We explore the effect of varying binary physics assumptions by repeating our analysis for 20 different models and comparing the effect on the detection rate and distributions of source parameters. We use a model for the Milky Way that accounts for the chemical enrichment history and is calibrated on the latest APOGEE survey (Majewski et al. 2017; Frankel et al. 2018), whereas most others did not consider the effect of metallicity in detail (see however Lamberts et al. 2019; Sesana et al. 2020). We provide a full treatment of the eccentricity of detectable sources both for the inspiral evolution as well as gravitational wave signal during the LISA mission. Moreover, our binary population synthesis simulation is the most extensive of its kind to date and makes use of the adaptive sampling algorithm STROOPWAFEL (Broekgaarden et al. 2019, 2021). Overall we simulate over 2 billion massive binaries to produce the DCO populations used in this work. We find that this large number of simulations is important to reduce the sampling noise.

All data produced in this study is publicly available on Zenodo <sup>1</sup>[1](https://zenodo.org/record/4699713) as is the population used in our simulations <sup>2</sup>[2](https://zenodo.org/record/4574727). We make all code used to produce our re-

sults available in a Github repository <sup>3</sup>[3](https://github.com/TomWagg/detecting-DCOs-in-LISA). In addition, the repository contains step-by-step Jupyter notebooks that explain how to reproduce and change each figure in the paper. In a companion paper, Wagg et al. (in prep), we present **LEGWORK**<sup>4</sup>, the **LISA Evolution and Gravitational Wave Orbit Kit**, a python package designed for making predictions for the detection of sources with LISA, which we use in this work.

Our paper is structured as follows. In Section 2, we describe our methods for synthesising a population of binaries, the variations of physical assumptions that we consider, how we simulate the Milky Way distribution of DCOs and our methods for calculating a detection rate for LISA. We present the main results for our fiducial model in Section 3, before exploring the variations in the detectable population when changing our physical assumptions in Section 4. In Section 5 we discuss these results. In Section 6, we compare and contrast our methods and findings to previous work and finish with our conclusions in Section 7.

## 2. METHOD

To produce predictions for the DCOs that are detectable with LISA, we synthesise a population of DCOs using the population synthesis methods described in Section 2.1. In Section 2.2 we describe our model for the Milky Way and how we place DCOs in randomly sampled Milky Way instances. We evolve the orbit of each DCO in a Milky Way instance up to the LISA mission and calculate the detection rate for that instance using the methods presented in Section 2.3.

### 2.1. *Binary population synthesis*

We use the grid of 20 binary population synthesis simulations recently presented in Broekgaarden et al. (2021) and Broekgaarden et al. (in prep). This grid of simulations is synthesised using the rapid population synthesis code **COMPAS** (Stevenson et al. 2017; Vigna-Gómez et al. 2018; Stevenson et al. 2019; Broekgaarden et al. 2019). COMPAS follows the approach of the population synthesis code BSE (Hurley et al. 2000, 2002) and uses fitting formula and rapid algorithms to efficiently predict the final fate of binary systems. The code is open source and documented in the papers listed above, the online documentation<sup>5</sup> and in the upcoming methods paper (Team COMPAS: J. Riley et al. in prep). We summarise the main assumptions and settings relevant for this work in Appendix A.

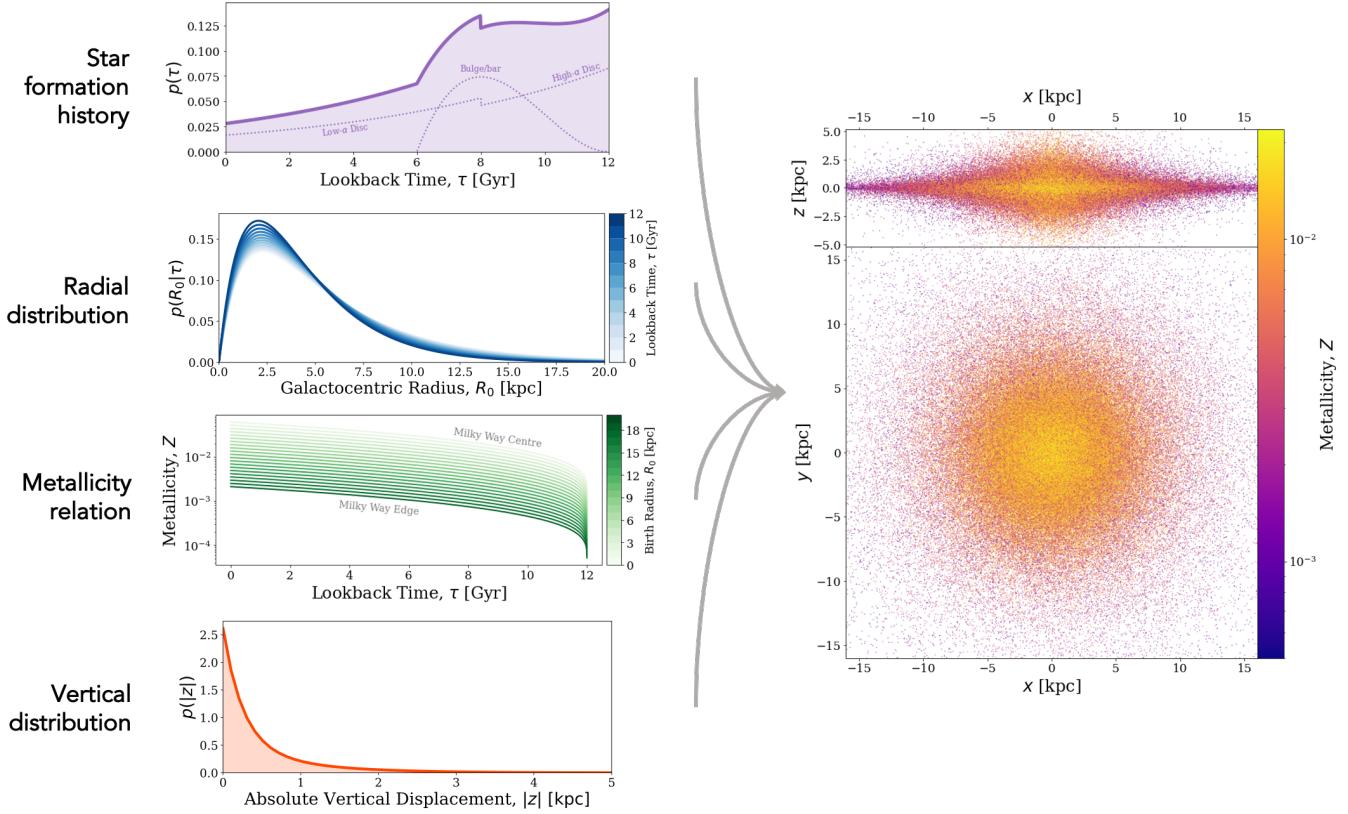
<sup>1</sup><https://zenodo.org/record/4699713>

<sup>2</sup><https://zenodo.org/record/4574727>

<sup>3</sup><https://github.com/TomWagg/detecting-DCOs-in-LISA>

<sup>4</sup><https://legwork.readthedocs.io>

<sup>5</sup><https://compas.science>



**Figure 1.** A schematic illustrating how we model the Milky Way. The left panel illustrates the different model aspects: star formation history of 3 galactic components (individually shown in the dotted lines), radial distribution, metallicity-radius-time relation, and height distribution. The right panel shows an example instance of the Milky Way with 250000 binaries shown as points coloured by metallicity. The top panel shows a side-on view and the bottom panel a face-on view.

The result of the simulations is a sample of binaries, which, for each metallicity  $Z$ , have  $N_{\text{binaries}}$  binaries with parameters

$$\mathbf{b}_{Z,i} = \{m_1, m_2, a_{\text{DCO}}, e_{\text{DCO}}, t_{\text{evolve}}, t_{\text{inspiral}}, w\}, \quad (1)$$

for  $i = 1, 2, \dots, N_{\text{binaries}}$ , where  $m_1$  and  $m_2$  are the primary and secondary masses,  $a_{\text{DCO}}$  and  $e_{\text{DCO}}$  are the semi-major axis and eccentricity at the moment of double compact object (DCO) formation,  $t_{\text{evolve}}$  is the time between the binary's zero-age main sequence and DCO formation,  $t_{\text{inspiral}}$  is the time between DCO formation (that is immediately after the second supernova in the system) and gravitational wave merger and  $w$  is the adaptive importance sampling weight assigned by STROOPWAFEL (Broekgaarden et al. 2019, Eq. 7). We sample from these sets of parameters when creating synthetic galaxies.

## 2.2. Galaxy synthesis

In order to estimate a detection rate of DCOs with statistical uncertainties, we create a series of random instances of the Milky Way, each populated with a sub-sample drawn (with replacement) from the synthesised binaries described in Section 2.1.

Most previous studies that predict a detection rate for LISA place binaries in the Milky Way independently of their age or evolution. We improve upon this as the first study to use an empirically-informed analytical model of the Milky Way that takes into account the galaxy's enrichment history by applying the metallicity-radius-time relation from Frankel et al. (2018). The authors developed this relation in order to measure the global efficiency of radial migration in the Milky Way and calibrated it using a sample of red clump stars measured with APOGEE (Majewski et al. 2017).

In Section 2.2.1, we outline our model for the Milky Way and in Section 2.2.2 we explain how we combine our population of synthesised DCOs with this Milky Way model.

### 2.2.1. Milky Way model

Fig. 1 shows the distributions and relations outlined in this section and also displays an example random galaxy drawn using this model.

Our model for the Milky Way accounts for the low-[ $\alpha/\text{Fe}$ <sup>6</sup>] disc, high-[ $\alpha/\text{Fe}$ ] disc and a central component approximating a bar/bulge. The low- and high-[ $\alpha/\text{Fe}$ ] discs are often also referred to as the thin and thick discs because the stellar vertical distribution is better fit by a double exponential rather than a single one. However, this doesn't allow one to assign a star to either the thin or thick disk purely based on its height above the Galactic plane. Therefore, we use the chemical definition of the two disks with the [ $\alpha/\text{Fe}$ ] nomenclature because there is a clear bimodal distribution in the chemical plane, allowing stars to be assigned to each of the disc components based on their chemical abundances. For each of the three components, we use a separate star formation history and spatial distribution, which we combine into a single model, weighting each component by its stellar mass. [Licquia & Newman \(2015\)](#) gives that the stellar mass of the bulge is  $0.9 \times 10^{10} \text{ M}_\odot$  and the stellar mass of the disc is  $5.2 \times 10^{10} \text{ M}_\odot$ , which we split equally between the low- and high-[ $\alpha/\text{Fe}$ ] discs (e.g., [Snaith et al. 2014](#)).

*Star formation history:* We use an exponentially declining star formation history ([Frankel et al. 2018](#)) (inspired by the average cosmic star formation history) for the combined low- and high-[ $\alpha/\text{Fe}$ ] discs,

$$p(\tau) \propto \exp\left(-\frac{(\tau_m - \tau)}{\tau_{\text{SFR}}}\right), \quad (2)$$

where  $\tau$  is the lookback time (the amount of time elapsed between the binary's zero-age main sequence and today),  $\tau_m = 12 \text{ Gyr}$  is the assumed age of the Milky Way and  $\tau_{\text{SFR}} = 6.8 \text{ Gyr}$  is the star formation timescale ([Frankel et al. 2018](#)). The two discs form stars in mutually exclusive time periods, such that the high-[ $\alpha/\text{Fe}$ ] disc forms stars in the early history of the galaxy (8–12 Gyr ago) and the low-[ $\alpha/\text{Fe}$ ] disc forms stars more recently (0–8 Gyr ago). We normalize both distributions so that an equal amount of mass is formed in each of the two components over their respective star forming periods.

The star formation history of the Milky Way bulge (which we assume here to be dominated by the central bar) has many uncertainties due to the (1) sizeable age measurement uncertainties at large ages in observational studies, (2) complex selection processes affecting the observed age distributions, and (3) formation mechanisms that are still under debate. However, the central bar was shown to contain stars with an extended age range, with most observed stars between 6 and 12 Gyr with

a younger tail of ages that could come from the subsequent secular growth of the Galactic bar (e.g., [Bovy et al. 2019](#)). To model the bar's age distribution more realistically than in previous studies (which assume an old bulge coming from a single starburst), we choose to adopt a more extended star formation history using a  $\beta(2, 3)$  distribution, shifted and scaled such that stars are only formed in the range [6, 12] Gyr. We show these distributions in the top left panel of Fig. 1.

*Radial distribution:* For each of the three components we employ the same single exponential distribution (but with different scale lengths)

$$p(R) = \exp\left(-\frac{R}{R_d}\right) \frac{R}{R_d^2}, \quad (3)$$

where  $R$  is the Galactocentric radius and  $R_d$  is the scale length of the component. For the low-[ $\alpha/\text{Fe}$ ] disc, we set  $R_d = R_{\text{exp}}(\tau)$ , where  $R_{\text{exp}}(\tau)$  is the scale length presented in [Frankel et al. \(2018, Eq. 5\)](#)

$$R_{\text{exp}}(\tau) = 4 \text{ kpc} \left(1 - \alpha_{R_{\text{exp}}} \left(\frac{\tau}{8 \text{ Gyr}}\right)\right), \quad (4)$$

where  $\alpha_{R_{\text{exp}}} = 0.3$  is the inside-out growth parameter<sup>7</sup>. This scale length accounts for the inside-out growth of the low-[ $\alpha/\text{Fe}$ ] disc and hence is age dependent. We assume  $R_d = (1/0.43) \text{ kpc}$  for the high-[ $\alpha/\text{Fe}$ ] disc ([Bovy et al. 2016, Table 1](#)) and  $R_d = 1.5 \text{ kpc}$  for the bar component ([Bovy et al. 2019](#)). We show the combination of these distributions in the second panel on the left in Fig. 1.

*Vertical distribution:* Similar to the radial distribution, we use the same single exponential distribution (but with different scale heights) for each component, given by

$$p(|z|) = \frac{1}{z_d} \exp\left(-\frac{z}{z_d}\right), \quad (5)$$

where  $z$  is the vertical displacement above the Galactic plane and  $z_d$  is the scale height. We set  $z_d = 0.3 \text{ kpc}$  for the low-[ $\alpha/\text{Fe}$ ] disc ([McMillan 2011](#)) and  $z_d = 0.95 \text{ kpc}$  for the high-[ $\alpha/\text{Fe}$ ] disc ([Bovy et al. 2016](#)). For the bar, we set  $z_d = 0.2 \text{ kpc}$  ([Wegg et al. 2015](#)). We show the combination of these distributions in the bottom left panel of Fig. 1.

*Metallicity-radius-time relation:* To account for the chemical enrichment of star forming gas as the Milky Way evolves, we adopt the relation given by ([Frankel](#)

<sup>6</sup>Nomenclature used to describe the enhancement of  $\alpha$  elements compared to iron in stellar atmospheres

<sup>7</sup>In  $R_{\text{exp}}(\tau)$ , we use 4 kpc instead of 3 kpc for the 0 Gyr exponential scale-length of the disc as NF finds that it provides a better fit to the original data

et al. 2018, Eq. 7)

$$\begin{aligned} [\text{Fe}/\text{H}](R, \tau) &= F_m + \nabla[\text{Fe}/\text{H}]R \\ &- \left( F_m + \nabla[\text{Fe}/\text{H}]R_{[\text{Fe}/\text{H}]=0}^{\text{now}} \right) f(\tau), \end{aligned} \quad (6)$$

where

$$f(\tau) = \left( 1 - \frac{\tau}{\tau_m} \right)^{\gamma_{[\text{Fe}/\text{H}]}} , \quad (7)$$

$F_m = -1$  dex is the metallicity of the gas at the center of the disc at  $\tau = \tau_m$ ,  $\nabla[\text{Fe}/\text{H}] = -0.075 \text{ kpc}^{-1}$  is the metallicity gradient,  $R_{[\text{Fe}/\text{H}]=0}^{\text{now}} = 8.7 \text{ kpc}$  is the radius at which the present day metallicity is solar and  $\gamma_{[\text{Fe}/\text{H}]} = 0.3$  set the time dependence of the chemical enrichment. We can convert this to the representation of metallicity that we use in this paper by applying (e.g. Bertelli et al. 1994)

$$\log_{10}(Z) = 0.977[\text{Fe}/\text{H}] + \log_{10}(Z_\odot). \quad (8)$$

Although Frankel et al. (2018) only fit this model for the low- $[\alpha/\text{Fe}]$  disc, we also use this metallicity-radius-time relation for the high- $\alpha$  disc and the bar, but focusing on the chemical tracks more representative to the inner disc and large ages. Sharma et al. (2020) showed that using a simple continuous model for both the low- and high- $[\alpha/\text{Fe}]$  discs, the Milky Way abundance distributions could be well reproduced. Empirically, the abundance tracks in the  $[\alpha/\text{Fe}]\text{-}[\text{Fe}/\text{H}]$  plane (and other elements) of the stars in the bulge/bar follow the same track as those of the old stars in the Solar neighbourhood (Griffith et al. 2021; Bovy et al. 2019, Fig. 7.), which motivates our modelling choice to use the same metallicity-radius-time relation.

### 2.2.2. Combining population and galaxy synthesis

For each Milky Way instance, we randomly sample the following set of parameters

$$\mathbf{g}_j = \{\tau, R, Z, z, \theta\} \quad (9)$$

for  $j = 1, 2, \dots, N_{\text{MW}}$ , where we set  $N_{\text{MW}} = 2 \times 10^5$ ,  $\tau, R, Z$  and  $z$  are defined and sampled using the distribution functions specified in Section 2.2.1,  $\theta$  is the polar angle sampled uniformly on  $[0, 2\pi]$  and  $Z$  is the metallicity. Figure 1 shows an example of a random Milky Way instance created with these distributions. This shows how these distributions translate to positions and illustrates the gradient in metallicity over radius.

We match each set of galaxy parameters  $\mathbf{g}_j$ , to a random set of binary parameters  $\mathbf{b}_{Z,i}$ , by drawing a binary from the closest metallicity bin to the metallicity in  $\mathbf{g}_j$ .

Each binary is likely to move from its birth orbit. Although all stars in the Galactic disc experience radial migration (Sellwood & Binney 2002; Frankel et al.

2018), DCOs generally experience stronger dynamical evolution as a result of the effects of both Blaauw kicks (Blaauw 1961) and natal kicks (e.g. Hobbs et al. 2005).

The magnitude of the systemic kicks are typically small compared to the initial circular velocity of a binary at each Galactocentric radius. Therefore, we expect that kicks will not significantly alter the overall distribution of their positions (see however e.g. Brandt & Podsiadlowski 1995; Abbott et al. 2017b). Given this, and for the sake of computational efficiency, we do not account for the displacement due to systemic kicks in our analysis.

### 2.3. Gravitational wave detection

We use the Python package **LEGWORK** to evolve binaries and calculate their LISA detectability. For a full derivation of the equations given below please see the **LEGWORK** release paper, Wagg et al. 2021b (in prep), or the documentation.

#### 2.3.1. Inspiral evolution

Each binary loses orbital energy to gravitational waves throughout its lifetime. This causes the binary to shrink and circularise over time. In order to assess the detectability of a binary, we need to know its eccentricity and frequency at the time of the LISA mission. For each binary in our simulated Milky Way, we know that the time from DCO formation to today is  $\tau - t_{\text{evolve}}$  and that the initial eccentricity and semi-major axis are  $e_{\text{DCO}}$  and  $a_{\text{DCO}}$ . We find the eccentricity of the binary at the start of the LISA mission,  $e_{\text{LISA}}$ , by numerically integrating its time derivative (Peters 1964, Eq. 5.13) given the initial conditions. This can be converted to the semi-major axis at the start of LISA,  $a_{\text{LISA}}$  (Peters 1964, Eq. 5.11), which in turn gives the orbital frequency,  $f_{\text{orb,LISA}}$ , by Kepler's third law since we know the component masses.

#### 2.3.2. Binary detectability

We define a binary as detectable if its gravitational wave signal has a signal-to-noise ratio (SNR) of greater than 7 (e.g. Breivik et al. 2020; Korol et al. 2020). The sky-, polarisation- and orientation-averaged signal-to-noise ratio,  $\rho$ , of an inspiraling binary can be calculated with the following (e.g. Finn & Thorne 2000)

$$\rho^2 = \sum_{n=1}^{\infty} \int_{f_{n,i}}^{f_{n,f}} \frac{h_{c,n}^2}{f_n^2 S_n(f_n)} df_n , \quad (10)$$

where  $n$  is a harmonic of the gravitational wave signal,  $f_n = n \cdot f_{\text{orb}}$  is the frequency of the  $n^{\text{th}}$  harmonic of the gravitational wave signal,  $f_{\text{orb}}$  is the orbital frequency,  $S_n(f_n)$  is the LISA sensitivity curve at frequency  $f_n$  (e.g. Robson et al. 2019) and  $h_{c,n}$  is the characteristic strain

of the  $n^{\text{th}}$  harmonic, given by (e.g. Barack & Cutler 2004)

$$h_{c,n}^2 = \frac{2^{5/3}}{3\pi^{4/3}} \frac{(G\mathcal{M}_c)^{5/3}}{c^3 D_L^2} \frac{1}{f_{\text{orb}}^{1/3}} \frac{g(n,e)}{nF(e)}, \quad (11)$$

where  $D_L$  is the luminosity distance to the source,  $f_{\text{orb}}$  is the orbital frequency,  $g(n,e)$  and  $F(e)$  are given in Peters & Mathews (1963) and  $\mathcal{M}_c$  is the chirp mass, defined as

$$\mathcal{M}_c = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}. \quad (12)$$

Note that the length of the LISA mission affects how much a binary will evolve over the mission and thus changes the frequency limits in Eq. 10. The SNR generally scales as  $\sqrt{T_{\text{obs}}}$  (with exceptions for sources very close to merging) and thus the SNR of a typical source in a 10-year LISA mission is approximately 1.58 ( $=\sqrt{10/4}$ ) times stronger than in a 4-year mission.

We use LEGWORK to calculate the signal-to-noise ratio for each binary and the package ensures that enough harmonics are computed for each binary such that the error on the gravitational wave luminosity remains below 1%.

### 2.3.3. Detection rate calculation

For each physics variation model and DCO type, we first convert the COMPAS simulation results into a total number of DCOs in the Milky Way,  $N_{\text{DCO}}$ . We do this by integrating the full mass and period distributions and stars and normalising to the total Milky Way mass. For more details see Appendix B.

We then determine the fraction of binaries that are detectable in each Milky Way instance by summing the adaptive importance sampling weights of the binaries that have an SNR greater than 7 and dividing by the total weights in the simulation. We multiply this fraction by the total number in the Milky Way to find a detection rate (which we write as a total number of detections per LISA mission)

$$N_{\text{detect}} = \frac{\sum_{i=0}^{N_{\text{MW}}} w_i \phi(i)}{\sum_{i=0}^{N_{\text{MW}}} w_i} N_{\text{DCO}}, \quad (13)$$

where  $\phi(i) = 1$  if a binary is detectable and 0 otherwise. We calculate the detection rate by Monte Carlo sampling 2500 Milky Way instances (each containing 200,000 DCOs) for each DCO type and every physics variation in order to obtain values for the uncertainty on the expected detection rate.

## 3. RESULTS I - PREDICTIONS FOR LISA SOURCES

In this section we present our predictions for the population of detectable LISA sources. In total we expect 124 detections in a 4-year LISA mission, of which 74, 42 and 8 are BHBHs, BHNSs and NSNSs respectively, based on our fiducial simulations. In the remainder of this section, we discuss where the sources are expected relative to LISA's sensitivity curve (Sec. 3.1), their properties (Sec. 3.2), their locations in the Milky Way (Sec. 3.3), their progenitors and formation channels (Sec. 3.4) and finally we discuss the expected SNR and how accurately we expect that the parameters can be measured (Sec. 3.5). Note that all results shown in this section are based on our fiducial simulations. A discussion of the impact of variations in the physics assumptions is provided in Sec. 4.

### 3.1. The LISA sensitivity curve and the population of detectable sources

We show the expected LISA sensitivity curve (Robson et al. 2019) in Fig. 2 and overplot our predictions for the distribution of detectable sources. Eccentric systems emit gravitational waves in multiple harmonic frequencies ( $nf_{\text{orb}}$ , with  $n = 2, 3, \dots$ ). We choose to plot them at the  $x$ -coordinate that corresponds to the frequency of the harmonic that dominates the SNR. For circular systems, the  $x$ -coordinate simply corresponds to  $2f_{\text{orb}}$ . The  $y$ -coordinate indicates the strength of the signal (or to be more precise, the amplitude spectral density, ASD), including the contribution from *all* harmonics.

For reference, we show dotted lines to indicate where a hypothetical binary system would reside assuming a given distance (diagonal lines) and a fixed remaining inspiral time (vertical lines). For each line we assume the binary is circular and has a chirp mass equal to the average of the sample ( $\langle \mathcal{M}_c \rangle$ , annotated in each panel), unless indicated otherwise. We also overplot the LISA verification binaries (star symbols, Kupfer et al. 2018).

We observe several features in Fig. 2 that are worth discussing and explaining. We note that some of these have also been described in earlier studies (see Sec. 6). Firstly, we note the empty band that separates the LISA sensitivity curve and the detectable population. This is the result of the criteria for detection where we require  $\text{SNR} > 7$ .

We further note that the detectable population is concentrated on the left side of LISA's sensitivity window. The peak is located at a frequency of about 0.4 mHz, which is much lower than the frequency where LISA will be most sensitive (about 10 mHz). This can be understood as an evolutionary effect. DCOs typically form with wide orbits (low frequencies) that would not be detectable yet. Their orbits shrink as they lose angular



**Figure 2.** The LISA sensitivity curve is shown together with the density distribution of the characteristic strain and the dominant frequency for all detectable sources in our simulations (top) and separated by type (bottom). Contours show the percentage of the population enclosed. The remaining 2% of the population is shown as dots with a size that scales with the sampling weight. To guide the interpretation, we show reference lines marking where a circular binary would reside for a given distance (diagonal line) and remaining inspiral time (vertical lines), assuming an average chirp mass  $\langle M_c \rangle$ , orientation and sky location. The curved reference line shows the signal expected for an eccentric binary at 30 kpc. The coloured lines in the bottom panels show a contour that encloses 90% of the population that is circular. LISA verification binaries are overplotted in the top panel (star symbols). See also Fig. 15 and Sec. 3.1 for a discussion.

momentum in the form of gravitational waves leading to an increase of their orbital frequency until they become detectable. Even though the signal emitted by a more compact binary is stronger, such systems are increasingly rare because they evolve faster and faster towards higher frequency as the inspiral accelerates (Peters 1964). The vertical grid lines show these rapidly decreasing inspiral times at increasing frequencies. Most of the population is thus expected to reside at low frequencies.

In the lower three panels, we show the contributions of the different types of sources. Comparing them, one can observe the shift in the frequency at which the peak is located, at 0.3 mHz, 0.7 mHz and 1 mHz for BHBH, BHNS and NSNS systems respectively. This is a result of the difference in chirp mass. A higher mass system can emit at lower frequency and still produce the minimum SNR needed for detection. We note that this effect can be used to distinguish the heavier DCOs that we discuss in this work from WDWD systems, at least probabilistically (e.g. Sesana et al. 2020). In the same way, this also explains the offset in frequency between the population of sources we predict and the LISA verification binaries (star symbols).

Inspecting the dotted reference lines, we note that the peak of the density distribution of observable sources coincides with the location expected for circular systems at 8 kpc, which is the distance to the centre of the Milky Way. As can be seen best in the lower panels, the reference lines for 0.1 and 30 kpc enclose the majority of systems as expected given the dimensions of our Milky Way.

There is a distinct subpopulation of binaries, most clearly visible in the lower panels as an offshoot that extend downwards to ASDs of  $10^{-19} \text{ Hz}^{-1/2}$ , especially around 2 mHz. This offshoot is almost uniquely composed of eccentric binaries, as can be seen in Fig. 15 in our appendix, which shows a similar figure but colouring individual systems by their eccentricity. This can also be seen, albeit indirectly, from the contours shown in the bottom panels of Fig. 2, which encompass 90% of the *circular* sources in each population. This contour does not include the offshoot. We conclude that eccentric sources occupy a very different region in this diagram.

### 3.2. Properties of the detectable systems

In Figure 3, we show the 1D distribution of several individual parameters of the population of detectable binaries together with the 1- and  $2\sigma$  uncertainties obtained via bootstrapping. The distributions shown here are approximated by kernel density estimators, cor-

rected for edge effects by mirroring the sample (Schuster 1985).

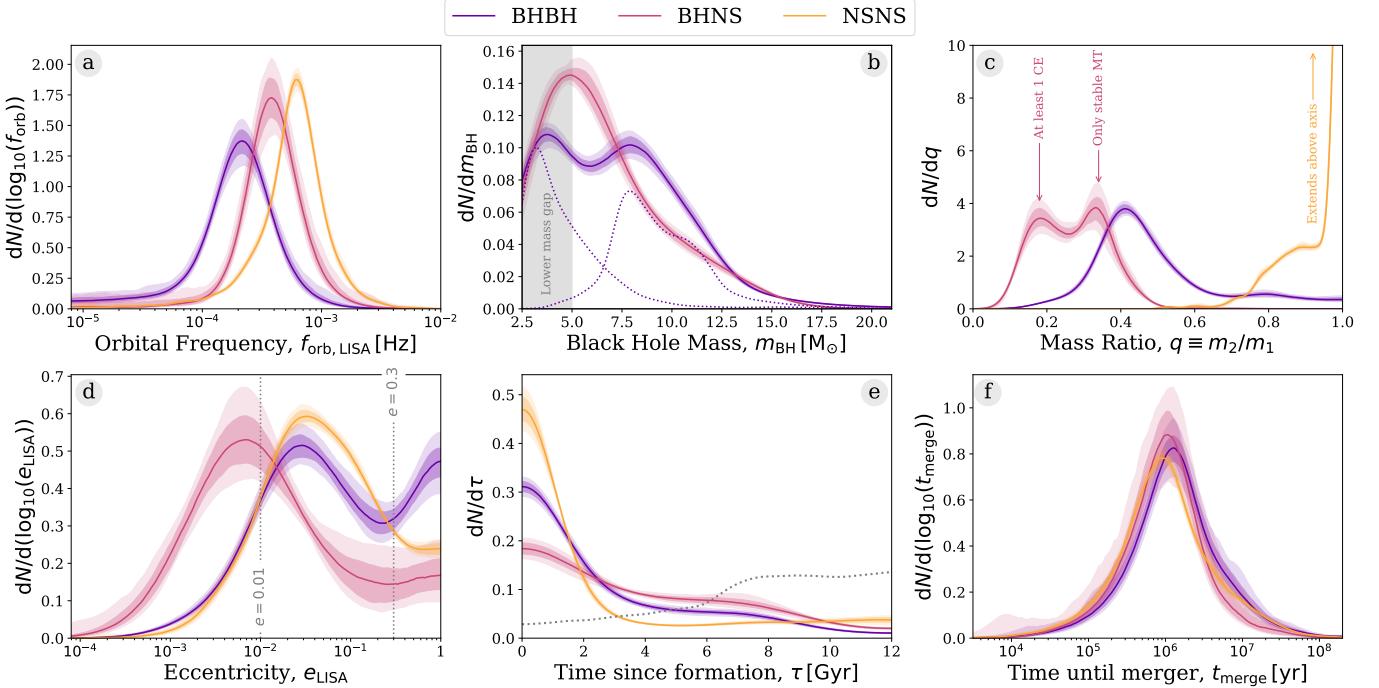
*Orbital Frequency*—The orbital frequency distributions for BHBHs, BHNSs and NSNSs (Fig. 3a) peak at progressively increasing frequencies. As already mentioned above, this is because a higher mass DCO at the same distance and eccentricity requires a lower frequency to produce the same signal-to-noise ratio and thus be detected. The distributions appear nearly symmetric, but closer inspections shows that the left hand side is more populated, which can be seen most clearly in the curve for the BHBHs. This is because of the contribution of highly eccentric binaries, which are most abundant in the BHBH population. These systems are still detectable by LISA, despite their low orbital frequency, as the high eccentricity means that the majority of the GW signal is emitted at higher harmonics, where LISA is more sensitive.

*Black Hole Mass*—In Fig. 3b, we show the distribution masses of individual black holes that are part of BHNS and BHBH systems. The distribution for BHBH shows a bimodality. This results from the two contributions of the more and less massive BHs in BHBH systems, which peak at around  $8 M_\odot$  and  $3.5 M_\odot$  respectively as shown by the dotted lines (see also our discussion of the mass ratios below).

For both the BHBHs and BHNSs, we see that the black hole mass distribution favours low masses. About 90% of BHs have masses below  $11 M_\odot$ , in our fiducial simulations shown here. This is in stark contrast with the masses of the BHs that have been detected so far by ground-based GW detectors, where heavy BHs, with masses of about  $30 M_\odot$  and higher have been common. Ground-based detectors are strongly biased towards high mass systems, since they can be detected out to larger distances. Instead, LISA is, in principle, sensitive to inspiraling BHBHs throughout the entire Milky Way regardless of their mass, as long as we catch them at the right time when their dominant frequencies coincide with the range where LISA is sensitive.

A further reason for the preference for low masses is that the population of BHs detectable by LISA primarily come from progenitors that formed from high metallicity gas according to our simulations (see Sec. 3.4 also Fig. 17). Stellar winds are stronger at high metallicity leading to increased mass loss. This affects the mass of the most massive black holes that can be formed (Berczynski et al. 2010).

We also note that our BH mass distribution extends down below  $5 M_\odot$  to  $2.5 M_\odot$  which is our fiducial maximum neutron star mass. The BHs in this simulation fill



**Figure 3.** Properties of detectable systems for a 4-year LISA mission in our fiducial model. Each panel shows a kernel density estimator for a single property, coloured by DCO type. The shaded areas show the 1- and 2- $\sigma$  uncertainties (obtained via bootstrapping). The dotted lines in the black hole mass panel show the individual primary and secondary mass distributions. See Sec. 3.2 for a discussion.

the so-called “lower mass gap” marked as a grey band (Özel et al. 2010; Farr et al. 2011), see also Shao & Li (2021) who recently pointed this out. This prediction is sensitive to adopted model for fall back during the SN explosion as we discuss this further in Sec. 4.2.1

**Mass Ratio**—The mass ratio distributions for each DCO type are very distinct from one another, as can be seen in Fig. 3c. The majority of NSNSs have a mass ratio close to unity, with 90% of systems having  $q > 0.8$ . The reason for the concentration around equal masses is that most NSs are formed either through electron-capture supernovae (ECSN) or from low mass stars in our simulations. We assume a remnant mass for any NS formed through ECSN of  $1.26 M_\odot$  (see Sec. A.2). The remnant mass prescription that we use assumes a fixed fallback mass for any star with a CO core mass less than  $2.5 M_\odot$ , such that many NSs end up with an identical mass of  $1.278 M_\odot$  (see Fryer et al. 2012, Eq. 19). This means that many NSs are formed with equal masses and hence we see a mass ratio distribution peaked around unity.

In contrast, only 8% of detectable BHBHs are formed with  $q > 0.8$  and the distribution peaks around  $q = 0.4$ . After closer inspection of the binary simulations, we find that the strong stellar winds in our (typically high-metallicity) progenitors are the reason behind this.

BHBHs with unequal masses typically come from progenitors that also had more extreme mass ratios at birth ( $90$  and  $30 M_\odot$  are typical for the progenitors of detectable BHBH systems in our simulations, see Sec. 3.4). The primary in such systems experiences strong mass loss by winds before filling its Roche lobe. This mostly happens during its early hydrogen-shell burning phase. The wind mainly reduces the mass of the envelope, but does not have a very strong effect on the core. By the time the primary fills its Roche lobe, it has become less massive and the mass ratio is closer to one. This favours stable mass transfer. The massive core of the primary star typically becomes the more massive BH. Accretion on the secondary star is limited and the secondary eventually provides the less massive black hole.

At the same time, stellar wind mass loss disfavours the formation of black holes with comparable masses. Such systems would have originated from progenitors that also started with comparable masses. The rather massive secondaries in these systems (especially after they accreted from the primary) experience very strong stellar wind mass loss. This limits the amount by which they can expand. At the same time, wind mass loss (and also stable and more conservative mas transfer in these systems) lead to widening of the orbit. Both effects, the reduced expansion and increased widening of the

orbit, tend to prevent the secondaries from being able to fill their Roche lobe. This thus limits the number of systems that experience the reverse mass transfer or common envelope phase needed shrink the orbits and make them tight enough to be detected as gravitational wave sources.

We find that detectable BHNSs have even more unequal mass ratios. Moreover, the mass ratio distribution is bimodal, where the two peaks arise from two distinct formation scenarios. Around two thirds of detectable BHNSs experience at least one common-envelope event, whilst the last third are formed through only stable mass transfer. The first peak at  $q = 0.18$  is from systems that experience at least one common-envelope and occurs at the expected mass ratio, which approximately follows the mean BH mass ( $\sim 6.5 M_{\odot}$ ) and NS mass ( $\sim 1.2 M_{\odot}$ ). Yet we also see a second peak at higher mass ratios around  $q = 0.34$ , which arises from the fraction of the population that underwent only stable mass transfer phases. The stability of mass transfer depends on the mass ratio, leading to a bias for more equal mass ratio systems, resulting in a peak at higher  $q$ .

**Eccentricity**—In Fig. 3d we show the eccentricity distributions. We find that most systems (73%) will have eccentricities larger than 0.01 during the LISA mission, which should in principle be detectable according to Nishizawa et al. (2016). This means that we will potentially be able to use eccentricity to distinguish these sources from WD sources (see Sec. 5.1.1). We note that several previous studies assumed all systems to be circular when calculating the detection rates (e.g. Liu & Zhang 2014; Lamberts et al. 2018; Sesana et al. 2020). See Sec. 6 for a discussion of the impact of this assumption

For systems with eccentricities higher than  $e \gtrsim 0.3$ , most gravitational wave energy is emitted in higher harmonics. Such systems are more rare, but we find them to be significant among the BHBH population, where they account for 21% of systems.

These BHBHs that are still eccentric when reaching the LISA band are primarily systems formed through the stable mass transfer channel (van den Heuvel et al. 2017). These systems are still relatively wide just prior to formation of the second BH. If the kick is oriented roughly in opposite direction to the orbital motion and has a velocity that is of similar magnitude as the orbital velocity, it will lead to the formation of a highly eccentric system.

It is rare to get such a “lucky kick”, but there are a few effects that favour this for BHBH. The kicks of BHs are reduced and they are thus less likely to disrupt the system. Moreover, because of their higher masses,

BHBHs can be observed already at lower orbital frequencies. This means that they have not had as much time to circularise and so still have significant eccentricity by the time of the LISA mission. Finally, LISA favours the detection of eccentric systems, if all other properties are held fixed. This is because the gravitational wave emission is stronger (Eq. C13) and the energy is emitted at higher frequencies (Peters & Mathews 1963, Eq. 20) where LISA is more sensitive.

The lower abundance of highly eccentric systems among the NSNS and BHNS systems may seem counter-intuitive since neutron stars receive stronger natal kicks, which one may expect to lead to more eccentric systems. However, in practice, these stronger kicks often disrupt these systems.

[TODO: @Selma suggests we return to this section. NSs do **NOT** receive stronger natal kicks in our simulations since most are formed through ECSN e.g. and have 30km/s instead of 265km/s.]

*Time since formation*—In Fig. 3e we show how long ago the DCOs detectable by LISA formed. For comparison we show the star formation history we have assumed in our Milky Way model (grey dotted line). Star formation was highest at early times 6 – 12 Gyr ago, after which it declined. In contrast, the DCOs detectable by LISA primarily formed in the relatively *recent* history of our Milky Way, about 2 Gyr ago. This reflects the fact that binaries in our simulation typically take about a Gyr to merge, a result as is commonly found also in other population synthesis studies.

When comparing the distribution of formation times for the three different DCO types we see that NSNSs are most strongly concentrated at recent times, followed by BHBH and then BHNS. To understand this it is helpful to consider how the inspiral time scales with various parameters

$$t_{\text{inspiral}} = \frac{a^4}{(m_1 + m_2)^3} \cdot \frac{q}{(1+q)^2} \cdot (1 - e^2)^{-7/2}. \quad (14)$$

The inspiral time depends most strongly on the separation at DCO formation ( $a$ ) and this is where the three types also differ most strongly, see Fig. 17. The NSNS systems have the tightest orbits at DCO formation. The median of the distribution of separations at DCO formation,  $\langle a_{\text{DCO}} \rangle_{\text{med}}$ , relate as 8:3:1 for BHBH:BHNS:NSNS in our simulations. This results in increase of the inspiral time by a factor of about 4000:80:1. The total masses affect the inspiral time to the third power and this where the heavier BHBH systems are favoured. The median total masses differ by ratios of 6:4:1 for BHBH:BHNS:NSNS in our simulations, impacting the inspiral times such that they are a factor of 200:60:1

shorter, partially counteracting the effect of the separations. The term depending on the mass ratio  $q$  only varies by about 30% for the mass ratio ranges considered here and is not of interest. The eccentricity term is not of importance for mildly eccentric systems,  $f(e_{\text{DCO}} \leq 0.3) \leq 1.4$  but of large importance for the very eccentric  $f(e_{\text{DCO}} \geq 0.9) \geq 300$ . The fraction of highly eccentric systems with  $e_{\text{DCO}} > 0.9$  is 33%, 16% and 8% of for BHBH, BHNS and NSNS respectively, see also Fig. 17. We conclude that the shorter median separations at DCO formation are the main reason why NSNS are most strongly peaked at short delay times, followed by BHBH where the inspiral timescales are short because of their high masses and substantial eccentricities.

In addition to the peak at recent times, we see that all three distributions have a large extended tail of systems that formed through out the history of the milky way, most prominently for the BHNS systems. For the BHBH and BHNS we find that the peak at very recent formation times comes from systems formed through stable mass transfer that obtained very high eccentricities. The long tail is a mixture of channels.

[TODO: Can we do a little better in understanding the tail?]

[TODO: I also think this section has become *very* dense...but can't see how to make it better]

*Time until merger*—The final panel of Fig. 3 shows the remaining time until merger for each of the DCO types at the start of LISA mission. The distributions are strikingly similar and peak with merger times of around a Myr.

The merger time is a function of the mass, frequency and eccentricity of the sources, such that more massive, higher frequency and more eccentric sources merge faster (Peters 1964, Eq. 5.14). So, despite the fact that each DCO type often has higher values in any one of these properties, the convolution of all three tends to negate the differences. For example, NSNSs have the highest orbital frequencies and are mildly eccentric whilst BHNSs have moderate orbital frequencies and are more circular. However, BHNSs are more massive in general and so the overall merger times are distributed very similarly for both DCO types.

### 3.3. Distribution in the Milky Way

For our Milky way model we considered three different components, a low- $\alpha$  (“thin”) disc, an older high- $\alpha$  (“thick”) disc and a bulge/bar. In Table 1 we summarise the number of detections originating from each of these components. Despite the fact that only 42.5% of systems are formed in the low- $\alpha$  disc, we find that

Component	Formation		Detectable		
	All	All	BHBH	BHNS	NSNS
Low- $\alpha$ disc	42.5%	86%	89%	82%	85%
High- $\alpha$ disc	42.5%	6%	5%	8%	12%
Bulge/bar	15.0%	8%	6%	10%	3%

**Table 1.** Percentage of systems in each Galactic component.

86% of the detections originate from this component. This is because most detectable systems were formed relatively recently (see Fig. 3) and so the high- $\alpha$  disc and bulge are effectively too old to contribute many detectable systems. Nevertheless, we do find a significant fraction of detectable systems originate in the high- $\alpha$  disc and bulge, indicating that it is still important to include these components, as was ignored in some earlier works (see Sec. 6).

In the top panel of Fig. 4, we show the density distribution for detectable DCOs in the galaxy. We see that most detectable sources are concentrated towards the Galactic centre, with a strong bias towards sources that are on our side of the Milky Way in the vicinity of the solar system (indicated with the  $\odot$  symbol). In principle, systems are detectable out to large distances of about 20 kpc and more, although they become increasingly rare, as can be seen from the 95% contour.

The differences between different DCO types can be seen more clearly in the bottom panel of Fig. 4 where we show the distribution of the distances,  $D$ , from earth to the detectable systems. Each distribution peaks around 8 kpc, which is the distance to the centre of the Milky Way. The distribution for BHBH and BHNS systems follow a very similar shape, favouring the detection of systems with distance  $< 8$  kpc, but with a tail extending out to about 20 kpc.

The distribution for NSNS stands out by being flatter, making them more common nearby and, surprisingly, also at larger distances. This may seem counter intuitive as one might naively expect the less massive NSNS systems would not be observable out to larger distances than the more massive BHNS and NSNS systems. To understand the differences we need to consider not only the mass distribution of binaries, but also their eccentricity and frequency distributions. Since, each parameter contributes to the calculation of the SNR (and thus affects the maximum distance at which systems can be detected).

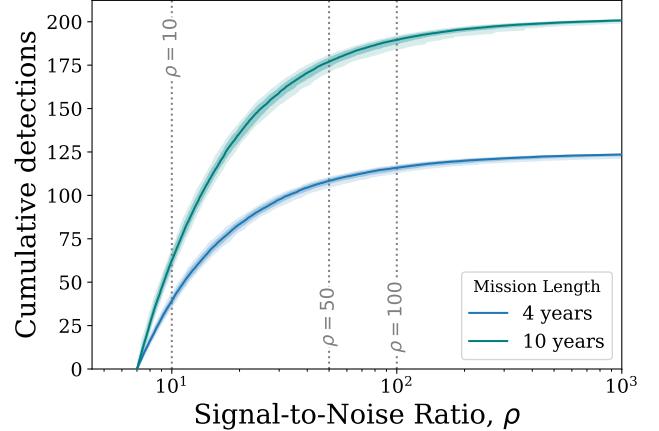
The reason that the NSNSs are favoured at higher distances is that the NSNS population has the highest fraction of “mildly” eccentric systems ( $0.01 < e < 0.03$ ). In contrast, the BHNS population has a much higher fraction of effectively circular systems ( $e < 0.01$ ), which emit



**Figure 4.** **Top:** A face-on view of the Galactic density distribution for detectable DCOs. We show the density distribution for the top 95% of the sources, the rest are indicated by scatter points whose sizes correspond to their sampling weights. The coloured lines show the 75% contour for each of the individual DCO types. The large cross passes through  $(0, 0)$  and helps to highlight the bias towards the position of the sun, which is indicated by the  $\odot$ . **Bottom:** As Figure 3, but for the luminosity distance.

weaker gravitational wave compared to equivalent eccentric systems. Therefore, despite their typically higher masses, the distance at which a BHNS source is detectable is generally lower than for the mildly eccentric NSNS.

Conversely, the BHBH population has the highest fraction of highly eccentric systems ( $e > 0.3$ ). Although one may naively expect that this would result in stronger signals (and so further distances), for a system to have these high eccentricities in LISA, it must still be early in its evolution (otherwise it would have circularised) and thus have a low orbital frequency. The result of this



**Figure 5.** Cumulative number of LISA detections with a given signal-to-noise ratio. Colours indicate LISA mission length and shading shows 1- and 2- $\sigma$  uncertainties (obtained via bootstrapping).

is that highest eccentricity systems tend to have lower SNRs and so cannot be detected at large distances.

Overall we see that the eccentricity distribution of NSNSs occupies a “sweet spot” where the gravitational wave power is increased compared to circular systems, but it isn’t too high that the frequency is significantly impacted. This means that NSNSs can be seen out to the largest distances of the three DCO types.

### 3.4. Progenitors and formation channels

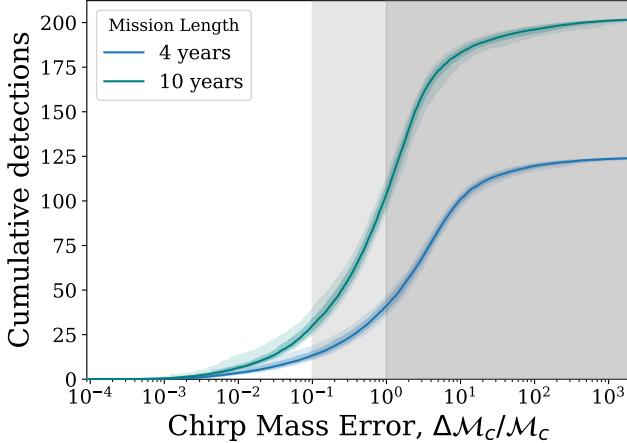
Formation channel figure is Fig. 16.

Progenitor distribution figure is Fig. 17.

### 3.5. How accurately will we be able to infer the parameters of detected systems?

So far we have discussed the properties of the detectable sources that can, in principle, be inferred from observations. However, many sources will only be barely detectable. Here we discuss for how many systems we expect to get high enough SNR to obtain accurate and useful measurements of these parameters. Below we discuss the expected SNR distribution and the typical uncertainties expected for the most relevant parameters, namely, the chirp mass and sky localisation.

*Signal-to-noise ratio*—In Fig. 5 we show the cumulative number of detections with a given SNR. Although many have SNRs around the our assumed detection threshold of 7, many systems are detection with very high SNRs. We find that for a 4(10)-year LISA mission, of the 124(202) detections, 85(138), 16(27) and 9(14) systems have SNRs greater than 10, 50 and 100 respectively. These high SNR systems are typically, but not only, the more massive BHBH systems.



**Figure 6.** As Fig. 5, but for the chirp mass uncertainty. The shaded areas indicate regions with more than 10% and 100% uncertainty. .

*Chirp mass*—The chirp mass is important for identifying the type of the source of a detected GW signal. The uncertainty of the chirp mass depends on the uncertainty in the measured orbital frequency, the time derivative of the orbital frequency and the eccentricity as detailed in Appendix C.

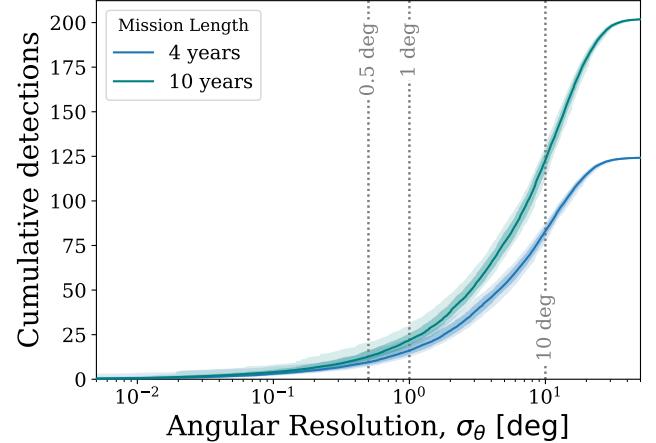
We find that for a 4(10)-year LISA mission, approximately 41(105) detections have measurable chirp masses ( $\Delta M_c/M_c < 1$ , indicated by the dark shaded region in Fig. 6) whilst 13(31) have chirp mass uncertainties below 10%. This uncertainty is generally dominated by the uncertainty on the time derivative of the frequency, since most of the binaries are too early in their inspiral for LISA to measure a strong chirp.

Note that increasing the mission length significantly increases the number of detections for which we can measure the mass. The total number of detections only scales as  $\sqrt{T_{\text{obs}}}$ , yet we find that the total number of detections *with well measured masses* scales approximately as  $T_{\text{obs}}$ .

*Eccentricities*—[TODO: SdM: Proposed additional paragraph.... can we say something compact? You made a large number of plots and estimates..]

*Sky localisation*—An accurate sky localisation will be essential to possibly identify electromagnetic counterparts or distinguish sources that come from different components of our Milky Way.

We quantify the sky localisation of a source by estimating the angular resolution for the detectable sources. Since all potential sources are effectively stationary on the timescale of the LISA mission, which can follow Mandel et al. (2018) and use the timing accuracy of LISA and the effective detector baseline to calculate the



**Figure 7.** As Fig. 5, but for the angular resolution. .

angular resolution,  $\sigma_\theta$ , as

$$\sigma_\theta = 16.6^\circ \left( \frac{7}{\rho} \right) \left( \frac{5 \times 10^{-4} \text{ Hz}}{f_{\text{dom}}} \right) \left( \frac{2 \text{ AU}}{L} \right), \quad (15)$$

where  $L$  is the effective detector baseline, which for LISA is 2 AU since it will orbit the Sun.

We plot the distribution of expected angular resolutions in Fig. 7. We see that, for a 4(10)-year LISA mission, approximately 82(123) sources can be resolved to an angular resolution better than 10 degrees and only 16(23) better than 1 degree.

For comparison, the size of a pencil beam for a 15 m diameter SKA dish observing at 1.4 GHz is roughly 0.67 square degrees (Smits et al. 2009), corresponding to an angular resolution of  $\sigma_\theta = \sqrt{(0.67/\pi)} = 0.46^\circ$  (similar to the angular size of the moon). We will further discuss the prospects of matching LISA detections to radio pulsars with SKA further in Sec. 5.2.

## 4. RESULTS II - IMPACT OF PHYSICS ASSUMPTIONS

In this section we explore the effect of varying the underlying binary physics assumptions, both on the detection rate (Section 4.1) and the properties of the detectable systems (Section 4.2).

### 4.1. Detection rates

In Figure 8, we show the expected number of LISA detections for each model variation and discuss the prominent trends in the following sections. For the exact rates and uncertainties plotted in this figure see Table 3 and for a plot of the rates relative to the fiducial model see Fig. 18.

The BHBH detection rate is markedly robust across variations of physics assumptions, as can be seen in the top panel of Figure 8, with the mean detections in three

quarters of our models staying within 25% of the fiducial rate. In contrast, the BHNS detection rate is very sensitive to all changes that we made in our binary physics assumptions, varying across two orders of magnitude in the middle panel of Figure 8. Finally, in the last panel of Figure 8 we show that the NSNS detection is only moderately sensitive to certain changes in physics assumptions, varying by up to an order of magnitude, whilst showing no change for many models.

#### 4.1.1. Mass transfer variation trends

In models B-D, we set the mass transfer efficiency  $\beta$  to a fixed value. For the BHBHs and BHNSs, as we increase  $\beta$  the detection rates steadily decrease. This may seem unintuitive since one may assume that a higher mass transfer efficiency would lead to more massive compact objects and thus a more detectable population. However, this contributes to the envelope mass and so without increasing the core mass or fallback rate, the final compact object mass is relatively unaffected. Moreover, one must consider that most of these DCOs are formed through a common-envelope event and so retaining more of the envelope during mass transfer means that the eventual ejection of the envelope is much more difficult, thus leading to more stellar mergers and fewer detectable systems (e.g. Kruckow et al. 2018; Broekgaarden et al. 2021).

In contrast, for the NSNSs, the detection rate increases with increasing  $\beta$ . This is for two main reasons: firstly the ejection of a common-envelope is less problematic for NSNSs as they are less massive (e.g. Kruckow et al. 2018). Moreover, the increased mass transfer efficiency means that systems that were previously below the mass necessary to become a NS can now accrete enough mass to form a NS. Although the same is true for more massive stars becoming BHs instead of NSs, due to the IMF, there is a net flux of more stars becoming NSs.

Enforcing that case BB mass transfer is always unstable (model E) has little effect on the BHBH detection rate whilst moderately and drastically decreasing the detection rates of BHNSs and NSNSs respectively. This is because a large fraction of NSs (and the majority of NSNS binaries) are formed through case BB mass transfer. Therefore, setting this mass transfer to be always unstable results in many of these binaries to merge before they could become DCOs since we assume the pessimistic common-envelope (CE) scenario by default.

If we instead use the optimistic CE scenario when setting case BB mass transfer to be unstable (model F) we find that the rates instead uniformly increase across

DCO types. This is both from the natural increases from using the optimistic CE scenario (see discussion of model K in Sec. 4.1.2), as well as additional increases from adding more DCOs formed through CE events after case BB mass transfer. This explains why we see more detections for BHNSs and NSNSs in model F than in model K.

#### 4.1.2. Common-envelope variation trends

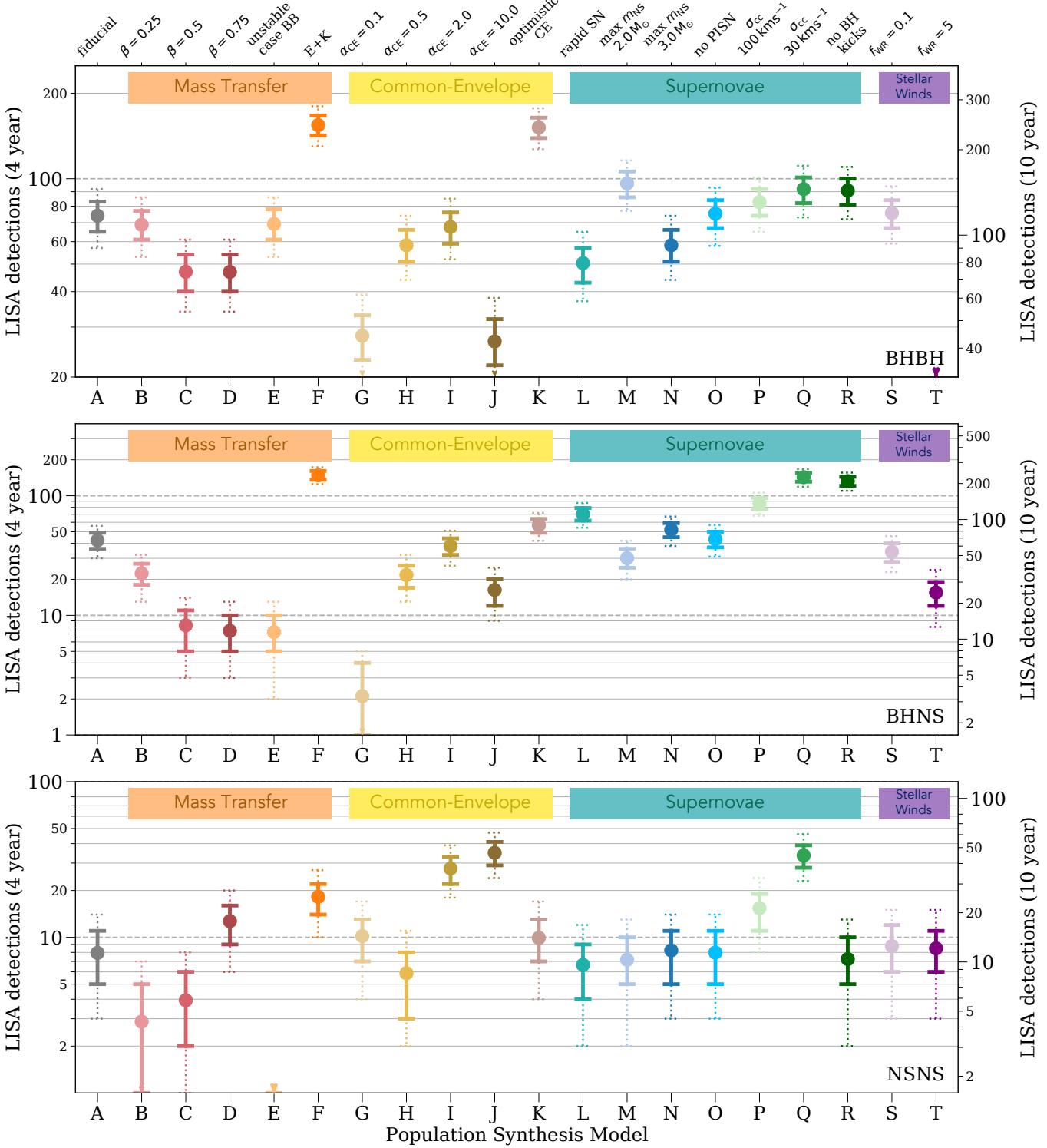
Models G-J alter  $\alpha_{\text{CE}}$ , the common-envelope efficiency parameter, to 0.1, 0.5, 2.0 and 10.0 respectively (where the fiducial model has  $\alpha_{\text{CE}} = 1.0$ ). Although we see that there are large variations in the LISA detection rates when changing  $\alpha_{\text{CE}}$ , the detectable *fraction* of systems remains largely unchanged. Therefore, any changes in the detection rate are due to changes in the *formation* rate of the DCOs.

Altering  $\alpha_{\text{CE}}$  has two major effects on the formation rate of double compact objects. Recall that  $\alpha_{\text{CE}}$  represents the fraction of the orbital energy that is available for the ejection of the common-envelope. Therefore, increasing  $\alpha_{\text{CE}}$  allows common-envelope ejection to occur at larger separations (since more energy is available at any given separation). A second effect of increasing  $\alpha_{\text{CE}}$  is that common-envelopes are less likely to result in stellar mergers as more energy is available for envelope ejection. We can understand the trends in formation rate of the three DCO types by considering these two effects.

The BHBH rate peaks for  $\alpha_{\text{CE}} = 1$  (model A) and reduces whether we increase *or* decrease  $\alpha_{\text{CE}}$ . This is because the increase in  $\alpha_{\text{CE}}$  shifts a significant fraction of the population to such large separations that they no longer merge within a Hubble time. Although the number of stellar mergers decreases as well, the overall change in the merging BHBH formation rate is negative. The inverse is also true for decreasing  $\alpha_{\text{CE}}$ , where BHBHs have shorter separations but stellar mergers are more frequent. Thus we find that  $\alpha_{\text{CE}} = 1$  seems to be optimal for BHBHs by balancing these two effects most effectively.

The BHNS rate follows the same pattern as BHBHs. This is expected since the BH is formed first for every BHNS in our sample and so the common-envelope event is due to a RLOF of BH progenitor, just as with BHBHs. However, given the similarity in the BHNS rates between  $\alpha_{\text{CE}} = 1$  and  $\alpha_{\text{CE}} = 2$  (models A and I), the ‘optimal’  $\alpha_{\text{CE}}$  for BHNSs seems to be somewhere between 1 and 2 rather than very close to 1 as with BHBHs.

In contrast,  $\alpha_{\text{CE}} = 1$  is far from optimal for NSNSs. Increasing  $\alpha_{\text{CE}}$  (models I and J) results in significantly higher rates. This is because LISA detectable NSNSs tend to have been formed very recently before detection



**Figure 8.** The number of expected detections in the LISA mission for different DCO types and model variations. Error bars show the 1- (solid) and 2- $\sigma$  (dotted) Poisson uncertainties. An arrow indicates that the error bar extends to zero. The left axis and grid lines show the number of detections in a 4-year LISA mission and the right axis shows an approximation of the number of detections in a 10-year mission (we scale the axis by  $\sqrt{T_{\text{obs}}}$ , see Table 3 for exact rates). Each model is described in further detail in Table 2 and details of the fiducial assumptions are in Section A.2. See also Fig. 18 and Sec. 4.1 for a discussion.

(see Fig. 3e). Therefore, though higher  $\alpha_{CE}$  results in larger *initial* separations, this does not adversely affect the NSNS rate as the systems are simply detected later after their formation compared to the fiducial model. So for models I and J, fewer systems result in stellar mergers and the overall population remains within the LISA band, producing much higher detection rates.

We explore the optimistic CE scenario in model K, meaning that we allow Hertzsprung gap donors to survive CE events. A large fraction of the progenitors of the BHs in our sample expand significantly during the Hertzsprung gap phase and initiate CE events. This means that, when we assume the optimistic CE scenario, the formation rate of BHs is greatly increased. Therefore, similar to models G-J, we find that though the detectable *fraction* does not change significantly, the increased formation rate of BHs in the Milky Way leads to this model predicting significantly more BHBH detections.

The progenitors of BHNSs tend to be less massive than those of BHBHs. Less massive stars expand less during the Hertzsprung gap and thus initiate CE events less frequently in this phase. For these reasons, as can be seen most clearly in Fig. 18, the relative increase in model K compared to the fiducial model is lower than BHBHs for BHNSs. For the same reason, the NSNS detection rate shows little significant change.

#### 4.1.3. Supernova variation trends

Changes to the remnant mass prescription (models L–N) have no effect on the NSNS detection rate for two reasons. A large fraction of the NSs in our sample are formed through ECSN, in which case their masses are set to  $1.26 M_{\odot}$  regardless of the remnant mass prescription. Secondly, the majority of the remaining NSs have such low progenitor masses that they are set to the lowest remnant mass in the prescription ( $1.28 M_{\odot}$  for the delayed prescription and  $1.1 M_{\odot}$  for the rapid prescription). This means that changing the upper portion of the mass range of NSs in the remnant mass prescription has little effect.

This is not the case for BHBHs and BHNSs. The different remnant mass prescriptions and maximum neutron star masses change whether a progenitor becomes a black hole or a neutron star. Hence, using the rapid prescription (model L) or increasing the maximum neutron star mass (model N), means that more progenitors form high mass NSs rather than low mass BHs and so more BHNSs are formed. This explains the higher detection rates for BHNSs and lower detection rates for BHBHs in these models. The inverse is true for model M, where instead more BHBHs are formed and detected.

Additionally, we find that not implementing pair-instability supernovae (PISN) pulsational pair-instability supernovae (PPISN) in model O has no effect on the results for any DCO type. This is because the average metallicity of the Milky Way is high enough than no progenitor retains enough mass to initiate a PISN or PPISN.

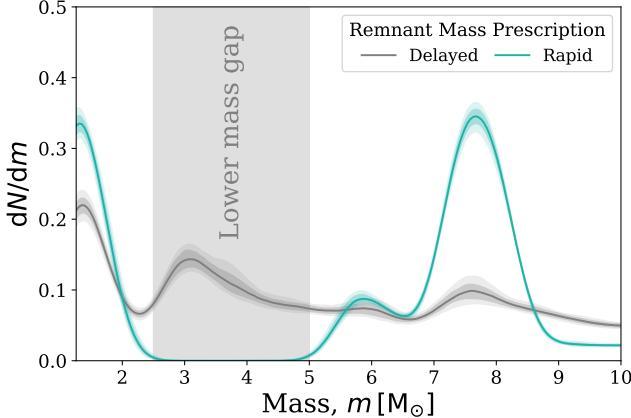
Decreasing the core-collapse supernova velocity dispersion (models P–Q) increases the detection rates for each DCO type since lower kicks result in fewer disrupted binaries and hence a more numerous detectable population. This increase is least prominent for BHBHs since their increased mass mean that disruptions are less frequent than, for instance, in NSNSs. It is also notable that these two models best match the detection rates from LIGO (Broekgaarden et al. in prep). Therefore we find the best matching models for the LIGO detection rates also produce the highest LISA detection rates.

In model R we do not give black holes natal kicks. For BHBHs, the rate is approximately equal to that of model Q since kicks of  $30 \text{ km s}^{-1}$  are equivalent to no kicks for massive BHs in our sample. For BHNSs, this model’s rate is still well above the fiducial but decreases slightly relative to model Q as the number of surviving binaries is limited by the neutron star kick more than the black hole kick. NSNSs are unaffected as one would assume.

#### 4.1.4. Stellar wind variation trends

We decrease the efficiency of Wolf-Rayet winds to 10% in model S. We find that decreasing the efficiency of these winds actually has very little effect on our predicted detection rates. This due to a combination of cancelling effects.

Firstly, the decreased winds mean that the DCOs (particularly those containing BHs) are generally more massive and so more detectable in LISA. Secondly, one may expect that LISA sources in model S would be higher frequency than in our fiducial model as decreased winds generally result in tighter binaries. However, though this *is* the case at DCO formation, we find by the time the sources have evolved until the LISA mission they are actually at *lower* frequencies than in our fiducial model. This is because the reduced winds allow DCOs to be formed at higher metallicity and, therefore, at more recent times. This means that the DCOs do not evolve for as long before the LISA mission and so remain at lower frequencies (wider separations) thus making them less detectable. Finally, we find that NSNSs are more eccentric and BHBHs are less eccentric than our fiducial model (with BHNSs relatively unchanged). The increase in eccentricity for NSNSs comes from the



**Figure 9.** Comparison of the component mass distribution of LISA detectable DCOs when using the Fryer delayed (model A) and rapid (model L) remnant mass prescriptions. Distributions are plotted in the same way as Fig. 3, except all DCOs types are shown in one curve and each type is weighted by its detection rate in the respective model.

same reason as the lower frequency, more recent birth times mean that binaries have less time to circularise. The same is not true for BHBHs as the more massive systems are less affected by supernova kicks and so fewer high eccentricity systems are formed.

Overall, despite the large differences in the system properties, these three effects in combination leave the detection rates relatively unchanged.

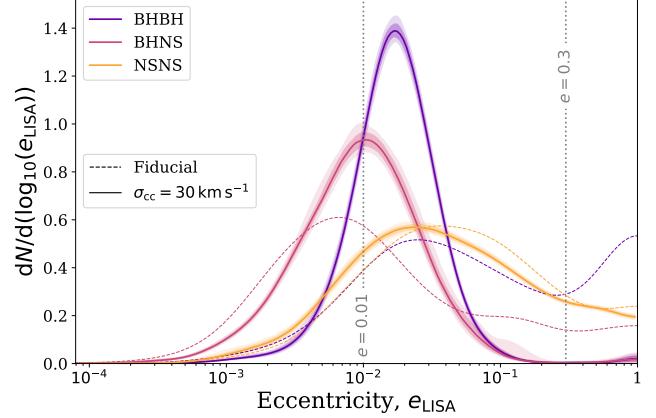
In model T we instead *increase* the efficiency of Wolf-Rayet winds by a factor of 5. In this model the detection rate of BHBHs decreases by over a factor of 10 and BHNSs by over a factor of 2 whilst the NSNS rate is relatively unchanged. Increasing the efficiency of WR winds strongly decreases the final masses of DCOs. This means that many progenitors that would have formed LISA sources under our fiducial assumptions would not have enough mass to produce a DCO, or produce a NS instead of a BH (therefore this effect is stronger on BHBHs than NSNSs). Moreover, DCOs that are formed tend to be less massive and therefore less detectable.

#### 4.2. Properties of detectable systems

In this section, we consider how varying underlying physics assumptions changes the properties of detectable systems. We focus on several key differences across physics variations rather than showing the differences in every model and thus this section is by no means exhaustive.

##### 4.2.1. Using LISA to investigate the lower mass gap

In Figure 9, we show the component mass distribution for all LISA detectable DCOs (BHBHs, BHNSs and NSNSs) for two different remnant mass prescriptions.



**Figure 10.** As Fig. 3d, but for model Q. For comparison, we show the mean distribution for the fiducial model (model A) as dashed lines.

The grey distribution uses the Fryer *delayed* remnant mass prescription (Fryer et al. 2012), which is our fiducial assumption. This prescription produces compact objects in the lower mass gap ( $2.5 M_{\odot} \leq m \leq 5 M_{\odot}$ ) and indeed we find that, of the LISA detectable DCOs, approximately 69% of BHBHs, 39% of BHNSs and 0% of NSNSs have at least one component in the lower mass gap. Overall, weighting by the relative detection rates, this gives that 55% of our predicted LISA DCO detections would have at least one component in the lower mass gap when using this remnant mass prescription.

Alternatively, the blue curve in Fig. 9 shows the same distribution but for the *rapid* remnant mass prescription (Fryer et al. 2012), which we use in model L. In this case, no compact objects are formed (and therefore, detected) in the lower mass gap. From the stark difference between these models, it is clear that it is difficult at this point to say with any certainty what fraction of systems LISA will detect in the lower mass gap given the highly uncertain formation rate of systems in this mass range.

However, it is important to highlight that if DCOs are formed with components in the lower mass gap, LISA will be able to detect them. And thus, LISA could be a useful instrument for providing evidence for the existence or non-existence of a lower mass gap based on the mass distribution of detected DCOs.

##### 4.2.2. Effect of natal kicks on eccentricity distribution

In Figure 10, we investigate how decreasing the magnitude of natal kicks from core-collapse supernovae affects the eccentricity distribution of LISA detectable DCOs. For reference, we show the mean fiducial distributions (model A) as dashed lines (see lower left panel of Fig. 3 for full comparison). In the main curves, we reduce

the velocity dispersion for core-collapse supernovae from  $265 \text{ km s}^{-1}$  to  $30 \text{ km s}^{-1}$  (model Q).

We find that the LISA detectable BHBHs are significantly less eccentric with weaker kicks, such that the population above  $e = 0.2$  is nearly completely eliminated. This is because BHBHs are often massive enough to withstand strong natal kicks without disrupting and these kicks tend to impart significant eccentricity. In model Q, very few systems are ever given such strong kicks and thus very few BHBHs are detected with significant eccentricity.

Since BHNSs are less massive than BHBHs and have more unequal mass ratios, they are more vulnerable to disruption during supernova kicks. BHNSs can only withstand strong kicks when they are aimed in the correct direction and so only a small ‘lucky’ fraction of the fiducial population is highly eccentric. Therefore in model Q, although we see that the population of highly eccentric BHNS systems is eliminated (similar to BHBHs), the peak of the distribution actually shifts to *higher* eccentricity. This is because a larger fraction of systems are given weaker kicks that BHNSs can withstand and these impart much more moderate eccentricities.

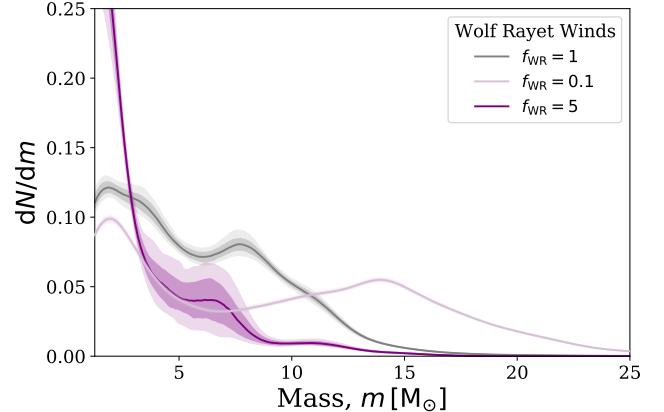
Finally, we find that the NSNS distribution is relatively unchanged between model A and Q. This is not surprising however since the majority of NSNSs are formed through electron-capture supernovae and ultra-stripped supernovae and for these types of supernovae we use  $\sigma_{\text{rms}}^{1D} = 30 \text{ km s}^{-1}$  already (see App. A.2) and thus there is little difference between the models.

Overall, we find that decreasing supernova natal kicks, though it strongly increases the number of detections (see Fig. 8), strongly *decreases* the fraction of highly eccentric systems that are detected.

#### 4.2.3. Effect of mass loss efficiency on mass distribution

In Fig. 11 we show the effect that changing the efficiency of Wolf-Rayet winds has on the individual component mass distribution. Decreasing the Wolf-Rayet wind efficiency allows the formation of more massive DCOs in the Milky Way and, indeed, we see that the distribution extends to  $25 M_\odot$  and fewer detectable systems are formed at low masses. By contrast, increasing the Wolf-Rayet efficiency by a factor of 5 strongly disfavours the formation of systems at high masses and approximately 85% of detectable systems have masses below  $5 M_\odot$ . These three distributions are very distinct and so it is possible that the mass distribution of LISA could help to constrain the efficiency of Wolf-Rayet winds.

## 5. DISCUSSION



**Figure 11.** As Fig. 9, but instead varying Wolf-Rayet wind efficiency (models S and T).

In this section we discuss the prospects of (and methods for) identifying LISA sources (Sec. 5.1), the possibility of matching LISA signals to SKA detections (Sec. 5.2) and the main caveats for this study (Sec. 5.3).

### 5.1. Identification of GW sources

It is important to note that, though we present predictions for the detection rates of specific DCO types, the nature of the source may not be immediately apparent from the gravitational wave signal. The population of Galactic WDWDs detectable with LISA will be several orders of magnitude larger than the population of more massive DCOs on which we focus in this paper (e.g. Korol et al. 2017). It is therefore imperative that we consider how to distinguish NS and BH binaries from this much more numerous population of sources. In addition to distinguishing them from WDWDs, we must consider how to discriminate between BHBHs, BHNSs and NSNSs themselves.

#### 5.1.1. Distinguishing from WDWD population

The simplest way to check whether a source is a WDWD is to check its chirp mass. The mass of a non-rotating white dwarf cannot be larger than the Chandrasekhar limit of  $1.4 M_\odot$  (Chandrasekhar 1931; Hamada & Salpeter 1961), so we can take the maximum chirp mass of a WDWD to be  $\sim 1.2 M_\odot$ . Therefore, any DCO with a chirp mass that satisfies  $M_c > 1.2 M_\odot + \Delta M_c$  must not be a WDWD (where  $\Delta M_c$  is calculated using Eq. C14). We find that for the detectable population of a 4(10)-year LISA mission, 24(38)% of BHBHs, 28(41)% of BHNSs and 4(5)% of NSNSs satisfy this condition. This method is not particularly effective for NSNSs since their average chirp mass,  $1.17 M_\odot$ , is below the Chandrasekhar limit.

Another discriminator between WDWDs and other DCOs is eccentricity. WDWDs formed in the disc are thought to be formed through isolated binary formation and have little to no eccentricity (e.g. Nelemans et al. 2001). Therefore, if any system is detected with anything other than one detectable harmonic, this suggests that the system is unlikely to be a WDWD. We find that for a 4(10)-year LISA mission, 55(61)% of BHBHs, 27(29)% of BHNSs and 66(68)% of NSNSs are detected with multiple harmonics. Both the absolute percentage and the relative improvement with an extended LISA mission is lower for the BHNSs with respect to other DCOs as we find that these BHNSs are less eccentric on average (see Fig. 3 and discussion in Sec. 3.2).

However, we should also consider that eccentric WDWDs could be formed through dynamical formation in Milky Way globular clusters (e.g. Willems et al. 2007; Kremer et al. 2018). This means that we cannot assume that eccentric binaries are not WDWDs unless they are detected in the Galactic plane. We can use the sky localisation, scale height of the disc and distance to the source to estimate what fraction of eccentric sources can be localised to the Galactic plane. This condition can be written as  $\sigma_\theta < \arcsin(z_{\text{plane}}/D_L)$  or  $D_L < z_{\text{plane}}$ , where we set the height of the Galactic plane,  $z_{\text{plane}} = 0.95$  kpc, to the scale height of the high- $\alpha$  disc. We apply this condition to find that the fraction of sources that are eccentric *and* localised within the disc for a 4(10)-year LISA mission are 40(40)% for BHBHs, 23(23)% for BHNSs and 59(59)% for NSNSs. Note that although the fractions are the same for the 10-year mission, the *absolute* number of detections is still greater.

Overall, combining these methods we find that for a 4(10)-year mission, LISA will detect at least 37(70) BHBHs, 18(38) BHNSs and 5(8) NSNSs that are distinguishable from the WDWD population. We highlight that, though the overall number of LISA detections in an extended mission only increases by a factor of  $\sqrt{T_{\text{obs}}}$ , the number of *distinguishable* detections increases by even more since each of the more numerous sources are better measured.

### 5.1.2. Discriminating between BHBHs, BHNSs and NSNSs

The problem of discriminating between the BHBH, BHNS and NSNS populations can be more difficult than distinguishing them from WDWDs. For NSNSs, we can follow a similar method to the WDWDs (see Sec. 5.1.1) by applying our knowledge of the maximum mass of a neutron star. Following our fiducial assumption, we can take the maximum mass of a neutron star as  $2.5 M_\odot$  and thus the maximum chirp mass that a system can attain without one of the components being a black hole

is  $\mathcal{M}_c = 2.2 M_\odot$ . For a 4(10)-year LISA mission, the fraction of systems that are above or below this limit (and thus *must* contain or not contain a BH component respectively) by more than  $\Delta\mathcal{M}_c$  is 21(33)% for BHBHs, 18(24)% for BHNSs and 47(62)% of NSNSs, which in terms of absolute detections is 16(39) for BHBHs, 7(17) for BHNSs and 4(8) for NSNSs.

For separating the BHBH and BHNS population one could do so probabilistically given the properties that are measured, particularly the orbital frequency, mass ratio and eccentricity, since these distributions are fairly different for the two DCO types (see Fig. 3). This method would pose a challenge however as it would likely only indicate which type was more likely rather than discriminate between them with strong evidence.

Another possible solution would be the existence of electromagnetic counterparts to the gravitational wave signal. In Section 5.2 we consider the possibility of detecting a pulsar within a BHNS or NSNS system. This could be used to identify the type of the source.

## 5.2. Matching LISA detections to pulsars with SKA

Since the vast majority of the LISA detectable population of DCOs will not merge for many years, the main form of electromagnetic counterpart for this population is pulsars. Therefore, for this section we focus only on BHNSs and NSNSs since no BHBH system will contain a pulsar. The joint detection of a binary pulsar with LISA and SKA would not only help to constrain the parameters of the binary, but also enable investigation of other compact object physics. A pulsar(PSR)+BH can provide stringent tests of theories of gravity, in particular the “No-hair theorem” (Keane et al. 2015). Alternatively, an ultrarelativistic PSR+NS system could be used to measure the neutron star equation of state up to an order of magnitude more accurately than other proposed observational constraints (Kyutoku et al. 2019; Thrane et al. 2020).

We estimate on average, given the number of detectable pulsars and SKA sky area, each pulsar in SKA occupies a region with an angular resolution of  $\sigma_\theta < 1.3^\circ$  or  $0.7^\circ$  for SKA-1 and SKA-2 respectively (see Appendix E). Therefore, any DCOs containing NSs localised by LISA with an angular resolution lower than these values can be unambiguously matched to the radio signal in SKA. By considering Fig. 7, approximately 11 and 6 (for SKA-1 and SKA-2) DCOs will satisfy this constraint.

If there is more than one pulsar in the region given by the LISA sky localisation, one can compare the measured parameters of the system in LISA and SKA. Both SKA and LISA will measure the orbital frequency to

high precision, as well as the time derivative of the frequency and chirp mass to a lesser precision, of each of these systems. Therefore, one could perform a targeted search with SKA that checks the sky location given by LISA and only looking for binary pulsars with orbital frequencies within the uncertainties. If there was *still* more than one possible pulsar one could also check against the chirp mass. In this way, we expect it will be possible to get a joint detection between SKA and LISA even when the sky area implied by the LISA detection contains more than one pulsar.

In order to assess the efficacy of this method, we would need to know the probability that two random binary pulsars would have orbital frequencies and chirp masses close enough that one could not tell which pulsar matches the LISA detection. This would require simulating the SKA population of pulsars with a code such as PSRPOPPy (Bates et al. 2014) to find the frequency and chirp mass distribution, which is beyond the scope of this paper. However, the uncertainty on the orbital frequency of a binary on the detection threshold ( $\text{SNR} = 7$ ) for a 4-year LISA mission is  $2.5 \times 10^{-9}$  Hz and  $1.0 \times 10^{-9}$  Hz for a 10-year mission (calculated using Eq. C15). Therefore, we expect that SKA could likely isolate the correct binary pulsar to match to a LISA detection even when several are present in the sky localisation region.

### 5.3. Caveats

*Binary fraction and other normalisation choices:* The choice of the binary fraction has a strong effect on our results. For the normalisation of the detection rate we conservatively choose to set the binary fraction of all stars to 50%. The binary fraction is uncertain and not well constrained, though we know it is higher for more massive stars (e.g. Sana et al. 2012). Increasing it to 70% and 100% would increase our expected detection rates for every physics variation by 30% and 67% respectively. Moreover, we assume that the mass ratio of secondary stars is uniformly distributed and this could also have strong effect on our results. We highlight this uncertainty in the normalisation to point out that the exact magnitude of the rates in each case could change drastically as our understanding shifts. However, changes to this normalisation would affect every variation equally and so the trends across variations are more robust.

*Population synthesis limitations:* As with any study involving a population synthesis code, our results rely on uncertain stellar and binary physics and the use of approximate fitting formulae. COMPAS uses fitting formulae and approximate prescriptions based on (sometimes limited) grids of detailed models to describe

the evolution of binary stars. Much of the underlying physics is uncertain, such as the common-envelope evolution and mass transfer physics. We attempt to understand the importance of these assumptions by varying over many different physics assumptions.

*Underlying helium star models:* One major weakness is that the Hurley et al. (2000) fitting formulae for the evolution of helium stars are based on a grid of models from  $0.3 M_{\odot}$  to  $10 M_{\odot}$ , for a single metallicity ( $Z = 0.02$ ) and thus the formulae have no metallicity dependence and are extrapolated for higher masses. A more comprehensive set of models in this regime could lead to large changes in the evolution of naked helium stars, a common progenitor of DCOs, and thus affect the detection rate of DCOs.

*Large progenitor masses:* Due to the high metallicity population within the Milky Way, many of the black holes that are formed in our sample come from progenitors with large initial masses. 39% of BHBs contain a BH with a progenitor mass in excess of  $100 M_{\odot}$ , as do 13% of BHNSs. The fitting formulae upon with COMPAS is based need to be extrapolated for stars of this mass and so their true evolution may not follow the results of population synthesis.

*Limited metallicity range:* Another limitation of the stellar evolution fitting formulae that COMPAS uses is that they are limited to a metallicity range of  $10^{-4} \leq Z \leq 0.03$  (due to the limits of the underlying models in Hurley et al. (2000)) and should not be extrapolated outside this region. However, in the Milky Way (based on the metallicity relation in Frankel et al. 2018), the metallicity distribution can extend as far as  $10^{-5} \leq Z \leq 0.06$ , with a significant fraction of star formation occurring past  $Z = 0.03$ . Therefore, for our study we had to re-assign any metallicities outside of COMPAS' range. We expect that stellar winds will be reduced to such a degree that they are effectively zero for any metallicity below the minimum. Hence we set any metallicity below the minimum,  $Z = 10^{-4}$ , equal to the  $Z = 10^{-4}$ . Similarly, DCO formation is less efficient at high metallicity (e.g. Broekgaarden et al. 2021) and so exploring metallicities above the COMPAS maximum is unlikely to contribute significantly to the observed rate. Therefore, we place any sampled metallicity above the maximum of  $Z = 0.03$  uniformly randomly in one of the top 5 highest bins that range across  $0.01416 < Z < 0.03$  (since using a single metallicity for many binaries leads to unphysical artifacts).

*Other formation channels:* We also note that our findings are only the result of a single formation channel (isolated binary formation). We do not consider other channels such as dynamical formation or chemically ho-

mogeneous evolution, which could increase the detection rate and alter the parameter distributions. For instance, Kremer et al. (2018) predict that around 21 systems could be detected in Milky Way globular clusters with LISA, formed through dynamical formation and thus different channels can still contribute significantly to the detection rate.

*Halo, globular clusters and young stellar clusters:* Our model for the Milky Way, though more extensive than many previous studies, does not consider the contributions from the Galactic halo, globular clusters or young massive and open stellar clusters. Lamberts et al. (2018) found that the halo’s contribution to the detection rate was minimal and, since the metallicity distribution of the halo is uncertain, we did not include it in our galaxy model. The impact of globular clusters would have required a more detailed look into dynamical formation that was beyond the scope of this paper but we again highlight the work of Kremer et al. (2018) that investigated these rates. Similarly, we did not consider DCO formation in young massive and open stellar clusters but Banerjee (2020) finds that the LISA detection rates could be even higher than contributions from isolated binary formation.

*Systemic kicks:* We do not include the effect of systemic kicks on the final location of the sources. This would require integrating the orbital evolution of the millions of binaries in our sample and thus was not computationally reasonable to include. We investigated the effect of kicks for a small grid of binaries and found that, though they would result in a more spread out distribution (with a smaller concentration in the Galactic centre) and larger heights above the plane, the overall distribution of positions would be relatively unchanged and very few sources have strong enough kicks to reach escape velocity for the Milky Way.

*Eccentricity measurement uncertainty:* The method that we use to determine the eccentricity uncertainty is pessimistic as it requires each harmonic to be individually detectable (e.g. Lau et al. 2020). In reality this may not be necessary depending on the efficacy of matched-filter analysis of LISA data. For an eccentric source to have been detected within the LISA data, several harmonics would already have to have been matched as originating from the same source. This could be done by looking in the same region of the sky for signals with similar chirp masses and distances to the most detectable harmonic in order to find other harmonics that are below the regular detection threshold. This would allow one to refine the measurement of the eccentricity uncertainty significantly by comparing the many different harmonics. Therefore, the eccentricity uncertainty

that we calculate in this study is a pessimistic estimate. Smaller eccentricity uncertainties would have two main effects on our results. Firstly, the chirp mass error would decrease slightly in the cases where it is dominated by the eccentricity uncertainty. However, it is mainly dominated by the frequency derivative uncertainty since most sources are essentially stationary in frequency. Secondly, it would improve our ability to distinguish between WD-WDs and higher mass DCOs. Until we know more about how LISA will search for eccentric sources, we rely upon our pessimistic estimates.

## 6. COMPARISON WITH PREVIOUS STUDIES

In Figure 12, we compare our results to similar previous studies that investigate the population of stellar-mass BHBHs, BHNSs and NSNSs that are detectable with LISA. Figure 12 details the expected detection rates predicted by each paper as well as their assumptions regarding their Milky Way galaxy model, binary population synthesis simulation and LISA mission specifications. We only include papers that are similar to our work, such that they use population synthesis and simulate sources in the Galactic plane. Moreover, Figure 12 does not include the numerous papers on the LISA WDWD population as we do not make predictions for these DCOs.

Nelemans et al. (2001)—were the first to investigate the population of LISA detectable stellar-mass double compact objects. We find a significantly higher detection rate for BHBHs and BHNSs, as well as a slightly lower rate for NSNSs. We can understand this difference from changes both to the specifications of LISA (such as the mission length and SNR threshold for detection) and our understanding of massive star evolution since the publication of their paper, which both strongly affect the expected detections rates.

Belczynski et al. (2010)—built upon the work of Nelemans et al. (2001), by using a different population synthesis code with two model variations and a multi-component model for the Milky Way. They find a much lower detection rate for BHNSs and NSNSs (and agreed on zero BHBHs) when compared to Nelemans et al. (2001). They state that this discrepancy from Nelemans et al. (2001) comes from differences in their population synthesis and an overall lower formation rate rather than any changes to LISA detectability. The low total detection rate for all DCOs in this paper compared to our work is unsurprising given the relatively high SNR threshold of 10 and short mission length of 1 year. The reduced mission length means that the source signal has much less time to accumulate, whilst also fewer WD-

		Predicted DCO Detection Rates						Population Synthesis												
Author	Year	BHBH			BHNS		NSNS		Code	Open Source Code	Metallicity	Binary Physics Variations								
<i>Wagg</i>	2021	74			117		42		66		8		13		COMPAS	✓	50 bins between [1e-4, 3e-2]			20
<i>Shao</i>	2021	12			2		X		BSE		✓		0.02		Fryer rapid, Mandel & Mueller remnant mass					
<i>Breivik</i>	2020	93			33		8		COSMIC		✓		0.02, 0.003		None					
<i>Lau</i>	2020	X			X		35		COMPAS		✓		0.0142		Case BB always unstable, Single SN, alpha=0.1					
<i>Sesana</i>	2020	4.2			6.5		X		X		BSE		✓		13 bins between [1e-4, 3e-2]				None	
<i>Lamberts</i>	2018	25			X		X		BSE		✓		13 bins between [1e-4, 3e-2]		None					
<i>Liu</i>	2014	6			3		16		BSE		✓		0.02		None					
<i>Belczynski</i>	2010	2.3	0	0.2	0	4	1.7	Startrack	X	0.02 (disc, bulge), 0.001 (halo)			Optimistic CE, Pessimistic CE							
<i>Nelemans</i>	2001	0			3		39		SeBa		X		0.02		None					

		Galaxy and Positioning							Detection				
Author	Year	Star formation history			Spatial distribution			Galactic Components	Metallicity Dependent Distributions	SNR Limit	LISA Mission Time (yr)	Eccentricity Treatment	
<i>Wagg</i>	2021	Exponential 8-0 Gyr ago (thin disc), Exponential 12-8 Gyr ago (thick disc), Skewed gaussian 0-6 Gyr (bulge)				Exponential radial and vertical, different scale length/height for each component, thin disc has inside-out growth			Thin disc, thick disc, bulge	✓	7	4, 10	Full
<i>Shao</i>	2021	Constant over 10 Gyr				Uniform flat disc of 15 kpc			Single disc	X	5	4	Full
<i>Breivik</i>	2020	Constant over 10 Gyr (thin disc), 1 Gyr burst 10 Gyr ago (bulge), 1 Gyr burst 11 Gyr (thick disc)				McMillan 2011			Thin disc, thick disc, bulge	X	7	4	Full
<i>Lau</i>	2020	Constant			Miyamoto & Nagai potential (disc), Wilkinson & Evans potential (halo)			Single disc or halo	X	8	4	Full	
<i>Sesana</i>	2020	FIRE simulation				FIRE simulation			Everything within 300kpc	✓	7	4, 10	Ignored during detection
<i>Lamberts</i>	2018	FIRE simulation				FIRE simulation			Everything within 300kpc	✓	5	4	Ignored during detection
<i>Liu</i>	2014	Constant over 13.7 Gyr			Exponential radial, sech^2 vertical (Benacquista+2007)			Single disc	X	7	2	Assumed circular	
<i>Belczynski</i>	2010	Constant over 10 Gyr (disc), 1 Gyr burst 10 Gyr ago (bulge), burst at 13 Gyr (halo)			Exponential sphere (bulge), exponential radial and vertical (disc), spherical shell (halo)			Disc, bulge, halo	X	10	1	Full	
<i>Nelemans</i>	2001	Exponential over 10 Gyr				Exponential radial, sech^2 vertical			Single disc	X	1, 5	1	Full

**Figure 12.** A table comparing previous studies of a similar nature to this work. The works listed in the table are Nelemans et al. (2001), Belczynski et al. (2010), Liu & Zhang (2014), Lamberts et al. (2018), Sesana et al. (2020), Lau et al. (2020), Breivik et al. (2020) and Shao & Li (2021).  .

WDs can be resolved in this time, leading to a weaker signal and an increased Galactic confusion noise relative to our work.

*Liu & Zhang (2014)*—performed a similar investigation using a different population synthesis code and find higher rates than earlier works. Their lower detection threshold and longer mission length compared to *Belczynski et al. (2010)* likely explains the relatively increased rates. Yet their rates are still significantly below what we find. This could be for several reasons; they assume all binaries are circular both in their evolution and for detection. This means that systems may not have inspiralled as far before the LISA mission or may appear to have weaker gravitational waves when eccentricity is not accounted for. They also use a simplified model for the Milky Way with a single disc of one metallicity and constant star formation, whilst also using a mission length half what we assume. Each of these factors likely contributes to the lower overall detection rates.

*Lamberts et al. (2018)*—presented a new approach to the problem by using the FIRE simulation (*Hopkins et al. 2014*) to distribute their sources rather than an analytical model of the Milky Way and were the first paper in this area to incorporate metallicity dependence into their Milky Way model. *Sesana et al. (2020)* followed up on this paper using the same simulated BHBH population and presented updated results for the number of expected BHBH detections. They find significantly fewer BHBHs than our fiducial model despite using the same SNR threshold and LISA mission length. The discrepancy between the results of *Sesana et al. (2020)* and those presented in this work could be caused by different treatments of eccentricity. Unlike our work, *Sesana et al. (2020)* assume that all binaries are circular for the purpose of detection in LISA, which could result in a lower number of detections by missing eccentric binaries that appear as weaker signals when assumed to be circular. This is especially relevant as we find that around 87% of LISA detectable BHBHs are not circular and around 21% have significant eccentricity (see Section 3.2). We also improve upon this work by using a larger number of metallicity bins compared to *Sesana et al. (2020)*, since a low number of metallicity bins can produce artificial features in the mass distribution of DCOs and possibly affect the detection rate (see Appendix D). Finally, it could be that different implicit assumptions in their population synthesis code lead to differences in our results (*Toonen et al. 2014*).

*Lau et al. (2020)*—focussed on the number of Galactic NSNS binaries that could be detected by LISA. Their study uses the same population synthesis code, COM-

PAS, as this work, though an earlier version. Despite this, their study finds a much larger number of detections. They make several different physical assumptions in their population synthesis, using the *Fryer et al. (2012)* rapid remnant mass prescription, limiting the maximum neutron star mass to  $2 M_{\odot}$  and not implementing PISM. However, we note that none of these assumptions strongly affect the NSNS LISA detection rate (see bottom panel of Fig. 8, models L, M and O) and so this is unlikely to entirely account for our differences. It is also important to highlight that COMPAS has received several improvements and bug fixes since *Vigna-Gómez et al. (2018)* (which contains the simulations used by *Lau et al. (2020)*) and these could possibly have affected the formation rate of NSNSs. Yet it is most likely that the remaining difference between our results is due to way in which we simulate the Milky Way. *Lau et al. (2020)* use a model for the Milky Way similar to that of *Breivik et al. (2020)*, which we use to estimate the impact of the choice of MW model is in Appendix D. We find that this simpler model for the Milky Way could result in an overestimation of the NSNS detection rate by at least a factor of two and so this may explain the discrepancy between our results. *Lau et al. (2020)* additionally only uses a single metallicity rather than the two used in *Breivik et al. (2020)* and so the overestimation could be even stronger as their results will have no contribution from lower metallicity systems.

*Breivik et al. (2020)*—introduced the population synthesis code COSMIC and presented detections for many different DCO types in LISA using this code. They find that LISA will detect 93 BHBHs, 33 BHNSs and 8 NSNSs in the Milky Way over a 4 year mission. *Breivik et al. (2020)* make many different physical assumptions from us, the most notable being that they assume the optimistic CE scenario and that case BB mass transfer is always unstable, whilst also using a simpler model for the Milky Way (see Appendix D). Thus for better comparison we ran our simulation using model F and the Milky Way model from *Breivik et al. (2020)*. This results in 97, 101 and 43 detections for BHBHs, BHNSs and NSNSs respectively. Therefore, though we are in very good agreement for BHBHs, we predict much higher rates for BHNSs and NSNSs. These differences are likely due to using a different population synthesis code (COSMIC), which has different underlying physics assumptions from COMPAS. Given our strong agreement for BHBHs, it is possible that COSMIC and COMPAS handle NSs differently and so lead to different detection rates for DCOs containing NSs. However checking this would require a more in-depth study of the in-

trinsic formation rate of DCOs containing NSs in the two codes.

*Shao & Li (2021)*—most recently investigated the detectability binaries containing BHs in LISA using BSE and a relatively simple model for the Milky Way (assuming a uniform flat disc, constant star formation and a single metallicity). They assume that kicks for NSs formed through ECSN are slightly higher than our work ( $50 \text{ km s}^{-1}$  instead of  $30 \text{ km s}^{-1}$ ). This may account for their particularly low BHNS rate (as the binaries would be more likely to disrupt), which is a factor of 20 lower than ours, but we expect their assumption of the optimistic CE scenario, reduced Wolf-Rayet winds and lower SNR detection threshold would offset this. As we show in Appendix D, their use of a simpler Milky Way model, especially with only a single metallicity, would lead to an underestimate of the BHBH and BHNS rates, which may explain the discrepancy in our results.

Overall, since the work of *Nelemans et al. (2001)*, in additions to the LISA mission specifications, the methods that we use to simulate binaries and the Milky Way have all changed significantly. We now predict that LISA detections of these massive DCOs are dominated by BHBHs, rather than NSNSs, whilst the absolute detection rates for BHBHs and BHNSs are much higher. Further studies in this area could improve on this work by including the effects of systemic kicks on the position of systems in the Milky Way and accounting for contributions from other formation channels.

## 7. CONCLUSION & SUMMARY

We provide predictions for the detection rate and population properties of LISA detectable BHBH, BHNS and NSNS. To this end, we use the rapid population synthesis code COMPAS to simulate over two billion massive binaries, to explore the effect of varying underlying physics assumptions that represent the most common uncertainties in binary physics. We use an new empirically-informed analytical model to distribute the resulting BHBH, BHNS and NSNS populations in a Milky-Way like galaxy based on their birth metallicity, in order to estimate and investigate the LISA detectable population of BH and NS binaries.

Our main conclusions can be summarised as:

- **Detection rate:** We find that on average, for our fiducial model, a 4(10)-year LISA mission will detect 74(118) BHBHs, 42(71) BHNSs and 8(13) NSNSs.
- **Black hole mass distribution:** We find that the black hole mass distribution for both BHBHs

and BHNSs that are detectable with LISA are not skewed towards heavier masses (as in current ground-based detectors), with 90% having BH masses less than  $11 M_{\odot}$ . Moreover, we find that for LISA detectable DCOs, 35% of BHBHs and 39% of BHNSs have masses in the theoretical lower mass gap between NSs and BHs. This implies that LISA will be able provide conclusive evidence about the existence of a lower mass gap.

- **Eccentricity distribution:** We find that BH and NS binaries still have significant eccentricities when observable by the LISA mission, unlike the more numerous WDWD LISA population. We predict that for the observed population 87%(21%) of BHBHs, 46%(8%) of BHNSs and 86%(12%)% of NSNSs have eccentricity of  $e > 0.01(0.3)$  at the start of the LISA mission.
- **Physics variations:** For BHBHs, we find that the detection rate is largely unaffected by changes in underlying physics assumptions, except for the optimistic CE scenario and increased Wolf-Rayet winds. Conversely, the BHNS detection rate varies widely with physics variation, ranging across three orders of magnitude. The NSNS detection rate shows some variation but only shows large changes under the assumption that case BB mass transfer is unstable and that core-collapse supernovae produce weaker kicks.
- **Distinguishable sources:** For a 4(10)-year LISA mission we estimate that, of the detectable population, at least 37(70) BHBHs, 18(38) BHNSs and 5(8) NSNSs will produce signals that are distinguishable from a signal produced by a WDWD. Additionally, we predict that we will be able to determine whether 16(39) BHBHs, 7(17) BHNSs and 4(8) NSNSs are signals from a binary that contains a black hole.
- **Joint SKA-LISA detections:** We expect that if a LISA detectable DCO contains a pulsar that is beaming towards Earth, SKA-1 will be able to detect at least 10 LISA DCOs that contain a NS, whilst the increased number of detectable pulsars that could crowd the sky with SKA-2 sensitivity means that this total decreases to 6. This total could be much higher if we consider that SKA and LISA could match their orbital frequency estimates to select the correct pulsar from a crowded field.
- **Importance of choice of MW model:** Many studies use Milky Way models that use fixed

metallicity populations which are assigned irrespective of birth time or position, do not account for the inside-out growth of the thin disc and use constant star formation rates. The use of these simpler MW models may lead to an overestimation of the LISA NSNS detection rate. It may also introduce unphysical artifacts into DCO parameter distributions, particularly the mass distributions, which lead to inaccurate predictions.

## ACKNOWLEDGMENTS

We thank Mike Lau, Floris Kummer, Eva Laplace, Rob Farmer, Katie Sharpe, Mathieu Renzo, Katie Breivik, Alberto Sesana, Valeriya Korol, Will Farr, Tyson Littenberg and the members of the COMPAS collaboration for insightful discussions. TW also thanks Terence Lovatt for his advice and help on improving an earlier draft of this project. This project was funded in part by the National Science Foundation under Grant No. (NSF grant number 2009131), the European Union’s Horizon 2020 research and innovation program from the European Research Council (ERC, Grant agreement No. 715063), and by the Netherlands Organization for Scientific Research (NWO) as part of the Vidi research program BinWaves with project number 639.042.728. We further acknowledge the Black Hole Initiative funded by a generous contribution of the John Templeton Foundation and the Gordon and Betty Moore Foundation.

*Software:* We used LEGWORK to evolve sources over time and calculate signal-to-noise ratios, it is freely available at <https://legwork.readthedocs.io/en/latest/>. Simulations in this paper made use of the COMPAS rapid binary population synthesis code, which is freely available at <http://github.com/TeamCOMPAS/COMPAS> (Stevenson et al. 2017; 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 Floor Broekgaarden. The authors used STROOPWAFEL from (Broekgaarden et al. 2019), publicly available at <https://github.com/FloorBroekgaarden/STROOPWAFEL><sup>8</sup>. The authors also made use of Python from the Python Software Foundation. Python Language Reference, version 3.6. Available at <http://www.python.org>. In addition the following Python packages were used: matplotlib (<https://matplotlib.org/>), NumPy (<https://numpy.org/>), SciPy (<https://www.scipy.org/>), Jupyter Lab (<https://jupyter.org/>), seaborn (<https://seaborn.pydata.org/>), h5py (<https://www.h5py.org/>) and Astropy (<http://www.astropy.org>). This research has made use of NASA’s Astrophysics Data System Bibliographic Services. We also made use of the computational facilities from the Harvard FAS Research Computing cluster.

<sup>8</sup>For the latest pip installable version of STROOPWAFEL please contact Floor Broekgaarden.

## REFERENCES

- Abadie, J., Abbott, B. P., Abbott, R., et al. 2010, Classical and Quantum Gravity, 27, 173001, doi: [10.1088/0264-9381/27/17/173001](https://doi.org/10.1088/0264-9381/27/17/173001)
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, Physical Review X, 6, 041015, doi: [10.1103/PhysRevX.6.041015](https://doi.org/10.1103/PhysRevX.6.041015)
- . 2019, Physical Review X, 9, 031040, doi: [10.1103/PhysRevX.9.031040](https://doi.org/10.1103/PhysRevX.9.031040)
- . 2020a, ApJL, 892, L3, doi: [10.3847/2041-8213/ab75f5](https://doi.org/10.3847/2041-8213/ab75f5)
- . 2017a, PhRvL, 119, 161101, doi: [10.1103/PhysRevLett.119.161101](https://doi.org/10.1103/PhysRevLett.119.161101)
- . 2017b, ApJL, 850, L40, doi: [10.3847/2041-8213/aa93fc](https://doi.org/10.3847/2041-8213/aa93fc)
- Abbott, R., Abbott, T. D., Abraham, S., et al. 2020b, arXiv e-prints, arXiv:2010.14527. <https://arxiv.org/abs/2010.14527>
- . 2020c, ApJL, 896, L44, doi: [10.3847/2041-8213/ab960f](https://doi.org/10.3847/2041-8213/ab960f)
- . 2021, ApJL, 913, L7, doi: [10.3847/2041-8213/abe949](https://doi.org/10.3847/2041-8213/abe949)
- Abdul-Masih, M., Banyard, G., Bodensteiner, J., et al. 2020, Nature, 580, E11, doi: [10.1038/s41586-020-2216-x](https://doi.org/10.1038/s41586-020-2216-x)
- Abt, H. A. 1983, ARA&A, 21, 343, doi: [10.1146/annurev.aa.21.090183.002015](https://doi.org/10.1146/annurev.aa.21.090183.002015)
- Amaro-Seoane, P., Audley, H., Babak, S., et al. 2017, arXiv e-prints, arXiv:1702.00786. <https://arxiv.org/abs/1702.00786>
- Antonini, F., & Rasio, F. A. 2016, ApJ, 831, 187, doi: [10.3847/0004-637X/831/2/187](https://doi.org/10.3847/0004-637X/831/2/187)
- Antonini, F., Toonen, S., & Hamers, A. S. 2017, ApJ, 841, 77, doi: [10.3847/1538-4357/aa6f5e](https://doi.org/10.3847/1538-4357/aa6f5e)
- Banerjee, S. 2020, PhRvD, 102, 103002, doi: [10.1103/PhysRevD.102.103002](https://doi.org/10.1103/PhysRevD.102.103002)
- Barack, L., & Cutler, C. 2004, PhRvD, 69, 082005, doi: [10.1103/PhysRevD.69.082005](https://doi.org/10.1103/PhysRevD.69.082005)
- Bates, S. D., Lorimer, D. R., Rane, A., & Swiggum, J. 2014, MNRAS, 439, 2893, doi: [10.1093/mnras/stu157](https://doi.org/10.1093/mnras/stu157)
- Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, Nature, 287, 307, doi: [10.1038/287307a0](https://doi.org/10.1038/287307a0)
- Belczynski, K., Benacquista, M., & Bulik, T. 2010, ApJ, 725, 816, doi: [10.1088/0004-637X/725/1/816](https://doi.org/10.1088/0004-637X/725/1/816)
- Belczynski, K., Holz, D. E., Bulik, T., & O’Shaughnessy, R. 2016, Nature, 534, 512, doi: [10.1038/nature18322](https://doi.org/10.1038/nature18322)
- Belczynski, K., Kalogera, V., Rasio, F. A., et al. 2008, ApJS, 174, 223, doi: [10.1086/521026](https://doi.org/10.1086/521026)
- Belczynski, K., Taam, R. E., Kalogera, V., Rasio, F. A., & Bulik, T. 2007, ApJ, 662, 504, doi: [10.1086/513562](https://doi.org/10.1086/513562)
- Berger, E. 2014, ARA&A, 52, 43, doi: [10.1146/annurev-astro-081913-035926](https://doi.org/10.1146/annurev-astro-081913-035926)
- Bertelli, G., Bressan, A., Chiosi, C., Fagotto, F., & Nasi, E. 1994, A&AS, 106, 275
- Blaauw, A. 1961, BAN, 15, 265
- Bolton, C. T. 1972, Nature, 235, 271, doi: [10.1038/235271b0](https://doi.org/10.1038/235271b0)
- Bovy, J., Leung, H. W., Hunt, J. A. S., et al. 2019, MNRAS, 490, 4740, doi: [10.1093/mnras/stz2891](https://doi.org/10.1093/mnras/stz2891)
- Bovy, J., Rix, H.-W., Schlafly, E. F., et al. 2016, ApJ, 823, 30, doi: [10.3847/0004-637X/823/1/30](https://doi.org/10.3847/0004-637X/823/1/30)
- Brandt, N., & Podsiadlowski, P. 1995, MNRAS, 274, 461, doi: [10.1093/mnras/274.2.461](https://doi.org/10.1093/mnras/274.2.461)
- Breivik, K., Coughlin, S., Zevin, M., et al. 2020, ApJ, 898, 71, doi: [10.3847/1538-4357/ab9d85](https://doi.org/10.3847/1538-4357/ab9d85)
- Breivik, K., Rodriguez, C. L., Larson, S. L., Kalogera, V., & Rasio, F. A. 2016, ApJL, 830, L18, doi: [10.3847/2041-8205/830/1/L18](https://doi.org/10.3847/2041-8205/830/1/L18)
- Broekgaarden, F. S., Berger, E., Neijssel, C. J., et al. 2021, arXiv e-prints, arXiv:2103.02608. <https://arxiv.org/abs/2103.02608>
- Broekgaarden, F. S., Justham, S., de Mink, S. E., et al. 2019, MNRAS, 490, 5228, doi: [10.1093/mnras/stz2558](https://doi.org/10.1093/mnras/stz2558)
- Chandrasekhar, S. 1931, ApJ, 74, 81, doi: [10.1086/143324](https://doi.org/10.1086/143324)
- Chattopadhyay, D., Stevenson, S., Hurley, J. R., Bailes, M., & Broekgaarden, F. 2020, arXiv e-prints, arXiv:2011.13503. <https://arxiv.org/abs/2011.13503>
- Colpi, M., Holley-Bockelmann, K., Bogdanovic, T., et al. 2019, arXiv e-prints, arXiv:1903.06867. <https://arxiv.org/abs/1903.06867>
- Corral-Santana, J. M., Casares, J., Muñoz-Darias, T., et al. 2016, A&A, 587, A61, doi: [10.1051/0004-6361/201527130](https://doi.org/10.1051/0004-6361/201527130)
- de Kool, M. 1990, ApJ, 358, 189, doi: [10.1086/168974](https://doi.org/10.1086/168974)
- de Mink, S. E., & Belczynski, K. 2015, ApJ, 814, 58, doi: [10.1088/0004-637X/814/1/58](https://doi.org/10.1088/0004-637X/814/1/58)
- de Mink, S. E., Cantiello, M., Langer, N., et al. 2009, A&A, 497, 243, doi: [10.1051/0004-6361/200811439](https://doi.org/10.1051/0004-6361/200811439)
- Dewi, J. D. M., Pols, O. R., Savonije, G. J., & van den Heuvel, E. P. J. 2002, MNRAS, 331, 1027, doi: [10.1046/j.1365-8711.2002.05257.x](https://doi.org/10.1046/j.1365-8711.2002.05257.x)
- Di Carlo, U. N., Mapelli, M., Giacobbo, N., et al. 2020, MNRAS, 498, 495, doi: [10.1093/mnras/staa2286](https://doi.org/10.1093/mnras/staa2286)
- du Buisson, L., Marchant, P., Podsiadlowski, P., et al. 2020, MNRAS, 499, 5941, doi: [10.1093/mnras/staa3225](https://doi.org/10.1093/mnras/staa3225)
- Eichler, D., Livio, M., Piran, T., & Schramm, D. N. 1989, Nature, 340, 126, doi: [10.1038/340126a0](https://doi.org/10.1038/340126a0)
- El-Badry, K., & Quataert, E. 2020, MNRAS, 493, L22, doi: [10.1093/mnrasl/slaa004](https://doi.org/10.1093/mnrasl/slaa004)
- Eldridge, J. J., Stanway, E. R., Breivik, K., et al. 2020, MNRAS, 495, 2786, doi: [10.1093/mnras/staa1324](https://doi.org/10.1093/mnras/staa1324)
- Farmer, R., Renzo, M., de Mink, S. E., Marchant, P., & Justham, S. 2019, ApJ, 887, 53, doi: [10.3847/1538-4357/ab518b](https://doi.org/10.3847/1538-4357/ab518b)
- Farr, W. M., Sravan, N., Cantrell, A., et al. 2011, ApJ, 741, 103, doi: [10.1088/0004-637X/741/2/103](https://doi.org/10.1088/0004-637X/741/2/103)

- Finn, L. S., & Thorne, K. S. 2000, PhRvD, 62, 124021, doi: [10.1103/PhysRevD.62.124021](https://doi.org/10.1103/PhysRevD.62.124021)
- Frankel, N., Rix, H.-W., Ting, Y.-S., Ness, M., & Hogg, D. W. 2018, ApJ, 865, 96, doi: [10.3847/1538-4357/aadba5](https://doi.org/10.3847/1538-4357/aadba5)
- Fryer, C. L., Belczynski, K., Ramirez-Ruiz, E., et al. 2015, ApJ, 812, 24, doi: [10.1088/0004-637X/812/1/24](https://doi.org/10.1088/0004-637X/812/1/24)
- Fryer, C. L., Belczynski, K., Wiktorowicz, G., et al. 2012, ApJ, 749, 91, doi: [10.1088/0004-637X/749/1/91](https://doi.org/10.1088/0004-637X/749/1/91)
- Gerosa, D., Ma, S., Wong, K. W. K., et al. 2019, PhRvD, 99, 103004, doi: [10.1103/PhysRevD.99.103004](https://doi.org/10.1103/PhysRevD.99.103004)
- Griffith, E., Weinberg, D. H., Johnson, J. A., et al. 2021, ApJ, 909, 77, doi: [10.3847/1538-4357/abdf6be](https://doi.org/10.3847/1538-4357/abdf6be)
- Hamada, T., & Salpeter, E. E. 1961, ApJ, 134, 683, doi: [10.1086/147195](https://doi.org/10.1086/147195)
- Hamann, W. R., Gräfener, G., Liermann, A., et al. 2019, A&A, 625, A57, doi: [10.1051/0004-6361/201834850](https://doi.org/10.1051/0004-6361/201834850)
- Hamann, W. R., & Koesterke, L. 1998, A&A, 335, 1003
- Hjellming, M. S., & Webbink, R. F. 1987, ApJ, 318, 794, doi: [10.1086/165412](https://doi.org/10.1086/165412)
- Hobbs, G., Lorimer, D. R., Lyne, A. G., & Kramer, M. 2005, MNRAS, 360, 974, doi: [10.1111/j.1365-2966.2005.09087.x](https://doi.org/10.1111/j.1365-2966.2005.09087.x)
- Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, MNRAS, 445, 581, doi: [10.1093/mnras/stu1738](https://doi.org/10.1093/mnras/stu1738)
- Hotokezaka, K., Nissanke, S., Hallinan, G., et al. 2016, ApJ, 831, 190, doi: [10.3847/0004-637X/831/2/190](https://doi.org/10.3847/0004-637X/831/2/190)
- Hulse, R. A., & Taylor, J. H. 1975, ApJL, 195, L51, doi: [10.1086/181708](https://doi.org/10.1086/181708)
- Hurley, J. R., Pols, O. R., & Tout, C. A. 2000, MNRAS, 315, 543, doi: [10.1046/j.1365-8711.2000.03426.x](https://doi.org/10.1046/j.1365-8711.2000.03426.x)
- Hurley, J. R., Tout, C. A., & Pols, O. R. 2002, MNRAS, 329, 897, doi: [10.1046/j.1365-8711.2002.05038.x](https://doi.org/10.1046/j.1365-8711.2002.05038.x)
- Ivanova, N., Heinke, C. O., Rasio, F. A., Belczynski, K., & Fregeau, J. M. 2008, MNRAS, 386, 553, doi: [10.1111/j.1365-2966.2008.13064.x](https://doi.org/10.1111/j.1365-2966.2008.13064.x)
- Ivanova, N., & Taam, R. E. 2004, ApJ, 601, 1058, doi: [10.1086/380561](https://doi.org/10.1086/380561)
- Kalogera, V., & Baym, G. 1996, ApJL, 470, L61, doi: [10.1086/310296](https://doi.org/10.1086/310296)
- Kalogera, V., Belczynski, K., Kim, C., O'Shaughnessy, R., & Willems, B. 2007, PhR, 442, 75, doi: [10.1016/j.physrep.2007.02.008](https://doi.org/10.1016/j.physrep.2007.02.008)
- Keane, E., Bhattacharyya, B., Kramer, M., et al. 2015, in Advancing Astrophysics with the Square Kilometre Array (AASKA14), 40. <https://arxiv.org/abs/1501.00056>
- Klein, A., Barausse, E., Sesana, A., et al. 2016, PhRvD, 93, 024003, doi: [10.1103/PhysRevD.93.024003](https://doi.org/10.1103/PhysRevD.93.024003)
- Klencki, J., Nelemans, G., Istrate, A. G., & Chruslinska, M. 2021, A&A, 645, A54, doi: [10.1051/0004-6361/202038707](https://doi.org/10.1051/0004-6361/202038707)
- Korol, V., Rossi, E. M., & Barausse, E. 2019, MNRAS, 483, 5518, doi: [10.1093/mnras/sty3440](https://doi.org/10.1093/mnras/sty3440)
- Korol, V., Rossi, E. M., Groot, P. J., et al. 2017, MNRAS, 470, 1894, doi: [10.1093/mnras/stx1285](https://doi.org/10.1093/mnras/stx1285)
- Korol, V., & Safarzadeh, M. 2021, MNRAS, 502, 5576, doi: [10.1093/mnras/stab310](https://doi.org/10.1093/mnras/stab310)
- Korol, V., Toonen, S., Klein, A., et al. 2020, A&A, 638, A153, doi: [10.1051/0004-6361/202037764](https://doi.org/10.1051/0004-6361/202037764)
- Kreidberg, L., Bailyn, C. D., Farr, W. M., & Kalogera, V. 2012, ApJ, 757, 36, doi: [10.1088/0004-637X/757/1/36](https://doi.org/10.1088/0004-637X/757/1/36)
- Kremer, K., Chatterjee, S., Breivik, K., et al. 2018, PhRvL, 120, 191103, doi: [10.1103/PhysRevLett.120.191103](https://doi.org/10.1103/PhysRevLett.120.191103)
- Kroupa, P. 2001, MNRAS, 322, 231, doi: [10.1046/j.1365-8711.2001.04022.x](https://doi.org/10.1046/j.1365-8711.2001.04022.x)
- Kruckow, M. U., Tauris, T. M., Langer, N., Kramer, M., & Izzard, R. G. 2018, MNRAS, 481, 1908, doi: [10.1093/mnras/sty2190](https://doi.org/10.1093/mnras/sty2190)
- Kupfer, T., Korol, V., Shah, S., et al. 2018, MNRAS, 480, 302, doi: [10.1093/mnras/sty1545](https://doi.org/10.1093/mnras/sty1545)
- Kyutoku, K., Kiuchi, K., Sekiguchi, Y., Shibata, M., & Taniguchi, K. 2018, PhRvD, 97, 023009, doi: [10.1103/PhysRevD.97.023009](https://doi.org/10.1103/PhysRevD.97.023009)
- Kyutoku, K., Nishino, Y., & Seto, N. 2019, MNRAS, 483, 2615, doi: [10.1093/mnras/sty3322](https://doi.org/10.1093/mnras/sty3322)
- Lamberts, A., Blunt, S., Littenberg, T. B., et al. 2019, MNRAS, 490, 5888, doi: [10.1093/mnras/stz2834](https://doi.org/10.1093/mnras/stz2834)
- Lamberts, A., Garrison-Kimmel, S., Hopkins, P. F., et al. 2018, MNRAS, 480, 2704, doi: [10.1093/mnras/sty2035](https://doi.org/10.1093/mnras/sty2035)
- Lau, M. Y. M., Mandel, I., Vigna-Gómez, A., et al. 2020, MNRAS, 492, 3061, doi: [10.1093/mnras/staa002](https://doi.org/10.1093/mnras/staa002)
- Li, L.-X., & Paczyński, B. 1998, ApJL, 507, L59, doi: [10.1086/311680](https://doi.org/10.1086/311680)
- Liequia, T. C., & Newman, J. A. 2015, ApJ, 806, 96, doi: [10.1088/0004-637X/806/1/96](https://doi.org/10.1088/0004-637X/806/1/96)
- Liu, J. 2009, MNRAS, 400, 1850, doi: [10.1111/j.1365-2966.2009.15574.x](https://doi.org/10.1111/j.1365-2966.2009.15574.x)
- Liu, J., Soria, R., Zheng, Z., et al. 2020, Nature, 580, E16, doi: [10.1038/s41586-020-2217-9](https://doi.org/10.1038/s41586-020-2217-9)
- Liu, J., Zhang, H., Howard, A. W., et al. 2019, Nature, 575, 618, doi: [10.1038/s41586-019-1766-2](https://doi.org/10.1038/s41586-019-1766-2)
- Liu, J., & Zhang, Y. 2014, PASP, 126, 211, doi: [10.1086/675721](https://doi.org/10.1086/675721)
- Lyne, A. G., & Lorimer, D. R. 1994, Nature, 369, 127, doi: [10.1038/369127a0](https://doi.org/10.1038/369127a0)
- Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, AJ, 154, 94, doi: [10.3847/1538-3881/aa784d](https://doi.org/10.3847/1538-3881/aa784d)
- Manchester, R. N., Hobbs, G. B., Teoh, A., & Hobbs, M. 2005, AJ, 129, 1993, doi: [10.1086/428488](https://doi.org/10.1086/428488)
- Mandel, I., & de Mink, S. E. 2016, MNRAS, 458, 2634, doi: [10.1093/mnras/stw379](https://doi.org/10.1093/mnras/stw379)

- Mandel, I., & Müller, B. 2020, MNRAS, 499, 3214, doi: [10.1093/mnras/staa3043](https://doi.org/10.1093/mnras/staa3043)
- Mandel, I., Sesana, A., & Vecchio, A. 2018, Classical and Quantum Gravity, 35, 054004, doi: [10.1088/1361-6382/aaa7e0](https://doi.org/10.1088/1361-6382/aaa7e0)
- Marchant, P., Langer, N., Podsiadlowski, P., et al. 2017, A&A, 604, A55, doi: [10.1051/0004-6361/201630188](https://doi.org/10.1051/0004-6361/201630188)
- Marchant, P., Langer, N., Podsiadlowski, P., Tauris, T. M., & Moriya, T. J. 2016, A&A, 588, A50, doi: [10.1051/0004-6361/201628133](https://doi.org/10.1051/0004-6361/201628133)
- Marchant, P., Renzo, M., Farmer, R., et al. 2019, ApJ, 882, 36, doi: [10.3847/1538-4357/ab3426](https://doi.org/10.3847/1538-4357/ab3426)
- Margalit, B., & Metzger, B. D. 2017, ApJL, 850, L19, doi: [10.3847/2041-8213/aa991c](https://doi.org/10.3847/2041-8213/aa991c)
- Massevitch, A., & Yungelson, L. 1975, Mem. Soc. Astron. Italiana, 46, 217
- McKernan, B., Ford, K. E. S., & O’Shaughnessy, R. 2020, MNRAS, 498, 4088, doi: [10.1093/mnras/staa2681](https://doi.org/10.1093/mnras/staa2681)
- McMillan, P. J. 2011, MNRAS, 414, 2446, doi: [10.1111/j.1365-2966.2011.18564.x](https://doi.org/10.1111/j.1365-2966.2011.18564.x)
- Metzger, B. D. 2017, Living Reviews in Relativity, 20, 3, doi: [10.1007/s41114-017-0006-z](https://doi.org/10.1007/s41114-017-0006-z)
- Miller, M. C., & Lauburg, V. M. 2009, ApJ, 692, 917, doi: [10.1088/0004-637X/692/1/917](https://doi.org/10.1088/0004-637X/692/1/917)
- Miller-Jones, J. C. A., Bahramian, A., Orosz, J. A., et al. 2021, Science, 371, 1046, doi: [10.1126/science.abb3363](https://doi.org/10.1126/science.abb3363)
- Morris, M. 1993, ApJ, 408, 496, doi: [10.1086/172607](https://doi.org/10.1086/172607)
- Narayan, R., Piran, T., & Shemi, A. 1991, ApJL, 379, L17, doi: [10.1086/186143](https://doi.org/10.1086/186143)
- Nelemans, G., Yungelson, L. R., & Portegies Zwart, S. F. 2001, A&A, 375, 890, doi: [10.1051/0004-6361:20010683](https://doi.org/10.1051/0004-6361:20010683)
- Nishizawa, A., Berti, E., Klein, A., & Sesana, A. 2016, PhRvD, 94, 064020, doi: [10.1103/PhysRevD.94.064020](https://doi.org/10.1103/PhysRevD.94.064020)
- Nissanke, S., Vallisneri, M., Nelemans, G., & Prince, T. A. 2012, ApJ, 758, 131, doi: [10.1088/0004-637X/758/2/131](https://doi.org/10.1088/0004-637X/758/2/131)
- Nomoto, K. 1984, ApJ, 277, 791, doi: [10.1086/161749](https://doi.org/10.1086/161749)
- . 1987, ApJ, 322, 206, doi: [10.1086/165716](https://doi.org/10.1086/165716)
- Öpik, E. 1924, Publications of the Tartu Astrofizica Observatory, 25, 1
- Özel, F., Psaltis, D., Narayan, R., & McClintock, J. E. 2010, ApJ, 725, 1918, doi: [10.1088/0004-637X/725/2/1918](https://doi.org/10.1088/0004-637X/725/2/1918)
- Paczyński, B., & Sienkiewicz, R. 1972, AcA, 22, 73
- Peters, P. C. 1964, Physical Review, 136, 1224, doi: [10.1103/PhysRev.136.B1224](https://doi.org/10.1103/PhysRev.136.B1224)
- Peters, P. C., & Mathews, J. 1963, Physical Review, 131, 435, doi: [10.1103/PhysRev.131.435](https://doi.org/10.1103/PhysRev.131.435)
- Pfahl, E., Podsiadlowski, P., & Rappaport, S. 2005, ApJ, 628, 343, doi: [10.1086/430515](https://doi.org/10.1086/430515)
- Pfahl, E., Rappaport, S., Podsiadlowski, P., & Spruit, H. 2002, ApJ, 574, 364, doi: [10.1086/340794](https://doi.org/10.1086/340794)
- Podsiadlowski, P., Langer, N., Poelarends, A. J. T., et al. 2004, ApJ, 612, 1044, doi: [10.1086/421713](https://doi.org/10.1086/421713)
- Pols, O. R., Schröder, K.-P., Hurley, J. R., Tout, C. A., & Eggleton, P. P. 1998, MNRAS, 298, 525, doi: [10.1046/j.1365-8711.1998.01658.x](https://doi.org/10.1046/j.1365-8711.1998.01658.x)
- Portegies Zwart, S. F., & McMillan, S. L. W. 2000, ApJL, 528, L17, doi: [10.1086/312422](https://doi.org/10.1086/312422)
- Rastello, S., Mapelli, M., Di Carlo, U. N., et al. 2020, MNRAS, 497, 1563, doi: [10.1093/mnras/staa2018](https://doi.org/10.1093/mnras/staa2018)
- Rastello, S., Mapelli, M., di Carlo, U. N., et al. 2021, arXiv e-prints, arXiv:2105.01669, <https://arxiv.org/abs/2105.01669>
- Robson, T., Cornish, N. J., & Liu, C. 2019, Classical and Quantum Gravity, 36, 105011, doi: [10.1088/1361-6382/ab1101](https://doi.org/10.1088/1361-6382/ab1101)
- Rodriguez, C. L., Amaro-Seoane, P., Chatterjee, S., & Rasio, F. A. 2018, PhRvL, 120, 151101, doi: [10.1103/PhysRevLett.120.151101](https://doi.org/10.1103/PhysRevLett.120.151101)
- Rodriguez, C. L., Morscher, M., Pattabiraman, B., et al. 2015, PhRvL, 115, 051101, doi: [10.1103/PhysRevLett.115.051101](https://doi.org/10.1103/PhysRevLett.115.051101)
- Ruiter, A. J., Belczynski, K., Benacquista, M., Larson, S. L., & Williams, G. 2010, ApJ, 717, 1006, doi: [10.1088/0004-637X/717/2/1006](https://doi.org/10.1088/0004-637X/717/2/1006)
- Sana, H., de Mink, S. E., de Koter, A., et al. 2012, Science, 337, 444, doi: [10.1126/science.1223344](https://doi.org/10.1126/science.1223344)
- Schuster, E. F. 1985, Communications in Statistics-theory and Methods, 14, 1123
- Sellwood, J. A., & Binney, J. J. 2002, MNRAS, 336, 785, doi: [10.1046/j.1365-8711.2002.05806.x](https://doi.org/10.1046/j.1365-8711.2002.05806.x)
- Sesana, A. 2016, PhRvL, 116, 231102, doi: [10.1103/PhysRevLett.116.231102](https://doi.org/10.1103/PhysRevLett.116.231102)
- Sesana, A., Lamberts, A., & Petiteau, A. 2020, MNRAS, 494, L75, doi: [10.1093/mnrasl/slaa039](https://doi.org/10.1093/mnrasl/slaa039)
- Shao, Y., & Li, X.-D. 2021, arXiv e-prints, arXiv:2107.03565, <https://arxiv.org/abs/2107.03565>
- Sharma, S., Hayden, M. R., & Bland-Hawthorn, J. 2020, arXiv e-prints, arXiv:2005.03646, <https://arxiv.org/abs/2005.03646>
- Shenar, T., Bodensteiner, J., Abdul-Masih, M., et al. 2020, A&A, 639, L6, doi: [10.1051/0004-6361/202038275](https://doi.org/10.1051/0004-6361/202038275)
- Shenar, T., Sablowski, D. P., Hainich, R., et al. 2019, A&A, 627, A151, doi: [10.1051/0004-6361/201935684](https://doi.org/10.1051/0004-6361/201935684)
- Sigurdsson, S., & Hernquist, L. 1993, Nature, 364, 423, doi: [10.1038/364423a0](https://doi.org/10.1038/364423a0)
- Silsbee, K., & Tremaine, S. 2017, ApJ, 836, 39, doi: [10.3847/1538-4357/aa5729](https://doi.org/10.3847/1538-4357/aa5729)

- Smarr, L. L., & Blandford, R. 1976, ApJ, 207, 574, doi: [10.1086/154524](https://doi.org/10.1086/154524)
- Smits, R., Kramer, M., Stappers, B., et al. 2009, A&A, 493, 1161, doi: [10.1051/0004-6361:200810383](https://doi.org/10.1051/0004-6361:200810383)
- Snaith, O. N., Haywood, M., Di Matteo, P., et al. 2014, ApJL, 781, L31, doi: [10.1088/2041-8205/781/2/L31](https://doi.org/10.1088/2041-8205/781/2/L31)
- Soberman, G. E., Phinney, E. S., & van den Heuvel, E. P. J. 1997, A&A, 327, 620. <https://arxiv.org/abs/astro-ph/9703016>
- Srinivasan, G. 1989, A&A Rv, 1, 209, doi: [10.1007/BF00873079](https://doi.org/10.1007/BF00873079)
- Stephan, A. P., Naoz, S., Ghez, A. M., et al. 2016, MNRAS, 460, 3494, doi: [10.1093/mnras/stw1220](https://doi.org/10.1093/mnras/stw1220)
- Stevenson, S., Sampson, M., Powell, J., et al. 2019, ApJ, 882, 121, doi: [10.3847/1538-4357/ab3981](https://doi.org/10.3847/1538-4357/ab3981)
- Stevenson, S., Vigna-Gómez, A., Mandel, I., et al. 2017, Nature Communications, 8, 14906, doi: [10.1038/ncomms14906](https://doi.org/10.1038/ncomms14906)
- Taam, R. E., & Sandquist, E. L. 2000, ARA&A, 38, 113, doi: [10.1146/annurev.astro.38.1.113](https://doi.org/10.1146/annurev.astro.38.1.113)
- Takahashi, R., & Seto, N. 2002, ApJ, 575, 1030, doi: [10.1086/341483](https://doi.org/10.1086/341483)
- Tauris, T. M., Kramer, M., Freire, P. C. C., et al. 2017, ApJ, 846, 170, doi: [10.3847/1538-4357/aa7e89](https://doi.org/10.3847/1538-4357/aa7e89)
- Tauris, T. M., Langer, N., Moriya, T. J., et al. 2013, ApJL, 778, L23, doi: [10.1088/2041-8205/778/2/L23](https://doi.org/10.1088/2041-8205/778/2/L23)
- Tauris, T. M., Langer, N., & Podsiadlowski, P. 2015, MNRAS, 451, 2123, doi: [10.1093/mnras/stv990](https://doi.org/10.1093/mnras/stv990)
- The LIGO Scientific Collaboration, the Virgo Collaboration, the KAGRA Collaboration, et al. 2021, arXiv e-prints, arXiv:2106.15163. <https://arxiv.org/abs/2106.15163>
- Thrane, E., Osłowski, S., & Lasky, P. D. 2020, MNRAS, 493, 5408, doi: [10.1093/mnras/staa593](https://doi.org/10.1093/mnras/staa593)
- Timmes, F. X., Woosley, S. E., & Weaver, T. A. 1996, ApJ, 457, 834, doi: [10.1086/176778](https://doi.org/10.1086/176778)
- Toonen, S., Claeys, J. S. W., Mennekens, N., & Ruiter, A. J. 2014, A&A, 562, A14, doi: [10.1051/0004-6361/201321576](https://doi.org/10.1051/0004-6361/201321576)
- Toonen, S., Portegies Zwart, S., Hamers, A. S., & Bandopadhyay, D. 2020, A&A, 640, A16, doi: [10.1051/0004-6361/201936835](https://doi.org/10.1051/0004-6361/201936835)
- Tout, C. A., Pols, O. R., Eggleton, P. P., & Han, Z. 1996, MNRAS, 281, 257, doi: [10.1093/mnras/281.1.257](https://doi.org/10.1093/mnras/281.1.257)
- Tutukov, A., & Yungelson, L. 1973, Nauchnye Informatsii, 27, 70
- Tutukov, A. V., & Yungelson, L. R. 1993, MNRAS, 260, 675, doi: [10.1093/mnras/260.3.675](https://doi.org/10.1093/mnras/260.3.675)
- van den Heuvel, E. P. J. 2011, Compact stars and the evolution of binary systems (World Scientific Publishing Co. Pte. Ltd), 55–73, doi: [10.1142/9789814374774\\_0006](https://doi.org/10.1142/9789814374774_0006)
- van den Heuvel, E. P. J., Portegies Zwart, S. F., & de Mink, S. E. 2017, MNRAS, 471, 4256, doi: [10.1093/mnras/stx1430](https://doi.org/10.1093/mnras/stx1430)
- Vigna-Gómez, A., Neijssel, C. J., Stevenson, S., et al. 2018, MNRAS, 481, 4009, doi: [10.1093/mnras/sty2463](https://doi.org/10.1093/mnras/sty2463)
- Vink, J. S. 2017, A&A, 607, L8, doi: [10.1051/0004-6361/201731902](https://doi.org/10.1051/0004-6361/201731902)
- Vink, J. S., & de Koter, A. 2005, A&A, 442, 587, doi: [10.1051/0004-6361:20052862](https://doi.org/10.1051/0004-6361:20052862)
- Vink, J. S., de Koter, A., & Lamers, H. J. G. L. M. 2000, A&A, 362, 295. <https://arxiv.org/abs/astro-ph/0008183>
- . 2001, A&A, 369, 574, doi: [10.1051/0004-6361:20010127](https://doi.org/10.1051/0004-6361:20010127)
- Webbink, R. F. 1984, ApJ, 277, 355, doi: [10.1086/161701](https://doi.org/10.1086/161701)
- Webster, B. L., & Murdin, P. 1972, Nature, 235, 37, doi: [10.1038/235037a0](https://doi.org/10.1038/235037a0)
- Wegg, C., Gerhard, O., & Portail, M. 2015, MNRAS, 450, 4050, doi: [10.1093/mnras/stv745](https://doi.org/10.1093/mnras/stv745)
- Willemse, B., Kalogera, V., Vecchio, A., et al. 2007, ApJL, 665, L59, doi: [10.1086/521049](https://doi.org/10.1086/521049)
- Woosley, S. E., Blinnikov, S., & Heger, A. 2007, Nature, 450, 390, doi: [10.1038/nature06333](https://doi.org/10.1038/nature06333)
- Xu, X.-J., & Li, X.-D. 2010a, ApJ, 716, 114, doi: [10.1088/0004-637X/716/1/114](https://doi.org/10.1088/0004-637X/716/1/114)
- . 2010b, ApJ, 722, 1985, doi: [10.1088/0004-637X/722/2/1985](https://doi.org/10.1088/0004-637X/722/2/1985)
- Yu, S., & Jeffery, C. S. 2010, A&A, 521, A85, doi: [10.1051/0004-6361/201014827](https://doi.org/10.1051/0004-6361/201014827)
- Ziosi, B. M., Mapelli, M., Branchesi, M., & Tormen, G. 2014, MNRAS, 441, 3703, doi: [10.1093/mnras/stu824](https://doi.org/10.1093/mnras/stu824)

## APPENDIX

## A. POPULATION SYNTHESIS

In this section we summarise the main assumptions and settings that we use when performing population synthesis for this work. For a more general overview of every setting see Broekgaarden et al. (2021).

A.1. *Initial conditions*

We simulate between 1 and 100 million massive binaries for each of 50 metallicities equally spaced in log space between  $Z \in [0.0001, 0.022]$ , where  $Z$  is the mass fraction of heavy elements. We simulate more binaries for higher metallicities so that large enough sample of DCOs at each metallicity (since DCOs are formed at a lower rate at higher metallicities). These metallicities span the allowed metallicity range for the original fitting formulae on which COMPAS is based (Hurley et al. 2000). This is repeated for 19 physics variations (see Section A.3) and so in total over two billion binaries were simulated.

Each binary is sampled from initial distributions for the primary and secondary masses as well as the separation. The primary mass, that is the mass of the initially more massive star, is restricted to  $m_1 \in [5, 150] M_\odot$ , which spans the range of interest for NS and BH formation in binary systems, and drawn from the Kroupa (2001) initial mass function (IMF),  $p(m_1) \propto m_1^{-2.3}$ . The secondary mass,  $m_2$ , is drawn using the initial mass ratio of the binary,  $q \equiv m_2/m_1$ , which we assume to be uniform on  $[0, 1]$ , therefore  $p(q) = 1$  (e.g. consistent with Sana et al. 2012). We additionally restrict the secondary masses  $m_2 \geq 0.1 M_\odot$ , which is approximately the minimum mass for a main sequence star. We assume that the initial separation follows a flat in the log distribution with  $p(a_i) \propto 1/a_i$  and  $a_i \in [0.01, 1000]$  AU (Öpik 1924; Abt 1983). We assume that all binary orbits are circular at birth to reduce the dimensions of initial parameters. Since we focus on post-interaction binaries which will have circularised after mass transfer we argue this is an reasonable assumption (as many studies have in the past) and is likely not critical for predicting detection rates (Hurley et al. 2002; de Mink & Belczynski 2015).

We apply the adaptive importance sampling algorithm STROOPWAFEL (Broekgaarden et al. 2019) to improve the yield of our sample. This algorithm increases the prevalence of target DCOs (BHBs, BHNSs and NSNSs in this case) in the sample and assigns each a weight,  $w$ , which represents the probability of drawing the DCO without STROOPWAFEL in effect.

A.2. *Physical assumptions in our fiducial model*

*Stellar Evolution:* To follow the evolution of massive stars, COMPAS relies on fitting formulae by Hurley et al. (2000) to detailed single star models by Pols et al. (1998). COMPAS models the evolution of stars that lose or gain mass closely following the algorithms originally described in Tout et al. (1996) and Hurley et al. (2002).

*Wind mass loss:* We follow the wind prescription from Belczynski et al. (2008), which was based on results from Monte Carlo radiative transfer simulation of Vink et al. (2000, 2001). We use the wind mass loss rates from Vink et al. (2001) for stars above 12500 K and the rates from Hurley et al. (2000) for cooler stars. Additionally, we use a separate, higher wind mass loss rate for luminous blue variable (LBV) stars, following Belczynski et al. (2008), to mimic observed LBV eruptions for stars with luminosities and effective temperatures above the Humphreys-Davidson limit. We use the Wolf-Rayet-like mass loss rate from Hamann & Koesterke (1998) with an additional metallicity scaling from Vink & de Koter (2005) for helium stars, and set  $f_{\text{WR}} = 1$ . See Team COMPAS: J. Riley et al. (in prep), Section 3 for the explicit equations.

*Mass Transfer:* In determining the stability of mass transfer we use the  $\zeta$ -prescription, which compares the radial response of the star with the response of the Roche lobe radius to the mass transfer (e.g. Hjellming & Webbink 1987). The mass transfer efficiency,  $\beta \equiv \Delta M_{\text{acc}}/\Delta M_{\text{don}}$ , is defined as the fraction of the mass transferred by the donor that is actually accreted by the accretor. We limit the maximum accretion rate for stars to  $\Delta M_{\text{acc}}/\Delta t \leq 10 M_{\text{acc}}/\tau_{\text{KH}}$ , where  $\tau_{\text{KH}}$  is the Kelvin-Helmholtz timescale of the star (Paczynski & Sienkiewicz 1972; Hurley et al. 2002). The maximum accretion rate for compact objects is limited to the Eddington accretion rate. If more mass than these rates is accreted then we assume that the excess is lost through isotropic re-emission in the vicinity of the accreting star (e.g. Massevitch & Yungelson 1975; Soberman et al. 1997). We assume that all mass transfer from a stripped post-helium-burning-star (case BB) onto a neutron star or black hole is unstable (Tauris et al. 2015).

*Common-Envelope:* A common-envelope phase follows dynamically unstable mass transfer and we parameterise this using the  $\alpha$ - $\lambda$  prescription from Webbink (1984) and de Kool (1990). We assume  $\alpha = 1$ , such that all of the gravitational binding energy is available for the ejection of the envelope. For  $\lambda$  we use

the fitting formulae from Xu & Li (2010a,b). We assume that any Hertzsprung gap donor stars that initiate a common-envelope phase will not survive this phase due to a lack of a steep density gradient between the core and envelope (Taam & Sandquist 2000; Ivanova & Taam 2004; Klencki et al. 2021). This follows the ‘pessimistic’ common-envelope scenario (c.f. Belczynski et al. 2007). We remove any binaries where the secondary immediately fills its Roche lobe upon the conclusion of the common-envelope phase as we treat these as failed common-envelope ejections, likely leading to a stellar merger.

*Supernovae:* We draw the remnant masses and natal kick magnitudes from different distributions depending on the type of supernova that occurs. For stars undergoing a general core-collapse supernova, we use the *delayed* supernova remnant mass prescription from Fryer et al. (2012). The *delayed* prescription does not reproduce a neutron star black hole mass gap and we use this as our default as it has been shown to provide a better fit for observed populations of DCOs (e.g. Vigna-Gómez et al. 2018). We draw the natal kick magnitudes from a Maxwellian velocity distribution with a one-dimensional root-mean-square velocity dispersion of  $\sigma_{\text{rms}}^{\text{1D}} = 265 \text{ km s}^{-1}$  (Lyne & Lorimer 1994; Hobbs et al. 2005). We assume that stars with helium core masses between  $1.6\text{--}2.25 M_{\odot}$  (Hurley et al. 2002) experience electron-capture supernovae (ECSN) (Nomoto 1984, 1987; Ivanova et al. 2008). We set all remnant masses to  $1.26 M_{\odot}$  in this case as an approximation of the solution to Equation 8 of Timmes et al. (1996). For these supernovae, we set  $\sigma_{\text{rms}}^{\text{1D}} = 30 \text{ km s}^{-1}$  (e.g. Pfahl et al. 2002; Podsiadlowski et al. 2004). We assume that stars that undergo case BB mass transfer (Dewi et al. 2002) experience extreme stripping which leads to an ultra-stripped supernova (Tauris et al. 2013, 2015). For these supernovae we calculate the remnant mass using the Fryer et al. (2012) prescription and use  $\sigma_{\text{rms}}^{\text{1D}} = 30 \text{ km s}^{-1}$  (as with ECSN). Stars with final helium core masses between  $35\text{--}135 M_{\odot}$  are presumed to undergo a pair-instability, or pulsational pair-instability supernova (e.g. Woosley et al. 2007; Farmer et al. 2019). We follow the prescription from Marchant et al. (2019) as implemented in (Stevenson et al. 2019) for these supernovae. We assume that kicks are isotropic in the frame of the collapsing star. We adopt a maximum neutron star mass of  $2.5 M_{\odot}$  (e.g. Kalogera & Baym 1996; Fryer et al. 2015; Margalit & Metzger 2017) for the fiducial model and change the Fryer et al. (2012) prescription accordingly.

### A.3. Model variations

In addition to our fiducial model for the formation of DCOs, we explore 19 other models in which we change various aspects of the mass transfer, common-envelope, supernova and wind mass loss physics assumptions in order to assess the effect of their uncertainties on the overall double compact object detection rates and distributions. Each of the models varies a single physics assumption (fiducial assumptions are outlined in Section A.2) and these models are outlined in Table 2.

Our fiducial model is labelled model A. Models B-F focus on changes to the mass transfer physics assumptions. We explore the effect of fixing the mass transfer efficiency  $\beta$  to a constant value, rather than allowing it to vary based on the maximum accretion rate. In models B, C, D, in which we set the value of  $\beta$  to 0.25, 0.5 and 0.75 respectively. In model E we investigate the consequence of assuming that case BB mass transfer onto a neutron star or black hole is always stable rather than always unstable.

Models G-K focus on altering the common-envelope physics. We change the common-envelope efficiency parameter to  $\alpha_{\text{CE}} = 0.1, 0.5, 2.0, 10.0$  in models G, H, I and J respectively. In model K, we relax our restriction that Hertzsprung gap donor stars cannot survive common-envelope events, thereby following the ‘optimistic’ common-envelope scenario. We combine this with model E in model F.

In models L-R we consider changes related to our assumptions about supernova physics. Model L uses the alternate *rapid* remnant mass prescription from Fryer et al. (2012) instead of the *delayed* prescription. We change the maximum neutron star mass in models M and N to 2 and  $3 M_{\odot}$  respectively to account for the range of predicted maximum neutron star masses. Model O removes the implementation of pair-instability and pulsational pair-instability supernovae. In models P and Q we decrease the root-mean-square velocity dispersion for core-collapse supernovae to explore the effect of lower kicks. Model R removes the natal kick for all black holes.

Finally, in models S-T we investigate the effect of changing our assumption about wind mass loss rates, specifically for Wolf-Rayet winds. We vary  $f_{\text{WR}}$  to 0.1 and 5.0 in models S and T respectively. These values approximately span the current range of possible Wolf-Rayet wind efficiencies suggested from observations (e.g. Vink (2017), Hamann et al. (2019), Shenar et al. (2019), Miller-Jones et al. (2021) and van Son et al. (in prep)).

Model	Physics Variation
A	Fiducial (see Section A.2)
B	Fixed mass transfer efficiency of $\beta = 0.25$
C	Fixed mass transfer efficiency of $\beta = 0.5$
D	Fixed mass transfer efficiency of $\beta = 0.75$
E	Case BB mass transfer is always unstable
F	Model E + Model K
G	CE efficiency parameter $\alpha = 0.1$
H	CE efficiency parameter $\alpha = 0.5$
I	CE efficiency parameter $\alpha = 2$
J	CE efficiency parameter $\alpha = 10$
K	HG donor stars initiating a CE survive CE
L	Fryer rapid SN remnant mass prescription
M	Maximum NS mass is fixed to $2 M_{\odot}$
N	Maximum NS mass is fixed to $3 M_{\odot}$
O	PISN and pulsational-PISN not implemented
P	$\sigma_{\text{rms}}^{\text{1D}} = 100 \text{ km s}^{-1}$ for core-collapse supernova
Q	$\sigma_{\text{rms}}^{\text{1D}} = 30 \text{ km s}^{-1}$ for core-collapse supernova
R	Black holes receive no natal kick
S	Wolf-Rayet wind factor $f_{\text{WR}} = 0.1$
T	Wolf-Rayet wind factor $f_{\text{WR}} = 5.0$

**Table 2.** A description of the 20 binary population synthesis models used in this study. A is the fiducial model, B-F change mass transfer physics, G-K change common-envelope physics, L-R change supernova physics and S-T change wind mass loss (c.f. Broekgaarden et al. 2021, Table 2).

## B. DETECTION RATE NORMALISATION

In this section we explain the normalisation process that we refer to in Section 2.3. From each simulated instance of the Milky Way we extract the fraction of targets that are detectable, where we define a target as one of BHBH, BHNS or NSNS that merges in a Hubble time. To convert the detectable fraction to a detection rate for the Milky Way, we write that the *number* of detectable targets in the Milky Way is

$$N_{\text{detect}} = f_{\text{detect}} \cdot N_{\text{target,MW}}, \quad (\text{B1})$$

where  $f_{\text{detect}}$  is the fraction of targets in the instance that were detectable and  $N_{\text{target,MW}}$  is the total number of targets that have been formed in the Milky Way’s history. We can further break this total down into

$$N_{\text{target,MW}} = \langle \mathcal{R}_{\text{target}} \rangle \cdot M_{\text{SF,MW}}, \quad (\text{B2})$$

where  $\langle \mathcal{R}_{\text{target}} \rangle$  is the average number of targets formed per star forming mass and  $M_{\text{SF,MW}}$  is the star forming mass of the Milky Way, meaning the total mass of every star ever formed in the Milky Way.

### B.1. Average target formation rate

Double compact object formation is metallicity dependent, so we find the average rate as the integral over metallicity, which is given by

$$\langle \mathcal{R}_{\text{target}} \rangle = \int_{Z_{\min}}^{Z_{\max}} p_Z \mathcal{R}_{\text{target,Z}} dZ, \quad (\text{B3})$$

where  $Z_{\min}, Z_{\max}$  are the minimum and maximum sampled metallicities,  $p_Z$  is the probability of forming a star at the metallicity  $Z$  (which can be found using the distribution in Frankel et al. 2018) and  $\mathcal{R}_{\text{target,Z}}$  is the number of targets formed per star forming mass,

$$\mathcal{R}_{\text{target,Z}} = \frac{N_{\text{target,Z}}}{M_{\text{SF,Z}}}. \quad (\text{B4})$$

In practice, this integral is instead approximated as a sum over the metallicity bins that we use in our simulation. The number of targets in our sample at a metallicity  $Z$ ,  $N_{\text{target,Z}}$ , can be written simply as the sum of the targets’ weights:

$$N_{\text{target,Z}} = \sum_{i=1}^{N_{\text{binaries,Z}}} w_i \theta_{\text{target,i}}, \quad (\text{B5})$$

where  $w_i$  is the binary’s adaptive importance sampling weight assigned,  $N_{\text{binaries,Z}}$  is the number of binaries at metallicity  $Z$  in our sample and  $\theta_{\text{target,i}}$  is only 1 when the binary is a target and otherwise 0.

The total star forming mass at a metallicity  $Z$ ,  $M_{\text{SF,Z}}$ , can be written as

$$M_{\text{SF,Z}} = \frac{\langle m \rangle_{\text{COMPAS,Z}}}{f_{\text{trunc}}} N_{\text{binaries,Z}}, \quad (\text{B6})$$

where  $\langle m \rangle_{\text{COMPAS}}$  is the average star forming mass of a binary in a simulation using our cutoffs (discussed in Section 2.1) and  $f_{\text{trunc}}$  is the fraction of the total stellar mass from which our COMPAS simulations sample, given our truncated mass and separation ranges (see Section 2.1). These truncations mean that only  $f_{\text{trunc}} \approx 0.17$  of the stellar mass in the galaxy is sampled from.

### B.2. Total star forming mass in the Milky Way

It is important to distinguish between the *total* mass of every star formed over the entire history of the Milky Way and the *current* stellar mass in the Milky Way. Many stars born in the Milky Way are no longer living and have lost much of their mass to stellar winds and supernovae, thus the current stellar mass in the Milky Way is an underestimate of the total star forming mass.

Licquia & Newman (2015) find that the total stellar mass today in the Milky Way is  $6.08 \pm 1.14 \times 10^{10} M_{\odot}$ .

This total includes all stars and stellar remnants (white dwarfs, neutrons stars and black holes) but *excludes* brown dwarfs. We can write that the total mass of every star every formed in the Milky Way is

$$M_{\text{SF,MW}} = (6.08 \pm 1.14) \times 10^{10} M_{\odot} \cdot \frac{\langle m \rangle_{\text{SF,total}}}{\langle m \rangle_{\text{SF,today}}}, \quad (\text{B7})$$

where  $\langle m \rangle_{\text{SF,total}}$  is the average mass of a star over the history of the Milky Way and is defined as

$$\langle m \rangle_{\text{SF,total}} = \int_0^{t_{\text{MW}}} p_{\text{birth}}(\tau) \int_{0.01}^{200} \zeta(m) m dm d\tau, \quad (\text{B8})$$

where  $t_{\text{MW}}$  is the age of the Milky Way,  $\zeta(m)$  is the Kroupa (2001) IMF function and  $p_{\text{birth}}(\tau)$  is the probability of a star being formed at a lookback time  $\tau$  (Eq. 2).  $\langle m \rangle_{\text{SF,today}}$  is the average mass of all stars and stellar remnants (excluding brown dwarfs) present in the Milky Way today is defined as follows (note that we integrate from 0.08 not 0.01 since observations of today's Milky Way mass exclude brown dwarfs)

$$\langle m \rangle_{\text{SF,today}} = \int_0^{t_{\text{MW}}} p_{\text{birth}}(\tau) \int_{0.08}^{200} \zeta(m) m_{\text{today}} dm d\tau, \quad (\text{B9})$$

where  $m_{\text{today}}(m, Z, \tau)$  is the current mass of a star that was formed  $\tau$  years ago at a metallicity  $Z$ . We calculate  $m_{\text{today}}(m, Z, \tau)$  by interpolating the final masses given by COMPAS for a grid of single stars over different masses and metallicities using the Fryer et al. (2012) delayed prescription and default wind mass loss settings. For  $Z$ , we use the average star forming metallicity in the Milky Way at a lookback time  $\tau$  using our galaxy model. Evaluating Equation B7, we find that the total mass of every star that has ever formed in the Milky Way is

$$\begin{aligned} M_{\text{SF,MW}} &= (6.1 \pm 1.1) \times 10^{10} M_{\odot} \cdot \frac{0.378 M_{\odot}}{0.221 M_{\odot}}, \\ &= (10.4 \pm 1.1) \times 10^{10} M_{\odot}, \end{aligned} \quad (\text{B10})$$

an increase of approximately 70% from the value still in stars today!

### B.3. Normalisation summary

Finally, we can substitute Equations B3 and B7 into B1 and write that the overall normalisation of the detection rate is calculated as

$$\begin{aligned} N_{\text{detect}} &= f_{\text{detect}} \cdot 10.4 \times 10^{10} M_{\odot} \\ &\times \sum_{Z=Z_{\min}}^{Z_{\max}} p_Z \left( \sum_{i=1}^{N_{\text{binaries,Z}}} w_i \theta_{\text{target,i}} \right) \\ &\times \left( \frac{\langle m \rangle_{\text{COMPAS,Z}}}{f_{\text{trunc}}} \sum_{i=1}^{N_{\text{binaries,Z}}} w_i \right)^{-1}. \end{aligned} \quad (\text{B11})$$

## C. CALCULATION OF THE UNCERTAINTIES IN THE CHIRP MASS FOR DETECTABLE SOURCES

How accurately the chirp mass of a detected binary can be determined depends on the signal to noise ratio, duration of the mission, its orbital frequency and the time derivative of the orbital frequency.

Here we describe how we estimate the uncertainty of the chirp mass. First, consider the chirp mass, which can be expressed as

$$\mathcal{M}_c = \frac{c^3}{G} \left( \frac{5\pi}{48n} \frac{f_n}{F(e)} \right)^{3/5} \frac{1}{(2\pi f_{\text{orb}})^{11/5}}, \quad (\text{C12})$$

where  $f_n$  is the frequency of the n-th harmonic,  $f_{\text{orb}}$  is the orbital frequency,  $\mathcal{M}_c$  is the chirp mass (defined in Eq. 12),  $e$  is the eccentricity and

$$F(e) = \frac{1 + \frac{73}{24}e^2 + \frac{37}{96}e^4}{(1 - e^2)^{7/2}}, \quad (\text{C13})$$

is the enhancement factor of gravitational wave emission for an eccentric binary over an otherwise identical circular binary (Peters & Mathews 1963, Eq. 17). In practice, we will use the dominating harmonic, with  $n = n_{\text{dom}}$  and  $f_n = f_{\text{dom}}$ . The dominating harmonic for circular binaries is  $n_{\text{dom}} = 2$  and the dominating frequency is twice the orbital frequency.

Therefore the chirp mass uncertainty can be estimated as

$$\frac{\Delta \mathcal{M}_c}{\mathcal{M}_c} = \frac{11}{5} \frac{\Delta f_{\text{orb}}}{f_{\text{orb}}} + \frac{3}{5} \frac{\Delta \dot{f}_{\text{dom}}}{\dot{f}_{\text{dom}}} + \frac{3}{5} \frac{\Delta F(e)}{F(e)}, \quad (\text{C14})$$

We estimate the frequency uncertainties using Takahashi & Seto (2002), such that

$$\frac{\Delta f_{\text{orb}}}{f_{\text{orb}}} = 4\sqrt{3} \cdot \frac{1}{\rho} \frac{1}{T_{\text{obs}}} \frac{1}{f_{\text{orb}}}, \quad (\text{C15})$$

$$\frac{\Delta \dot{f}_{\text{dom}}}{\dot{f}_{\text{dom}}} = 6\sqrt{5} \cdot \frac{1}{\rho} \left( \frac{1}{T_{\text{obs}}} \right)^2 \frac{1}{\dot{f}_{\text{dom}}}, \quad (\text{C16})$$

where  $\rho$  is the signal-to-noise ratio and  $T_{\text{obs}}$  is the LISA mission length. We estimate the eccentricity certainty,  $\Delta e$ , following the methods of Lau et al. (2020) and Korol & Safarzadeh (2021), which use the relative SNRs of different harmonics to work out the eccentricity. We propagate this uncertainty such that

$$\frac{\Delta F(e)}{F(e)} = \Delta e \cdot \frac{(1256 + 1608e^2 + 111e^4)e}{96 + 196e^2 - 255e^4 - 37e^6}. \quad (\text{C17})$$

We use Eq. C14 to calculate the chirp mass uncertainty for each DCO type in our sample and plot it in Fig. 6.

#### D. ASSESSING THE IMPACT OF MILKY WAY MODEL CHOICES

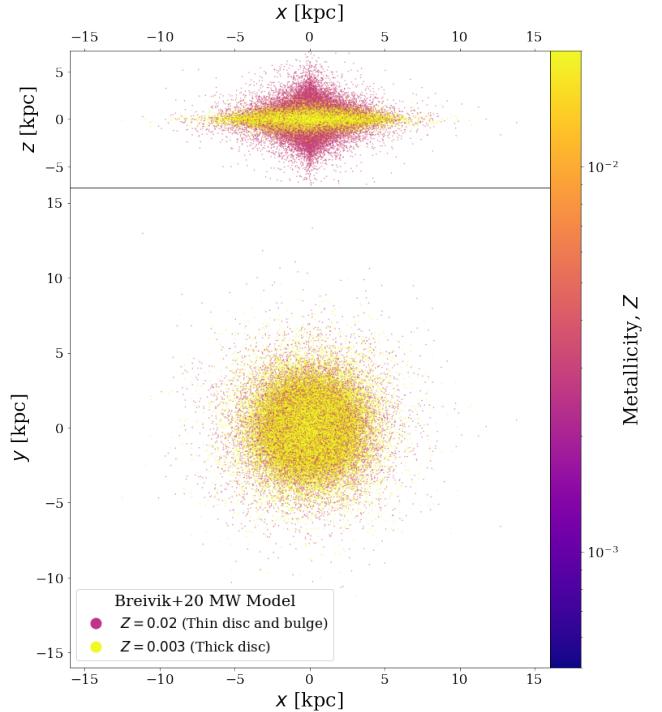
The model that we use for the Milky Way adds several layers of complexity, accounting for the inside-out growth of the thin disc, using empirically informed star formation histories that are a function of time and assigning metallicities based on the position and age of binaries. In this section, we repeat our main analysis but instead apply a simpler model for the Milky Way in order to assess the effect of these added features. For this purpose, we use model for the Milky Way used in Breivik et al. (2020) as this is representative of the models used in most previous works.

Their model can be summarised as follows: the Milky Way is assumed to comprise of three components, a thin disc, a thick disc and a bulge. The spatial distributions and relative masses for these components are given in McMillan (2011). Breivik et al. (2020) assume constant star formation over 10 Gyr for the thin disc, a 1 Gyr burst of star formation 11 Gyr ago for the thick disc and a 1 Gyr burst of star formation 10 Gyr ago for the bulge. A major difference is that only two metallicities are used and they are assigned to binaries independent of age or position. Binaries formed in the thin disc and bulge are assumed to have a metallicity of  $Z = 0.02$  and those formed in the thick disc are assumed to have  $Z = 0.003$ .

We show the spatial metallicity distribution for this model in Fig. 13 in the same form as Fig. 1 for ease of comparison between our models. The two main differences we can see between Fig. 1 and 13 are that the Breivik et al. (2020) model is more centrally concentrated and only has two fixed metallicity populations.

When applying this simpler Milky Way model in combination with our fiducial binary physics assumptions (model A), we find that the expected number of detections for BHBHs, BHNSs and NSNSs for a 4-year LISA mission is 52, 25 and 17 respectively. Thus the BHBH detection has decreased slightly compared to our main findings, whilst for BHNSs and NSNSs the rate has approximately halved and doubled respectively.

Moreover, the distribution of parameters within the population, particularly the mass distributions, are notably disparate. By using only two fixed metallicity populations, unphysical artifacts are introduced into distribution of DCO masses (Kummer et al. (in prep)). For example, in Fig. 14, we show the black hole mass distribution produced by the simulation using the simple Milky Way model. Despite the fact that these KDEs use the same bandwidth as Fig. 3, the distributions show many more sharp transitions, which is a result of pile-ups occurring at specific masses for specific metallicities.



**Figure 13.** As Fig. 1 (right panel), but for the Milky Way model used in Breivik et al. (2020).

Moreover, the lack of lower metallicity systems means that higher mass systems are not formed and so we see the distributions do not include a high mass tail such as in our fiducial results.

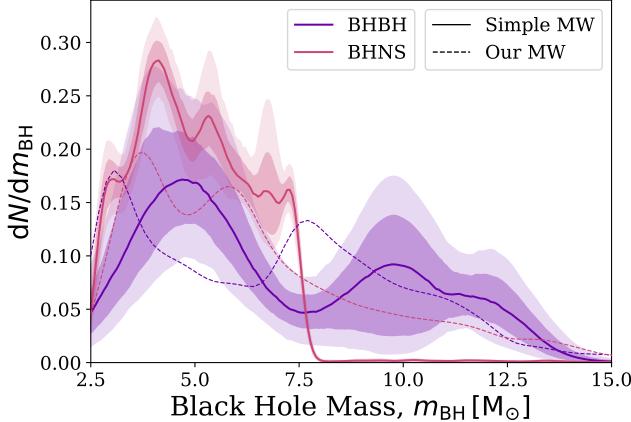
The unphysical artifacts present in the mass distributions can have far-reaching effects since the masses of DCOs affect most other parameters. The inspiral time and SNR are directly dependent on the mass, whilst the uncertainty estimates depend on the SNR. This means that the artifacts can affect the predictions for most distributions of LISA detectable populations.

Overall, we find that previous studies that use Milky Way models analogous to this simpler model may significantly underestimate the LISA BHNS rate whilst overestimating the NSNS detection rate. They may also miss higher mass systems (particular for BHNSs) and contain unphysical artifacts in their parameter distributions.

#### E. ESTIMATING THE NUMBER OF PULSARS FOR A GIVEN SKY AREA IN SKA

In this section, we perform some back-of-the-envelope calculations in order to estimate the number of pulsars that SKA will observe within a given sky area.

First, we consider how many pulsars SKA is likely to detect. Keane et al. (2015) uses PSRPOPPy (Bates et al. 2014) to simulate the Milky Way pulsar population. They find that for SKA-1, approximately 10000



**Figure 14.** As Fig. 3 (top left panel), but for the Milky Way model used in Breivik et al. (2020). Dotted lines show the distribution from Fig. 3 for comparison.

pulsars will be discovered. The second phase of SKA, which should be in operation by the time of the LISA mission, would yield a total of 35000-41000 pulsars (Keane et al. 2015). We use the average, 38000, in further estimates below. Moreover, we are only interested in pulsars that are part of a binary system. We estimate this pulsar binary fraction as the fraction of known pulsars that are in binaries using the ATNF Pulsar Catalogue<sup>9</sup> (Manchester et al. 2005). 290 of the 2872 currently known pulsars are in binary systems and thus we estimate the binary fraction of pulsars as 10%. Therefore, we expect that SKA-1 and SKA-2 will detect approximately 1000 and 3800 binary pulsars respectively.

Next, we can find the total number of pulsars SKA will detect in a patch on the sky. The total sky area that the SKA mission covers is approximately  $5700 \text{ deg}^2$ , which is calculated by integrating over the sky for all Galactic longitudes and Galactic latitudes limited to  $|b| < 10^\circ$  and  $\delta < 45^\circ$ , which are the limits on SKA-mid (Keane et al. 2015). If we assume that the pulsars are found uniformly across the sky, this means that roughly 0.2 and 0.7 binary pulsars are expected per square degree for SKA-1 and SKA-2 respectively. Note that the assumption of a uniform distribution is not realistic as pulsars will tend to be far more concentrated in the Galactic centre but we use it to provide a slightly optimistic estimate.

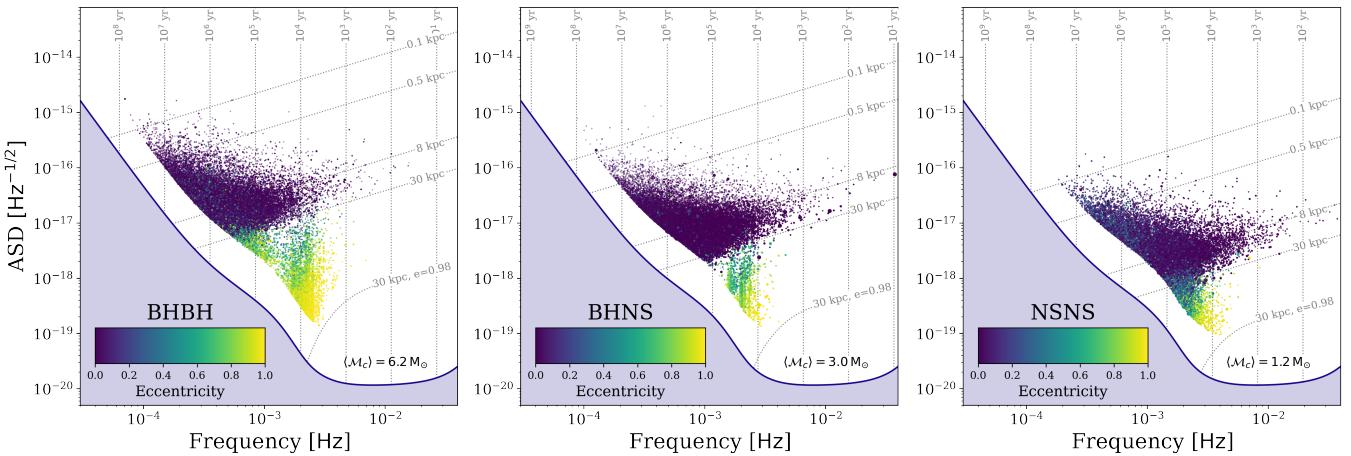
Overall, we therefore expect a single pulsar per  $5.7 \text{ deg}^2$  and  $1.5 \text{ deg}^2$  for SKA-1 and SKA-2 respectively, which correspond to angular resolutions of  $\sigma_\theta = 1.3^\circ$  and  $\sigma_\theta = 0.7^\circ$ .

<sup>9</sup><https://www.atnf.csiro.au/research/pulsar/psrcat>

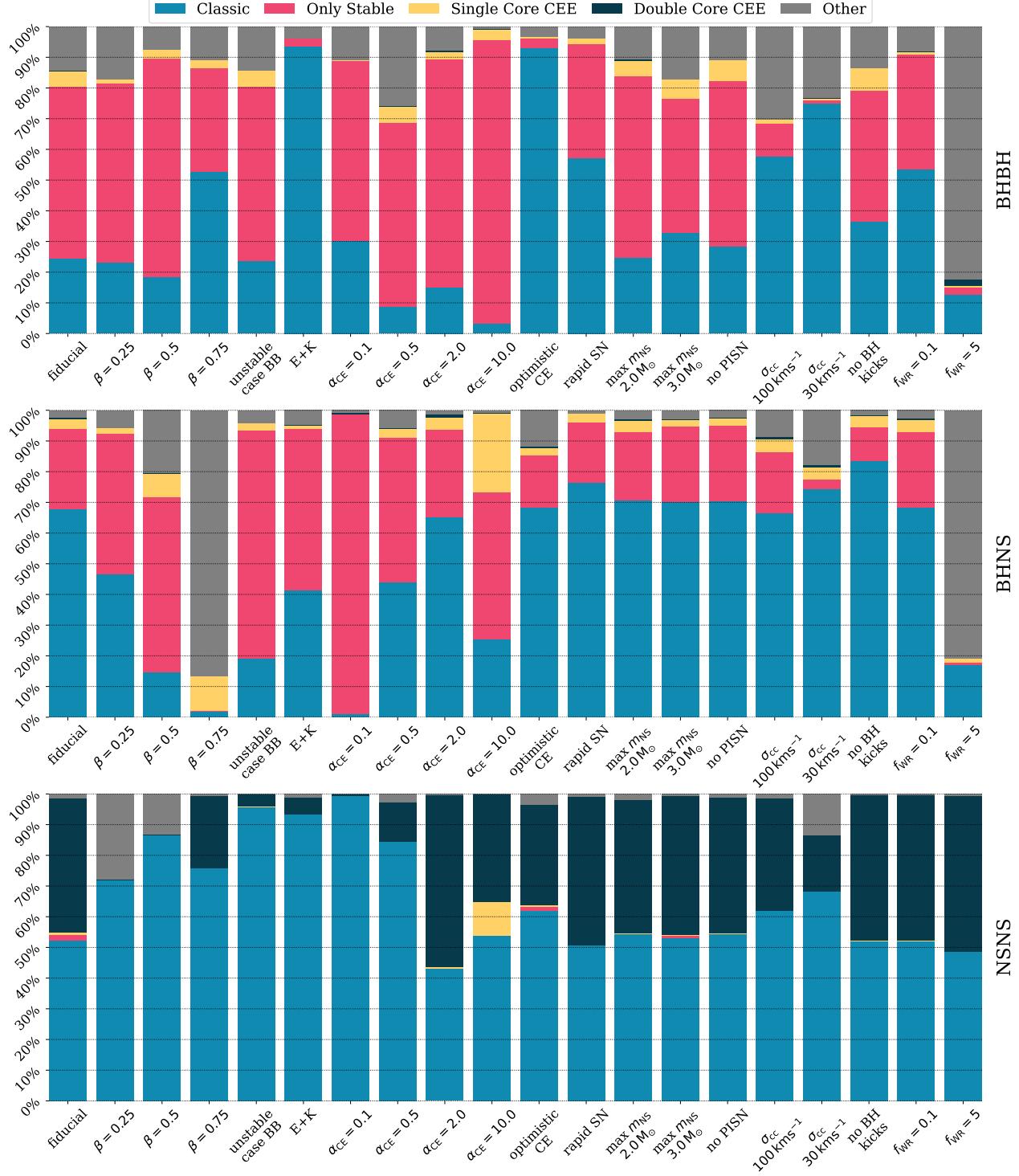
## F. SUPPLEMENTARY MATERIAL

**Table 3.** The number of detectable binaries in a 4- and 10-year LISA mission for the 20 different model variations and each DCO type. Each value shows the mean and the  $1-\sigma$  Poisson uncertainty.  $\square$

Model	Description	LISA detections (4 year)			LISA detections (10 year)		
		BHBH	BHNS	NSNS	BHBH	BHNS	NSNS
A	Fiducial	$74.0^{+9.0}_{-9.0}$	$42.4^{+6.6}_{-6.4}$	$7.9^{+3.1}_{-2.9}$	$117.9^{+11.1}_{-10.9}$	$71.3^{+8.7}_{-8.3}$	$13.0^{+4.0}_{-4.0}$
B	Fixed mass transfer efficiency of $\beta = 0.25$	$68.8^{+8.2}_{-7.8}$	$22.4^{+4.6}_{-4.4}$	$2.9^{+2.1}_{-1.9}$	$107.8^{+10.2}_{-10.8}$	$36.9^{+6.1}_{-5.9}$	$4.6^{+2.4}_{-2.4}$
C	Fixed mass transfer efficiency of $\beta = 0.5$	$47.0^{+7.0}_{-7.0}$	$8.3^{+2.7}_{-3.3}$	$3.9^{+2.1}_{-1.9}$	$75.8^{+9.2}_{-8.8}$	$13.6^{+3.4}_{-3.6}$	$6.4^{+2.6}_{-2.4}$
D	Fixed mass transfer efficiency of $\beta = 0.75$	$46.9^{+7.1}_{-6.9}$	$7.4^{+2.6}_{-2.4}$	$12.7^{+3.3}_{-3.7}$	$71.2^{+8.8}_{-8.2}$	$12.1^{+3.9}_{-3.1}$	$21.1^{+4.9}_{-4.1}$
E	Case BB mass transfer is always unstable	$69.3^{+8.7}_{-8.3}$	$7.3^{+2.7}_{-2.3}$	$0.2^{+0.8}_{-0.2}$	$109.3^{+10.7}_{-10.3}$	$11.8^{+3.2}_{-3.8}$	$0.4^{+0.6}_{-0.4}$
F	Model E + Model K	$154.3^{+12.7}_{-12.3}$	$148.4^{+12.6}_{-12.4}$	$18.2^{+3.8}_{-4.2}$	$239.5^{+15.5}_{-15.5}$	$216.8^{+15.2}_{-14.8}$	$27.4^{+5.6}_{-5.4}$
G	CE efficiency $\alpha = 0.1$	$27.9^{+5.1}_{-4.9}$	$2.1^{+1.9}_{-1.1}$	$10.2^{+2.8}_{-3.2}$	$43.9^{+7.1}_{-6.9}$	$3.5^{+1.5}_{-1.5}$	$16.7^{+4.3}_{-3.7}$
H	CE efficiency $\alpha = 0.5$	$58.3^{+7.7}_{-7.3}$	$21.8^{+4.2}_{-4.8}$	$5.9^{+2.1}_{-2.9}$	$91.6^{+9.4}_{-9.6}$	$34.9^{+6.1}_{-5.9}$	$9.7^{+3.3}_{-2.7}$
I	CE efficiency $\alpha = 2.0$	$67.6^{+8.4}_{-8.6}$	$38.0^{+6.0}_{-6.0}$	$27.7^{+5.3}_{-5.7}$	$109.6^{+10.4}_{-10.6}$	$62.7^{+8.3}_{-7.7}$	$46.0^{+7.0}_{-7.0}$
J	CE efficiency $\alpha = 10.0$	$26.7^{+5.3}_{-4.7}$	$16.3^{+3.7}_{-4.3}$	$34.9^{+6.1}_{-5.9}$	$42.4^{+6.6}_{-6.4}$	$26.6^{+5.4}_{-5.6}$	$57.2^{+7.8}_{-7.2}$
K	HG donor stars initiating a CE survive CE	$151.5^{+12.5}_{-12.5}$	$56.8^{+7.2}_{-7.8}$	$9.9^{+3.1}_{-2.9}$	$229.7^{+15.3}_{-14.7}$	$96.1^{+9.9}_{-10.1}$	$15.8^{+4.2}_{-3.8}$
L	Fryer rapid SN remnant mass prescription	$50.4^{+6.6}_{-7.4}$	$70.2^{+8.8}_{-8.2}$	$6.7^{+2.3}_{-2.7}$	$76.6^{+8.4}_{-8.6}$	$117.4^{+10.6}_{-10.4}$	$10.8^{+3.2}_{-2.8}$
M	Maximum NS mass = $2.0 M_{\odot}$	$96.2^{+9.8}_{-10.2}$	$30.1^{+5.9}_{-5.1}$	$7.2^{+2.8}_{-2.2}$	$153.7^{+12.3}_{-12.7}$	$49.9^{+7.1}_{-6.9}$	$11.6^{+3.4}_{-3.6}$
N	Maximum NS mass = $3.0 M_{\odot}$	$58.3^{+7.7}_{-7.3}$	$51.9^{+7.1}_{-6.9}$	$8.2^{+2.8}_{-3.2}$	$91.6^{+9.4}_{-9.6}$	$84.8^{+9.2}_{-8.8}$	$13.5^{+3.5}_{-3.5}$
O	PISN and pulsational-PISN not implemented	$75.3^{+8.7}_{-8.3}$	$43.4^{+6.6}_{-6.4}$	$8.0^{+3.0}_{-3.0}$	$120.5^{+10.5}_{-10.5}$	$72.3^{+8.7}_{-8.3}$	$12.8^{+3.2}_{-3.8}$
P	$\sigma_{\text{RMS}}^{\text{ID}} = 100 \text{ km s}^{-1}$ for core-collapse supernova	$82.7^{+9.3}_{-8.7}$	$86.6^{+9.4}_{-9.6}$	$15.4^{+3.6}_{-4.4}$	$130.1^{+11.9}_{-11.1}$	$145.1^{+11.9}_{-12.1}$	$25.6^{+5.4}_{-4.6}$
Q	$\sigma_{\text{RMS}}^{\text{ID}} = 30 \text{ km s}^{-1}$ for core-collapse supernova	$91.8^{+9.2}_{-9.8}$	$142.9^{+12.1}_{-11.9}$	$33.6^{+5.4}_{-5.6}$	$142.9^{+12.1}_{-11.9}$	$229.0^{+15.0}_{-15.0}$	$54.9^{+7.1}_{-7.9}$
R	Black holes receive not natal kick	$90.8^{+9.2}_{-9.8}$	$132.1^{+11.9}_{-11.1}$	$7.2^{+2.8}_{-2.2}$	$142.3^{+11.7}_{-12.3}$	$218.6^{+14.4}_{-14.6}$	$11.8^{+3.2}_{-3.8}$
S	Wolf-Rayet wind factor $f_{\text{WR}} = 0.1$	$75.7^{+8.3}_{-8.7}$	$34.0^{+6.0}_{-6.0}$	$8.8^{+3.2}_{-2.8}$	$112.4^{+10.6}_{-10.4}$	$55.8^{+7.2}_{-7.8}$	$14.3^{+3.7}_{-4.3}$
T	Wolf-Rayet wind factor $f_{\text{WR}} = 5.0$	$5.7^{+2.3}_{-2.7}$	$15.5^{+3.5}_{-3.5}$	$8.5^{+2.5}_{-2.5}$	$9.2^{+2.8}_{-3.2}$	$26.2^{+4.8}_{-5.2}$	$13.4^{+3.6}_{-3.4}$



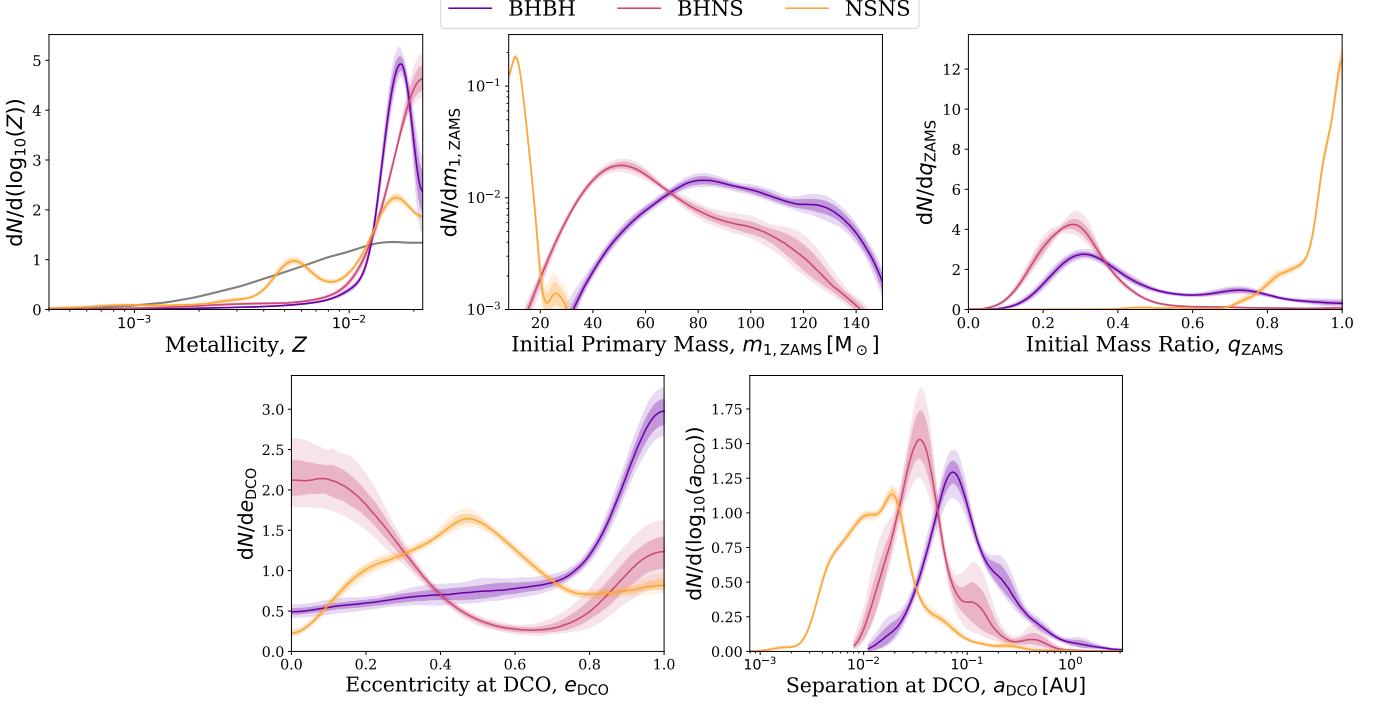
**Figure 15.** As the bottom panels of Fig. 2, but without the density distributions and scatter points are coloured by their eccentricity. We show eccentric sources are located in an offshoot below the 30 kpc around 2 mHz.  $\square$



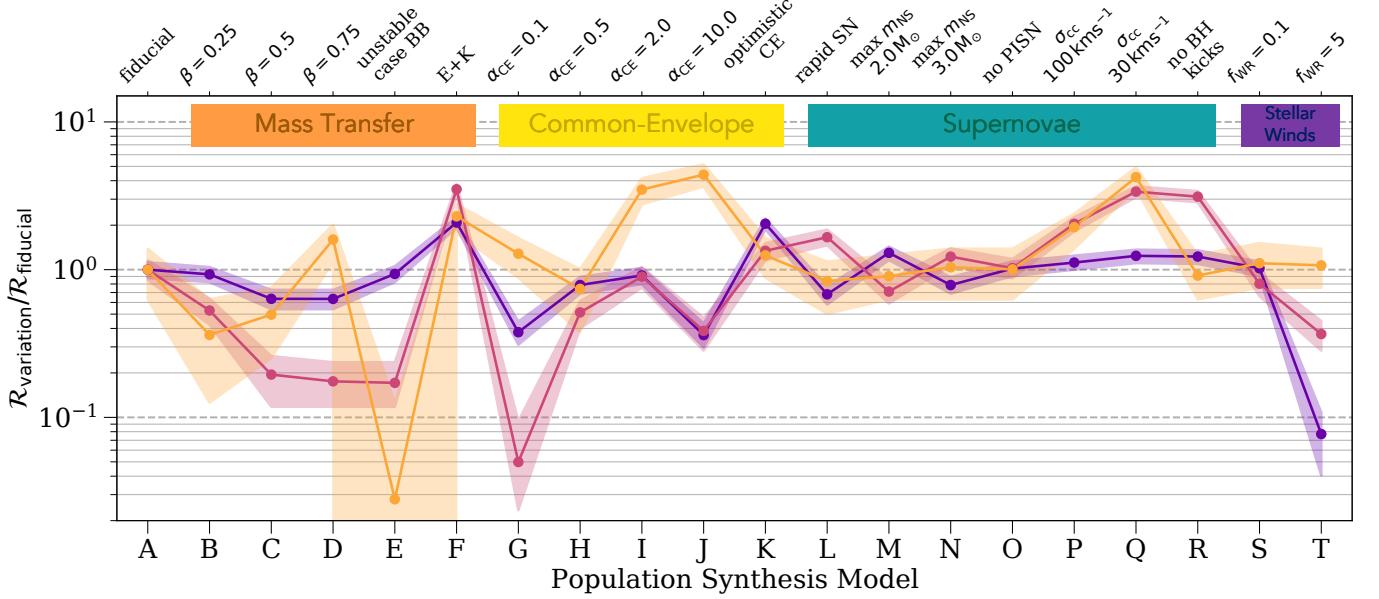
**Figure 16.** Fraction of each DCO type that is formed through different formation channels for all physics variations. Channels are described in detail in Broekgaarden et al. (2021). The classic, single core CEE and double core CEE channels all require at least one common-envelope event whilst only ‘only stable’ consists of only stable mass transfer and ‘other’ contains the remaining binaries which are mainly formed from ‘lucky’ supernova kicks that shrink the binary.

Model	4-year Mission						10-year Mission					
	$\Delta M_c/M_c < 0.1$			$\sigma_\theta < 0.1 \text{ deg}$			$\Delta M_c/M_c < 0.1$			$\sigma_\theta < 0.1 \text{ deg}$		
	BHBH	BHNS	NSNS	BHBH	BHNS	NSNS	BHBH	BHNS	NSNS	BHBH	BHNS	NSNS
A	0.15	0.11	0.68	0.23	0.15	0.65	0.24	0.61	0.12	0.10	0.59	0.18
B	0.15	0.10	0.68	0.22	0.16	0.66	0.23	0.48	0.13	0.10	0.59	0.17
C	0.15	0.10	0.69	0.20	0.14	0.68	0.24	0.49	0.13	0.10	0.59	0.16
D	0.18	0.09	0.64	0.21	0.13	0.65	0.31	0.24	0.44	0.16	0.08	0.54
E	0.15	0.12	0.70	0.14	0.15	0.76	0.22	0.25	0.76	0.12	0.10	0.61
F	0.19	0.09	0.63	0.08	0.22	0.91	0.04	0.35	0.96	0.16	0.07	0.53
G	0.15	0.10	0.68	0.17	0.17	0.75	0.27	0.24	0.55	0.13	0.09	0.59
H	0.14	0.11	0.69	0.19	0.20	0.70	0.32	0.24	0.37	0.12	0.09	0.59
I	0.14	0.10	0.69	0.24	0.16	0.65	0.28	0.24	0.64	0.12	0.09	0.60
J	0.14	0.11	0.70	0.17	0.16	0.75	0.29	0.24	0.49	0.12	0.10	0.60
K	0.19	0.08	0.64	0.25	0.15	0.63	0.25	0.24	0.67	0.16	0.08	0.53
L	0.18	0.08	0.63	0.22	0.15	0.68	0.27	0.24	0.66	0.16	0.08	0.53
M	0.14	0.13	0.72	0.23	0.16	0.64	0.28	0.24	0.59	0.12	0.11	0.62
N	0.16	0.10	0.66	0.22	0.14	0.66	0.28	0.23	0.61	0.14	0.09	0.57
O	0.15	0.11	0.69	0.23	0.15	0.65	0.28	0.24	0.61	0.13	0.10	0.59
P	0.19	0.09	0.64	0.24	0.14	0.64	0.27	0.22	0.61	0.15	0.08	0.53
Q	0.18	0.09	0.64	0.26	0.15	0.60	0.29	0.23	0.57	0.17	0.08	0.54
R	0.15	0.10	0.68	0.24	0.15	0.63	0.27	0.23	0.60	0.12	0.10	0.59
S	0.18	0.08	0.65	0.20	0.13	0.68	0.23	0.22	0.71	0.16	0.08	0.54
T	0.16	0.09	0.70	0.23	0.15	0.62	0.33	0.25	0.43	0.13	0.08	0.60

Table 4. A table



**Figure 17.** As Fig. 3, but for the properties of the progenitors of detectable systems (top row) and the properties of the detectable systems at DCO formation (bottom row). Note that the  $m_{1,\text{ZAMS}}$  panel uses a log scaled  $y$ -axis. .



**Figure 18.** Similar to Fig. 8, this shows the *relative* detection rates compared to the fiducial model in order to emphasise how much each DCO is affected by model variations compared to other DCOs. The scatter points show the mean and the shaded areas show the 1- $\sigma$  uncertainty range. .