



## UvA-DARE (Digital Academic Repository)

### Combined analysis of neutron star natal kicks using proper motions and parallax measurements for radio pulsars and Be X-ray binaries

Igoshev, A.P.; Chruslinska, M.; Dorozsmai, A.; Toonen, S.

**DOI**

[10.1093/mnras/stab2734](https://doi.org/10.1093/mnras/stab2734)

**Publication date**

2021

**Document Version**

Final published version

**Published in**

Monthly Notices of the Royal Astronomical Society

[Link to publication](#)

**Citation for published version (APA):**

Igoshev, A. P., Chruslinska, M., Dorozsmai, A., & Toonen, S. (2021). Combined analysis of neutron star natal kicks using proper motions and parallax measurements for radio pulsars and Be X-ray binaries. *Monthly Notices of the Royal Astronomical Society*, 508(3), 3345-3364. <https://doi.org/10.1093/mnras/stab2734>

**General rights**

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

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

*UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)*

# Combined analysis of neutron star natal kicks using proper motions and parallax measurements for radio pulsars and Be X-ray binaries

Andrei P. Igoshev<sup>1</sup>,<sup>\*</sup> Martyna Chruslinska<sup>2</sup>, Andris Dorozsmai<sup>3</sup> and Silvia Toonen<sup>3,4</sup>

<sup>1</sup>*Department of Applied Mathematics, University of Leeds, Leeds LS2 9JT, UK*

<sup>2</sup>*Institute of Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen, PO Box 9010, NL-6500 GL Nijmegen, the Netherlands*

<sup>3</sup>*Institute of Gravitational Wave Astronomy, School of Physics and Astronomy, University of Birmingham, Birmingham B15 2TT, UK*

<sup>4</sup>*Anton Pannekoek Institute for Astronomy, University of Amsterdam, NL-1090 GE Amsterdam, the Netherlands*

Accepted 2021 September 20. Received 2021 September 17; in original form 2021 August 15

## ABSTRACT

Supernova explosion and the associated neutron star (NS) natal kicks are important events on a pathway of a binary to become a gravitational wave source, an X-ray binary, or a millisecond radio pulsar. Weak natal kicks often lead to binary survival, while strong kicks frequently disrupt the binary. In this article, we aim to further constrain NS natal kicks in binaries. We explore binary population synthesis models by varying prescription for natal kick, remnant mass, and mass accretion efficiency. We introduce a robust statistical technique to analyse combined observations of different nature. Using this technique, we further test different models using parallax and proper motion measurements for young isolated radio pulsars and similar measurements for Galactic Be X-ray binaries (BeXs). Our best model for natal kicks is consistent with both measurements and contains a fraction of  $w = 0.2 \pm 0.1$  weak natal kicks with  $\sigma_1 = 45^{+25}_{-15} \text{ km s}^{-1}$ , the remaining natal kicks are drawn from the high-velocity component, same as in previous works:  $\sigma_2 = 336 \text{ km s}^{-1}$ . We found that currently used models for natal kicks of NSs produced by electron capture supernova (ecSN; combination of Maxwellian  $\sigma = 265 \text{ km s}^{-1}$  and  $\sigma = 30 \text{ km s}^{-1}$  for electron capture) are inconsistent or marginally consistent with parallaxes and proper motions measured for isolated radio pulsars. We suggest a new model for natal kicks of ecSN, which satisfy both observations of isolated radio pulsars and BeXs.

**Key words:** methods: statistical – binaries: general – stars: massive – stars: neutron – pulsars: general – X-rays: binaries.

## 1 INTRODUCTION

Neutron stars (NSs) often receive a significant birth velocity (natal kick) at the moment of supernova explosion (Lyne & Lorimer 1994). On one hand, this kick is frequently strong with amplitude  $|v| > 100 \text{ km s}^{-1}$  (in some cases reaching  $800 \text{ km s}^{-1}$ ) while typical velocity dispersion of NS progenitors is  $< 20 \text{ km s}^{-1}$ . On the other hand, formation of certain Galactic double neutron stars (DNSs) systems and presence of pulsars in globular clusters (with escape velocities  $\lesssim 50 \text{ km s}^{-1}$ ) clearly require small natal kicks (Tauris et al. 2017). It is unclear whether the NS natal kick velocity distribution is broad and extends from  $\approx 0$  to  $1000 \text{ km s}^{-1}$ , or if NSs with some particular characteristics are formed with much weaker natal kicks ( $0\text{--}30 \text{ km s}^{-1}$ ).

The problem of estimating the NS natal kick distribution has two parts, which were not sufficiently addressed in the previous research on the subject: (1) NSs receiving weak natal kick should be less common among isolated radio pulsars because they less frequently disrupt the binary and (2) different formation channels could lead to different velocity distributions. In studies of observed velocity distribution of radio pulsars (e.g. Hobbs et al. 2005; Verbunt, Igoshev & Cator 2017; Igoshev 2020), young pulsar velocities were assumed

to be equal to the isolated NS natal kick velocities and the influence of binarity of the NS progenitors was never analysed.

Individual velocities of radio pulsars (mode of activity typical for young, magnetized NSs with fast rotation; see e.g. Deller et al. 2019) are estimated based on VLBI measurements of parallaxes and proper motions. Individual velocities for high-mass X-ray binaries (HMXBs) and DNSs are estimated based on analysis of orbital properties (see e.g. Dewi, Podsiadlowski & Pols 2005; Beniamini & Piran 2016; Tauris et al. 2017).

On the one hand, young radio pulsars with reliably measured parallaxes and proper motions seem to be isolated (Igoshev & Perets 2019). Also, recent analysis of all radio pulsars recorded in the ATNF pulsar catalogue<sup>1</sup> (Manchester et al. 2005) found an upper limit of 5.3 per cent on the fraction of radio pulsars with wide binary companions (Antoniadis 2020, 2021). Close binary companions with orbital periods shorter than  $\sim 10 \text{ yr}$  (time-scales comparable with observation spans for radio pulsars) are easy to detect using the pulsar timing technique (Edwards, Hobbs & Manchester 2006). On the other hand, progenitors of NSs are massive ( $\geq 7 M_{\odot}$ ) and the majority of such stars are formed in binaries. A fraction of truly isolated massive stars is less than 15 per cent (Moe & Di Stefano 2017). This large discrepancy between the fraction of massive stars in binaries and the fraction of radio pulsars in binaries can be explained by a supernova

\* E-mail: [ignotur@gmail.com](mailto:ignotur@gmail.com)

<sup>1</sup> ATNF pulsar catalogue <http://www.atnf.csiro.au/research/pulsar/psrcat>.

explosion, which often disrupt a binary due to mass-loss and NS natal kick (see e.g. Renzo et al. 2019).

A stronger natal kick is more likely to disrupt a binary and leave behind an isolated radio pulsar, while an NS formed with natal kick weaker or comparable to orbital velocity of the secondary (typically less than  $100 \text{ km s}^{-1}$ ) is more likely to remain bound within the binary. Therefore, a universal natal kick distribution (if it exists) should be probed simultaneously using binaries with NS and isolated radio pulsars. A particular class of systems that can be used in such an analysis, and is explored in this study, is the population of Be X-ray binaries (BeXs). Those are HMXBs hosting NS, whose formation path involves a phase of (stable) mass transfer before the supernova explosion (e.g. Vinciguerra et al. 2020).

In this article, we aim to model the observed velocities of isolated radio pulsars and the peculiar velocities of Galactic BeXs using the binary population synthesis. Our approach is an extension to one suggested by Andrews et al. (2015). First, we identify the model that best describes the population of BeXs and then we vary the parameters of the natal kick distribution only. To allow for a qualitative comparison of different models, we introduce a new likelihood method for parallaxes and proper motion measurements, which allows us to directly use the results of binary stellar population synthesis as a model for peculiar velocities.

While we were preparing this article, Willcox et al. (2021) published a study on a similar topic. The author used a non-parametric description for the natal kicks of core-collapse NSs. They found that electron capture supernova (ecSN) events should only occur in stars stripped of hydrogen envelope to satisfy observed transverse velocities of isolated radio pulsars. This research does not take into account a comparison with BeXs and leaves us the opportunity to investigate low natal kicks.

This article is structured as follows: in Section 2, we give a brief overview for BeXs, ecSN, and modelling efforts; in Section 3, we introduce observational data sets for young isolated radio pulsars and BeXs, and perform simple analysis of peculiar velocities for BeXs. In Section 4, we describe our binary population synthesis. In Section 5, we present results for these exploratory models. In Section 6, we introduce our combined analysis and discuss ecSN explosion in details, which is followed by a discussion and conclusion in Sections 7 and 8, respectively.

## 2 BE X-RAY BINARIES AND ELECTRON CAPTURE SUPERNOVA EXPLOSIONS

Be stars are B spectral type stars with enhanced emission lines (see e.g. review by Reig 2011). Their emission spectral lines are formed in decretion disc. The secondary star rotates faster than 70 per cent of its critical spinning rate (Slettebak 1982; Porter 1996; Yudin 2001; Townsend, Owocki & Howarth 2004) and loses material from equatorial regions forming a large decretion disc. When combined with an NS in a binary these stars produced the BeX phenomenon. NS orbiting around the Be star enters the decretion disc close to the periastron position and accretes material. The accretion causes a strong heating of the polar regions, which is seen in X-rays.

The decretion disc around Be is thought to be formed due to fast rotation of B stars. Kriz & Harmanec (1975) and Rappaport & van den Heuvel (1982) suggested that the star is spun up due to accretion of large amount of material with angular momentum. It is important to note that masses of the isolated Be stars and Be X-ray stars differ. The most common masses of isolated Be stars are around  $3 M_{\odot}$  while Be stars in Be X-ray systems are more massive with masses of  $8 M_{\odot}$ . Vinciguerra et al. (2020) suggested that the

mass distribution of Be stars in binaries could be a good indicator for the mass accretion efficiency. A large sample of BeXs is observed in the Small Magellanic Cloud (SMC) where researchers discovered  $\approx 70$  active systems.

Prišegen (2020) recently analysed proper motions of BeXs in the Galaxy using the *Gaia* DR2. He found that transverse systemic velocities of BeXs are  $29 \pm 11 \text{ km s}^{-1}$  for short-period binaries and  $16 \pm 8 \text{ km s}^{-1}$  for long-period binaries. He confirmed that this difference is statistically significant. In another recent analysis by Bodaghee et al. (2021), researchers used cross-correlation function between position of HMXBs in SMC and OB associations. They found that initial systemic velocities (called kick velocities) are in the range  $2\text{--}34 \text{ km s}^{-1}$ . In this work, we plan to analyse the parallaxes and proper motions of BeXs using the new *Gaia* EDR3.

Over the years, it was suggested that NS born in binaries could receive very different natal kicks due to supernova explosion mechanisms that prevail or only operate in binaries, such as: (1) the ecSN explosion (Nomoto 1984, 1987) and (2) the explosion of ultrastripped helium stars (e.g. Tauris et al. 2013). The latter mechanism relies on a reduction of ejecta mass (Janka 2017), while the former mechanism leads to weaker natal kicks because stellar cores of ecSN progenitors are easier to explode and less asymmetries (responsible for formation of the natal kick) are developed (Podsiadlowski et al. 2004). All of these mechanisms are related to mass transfer in close interacting binaries, which constitute up to 70 per cent of all massive binaries (Sana et al. 2012). Once again, NSs formed through these mechanisms will more frequently stay in a binary due to reduced kick and mass-loss.

ecSN progenitors are believed to originate from stars that are too lightweight to ignite oxygen in their cores (stars that ignite oxygen eventually undergo core-collapse supernova explosion (CCSN)) and too heavy to end up as oxygen–neon white dwarfs. Such stars develop degenerate ONeMg cores after the carbon burning phase that exceed the Chandrasekhar limit  $\sim 1.37 M_{\odot}$  and they can collapse to NS in the process triggered by electron captures on magnesium and neon (e.g. Miyaji et al. 1980; Nomoto 1984; Nomoto & Kondo 1991; Woosley & Heger 2015; Jones et al. 2016). The range of initial masses allowed for ecSN progenitors is expected to be narrow for single stars, but can be effectively broadened by binary interactions, allowing for massive degenerate ONeMg core formation from initially more/less massive progenitors (e.g. Podsiadlowski et al. 2004; Doherty et al. 2017; Poelarends et al. 2017; Siess & Lebreuilly 2018). EcSNe are typically implemented in binary population synthesis codes based on the helium core mass of a star at the beginning of the AGB stage  $M_{C, \text{BAGB}}$  (with the commonly adopted ecSN progenitor  $M_{C, \text{BAGB}}$  range  $[1.8, 2.25] M_{\odot}$ , following Hurley, Pols & Tout 2000; Eldridge & Tout 2004 – as also used in this study). It should be noted that the exact limits on this mass range are very sensitive to the adopted assumptions about the stellar and binary physics (e.g. the treatment of convection and overshooting/mixing in stellar interiors, wind mass-loss rates, treatment of mass transfer, e.g. Poelarends et al. 2008; Jones et al. 2016; Doherty et al. 2017; Siess & Lebreuilly 2018) and depend on metallicity (e.g. Poelarends et al. 2008, but see Doherty et al. 2017 – where the metallicity dependence is shown to be very limited). Furthermore, a fraction of the ecSN progenitors estimated that way may not collapse to NS, but undergo a partial thermonuclear explosion and instead leave behind a white dwarf (e.g. Jones et al. 2016).

Overall, the evolution of the binary after supernova explosion was modelled multiple times in the literature (see e.g. Hills 1983; Dewey & Cordes 1987; Brandt & Podsiadlowski 1995; Kalogera 1996; Fryer, Burrows & Benz 1998). Multiple studies were focused

on the influence of NS natal kicks on the rate of compact binary mergers, especially in relation to gravitational wave progenitors (e.g. Lipunov, Postnov & Prokhorov 1997; Tauris et al. 2017). More recently, Kuranov, Popov & Postnov (2009) performed binary population synthesis assuming that ecSNe are only produced in binaries. They also assumed that NSs born in such explosions receive a significantly smaller natal kick than seen in truly isolated NS population. They demonstrated that a contribution of ecSN, which disrupted their parent binary, is negligible in comparison to the amount of isolated and binary origin NS with large velocities. The result of their population synthesis showed that the final velocity distribution of isolated NSs followed closely the distribution of natal kicks in isolated NSs besides a small region around  $30 \text{ km}^{-1}$ . In this region, they saw a slight increase (amplitude of a few percents) in the number of NSs. Such a difference is impossible to notice when sample size of parallaxes and proper motions of isolated radio pulsars is limited by 20–100 measurements.

### 3 DATA

In this article, we study parallaxes and proper motion measurements. For young (spin-down ages less than 3 Myr) isolated radio pulsars, we use the same catalogue as Igoshev (2020; see their table 1 and table A1 for details), which includes VLBI measurements by Brisken et al. (2002, 2003), Chatterjee et al. (2001, 2004, 2009), Kirsten et al. (2015), and Deller et al. (2009, 2019). For BeXs we use *Gaia* EDR3 (Gaia Collaboration 2016, 2021; Lindegren et al. 2021). We describe in detail our analysis of BeXs. The result of similar analysis for isolated radio pulsars is available in the article by Igoshev (2020).

#### 3.1 Gaia EDR3 parallaxes and proper motions of Galactic Be X-ray binaries

We use the *Gaia* EDR3 catalogue to identify BeXs. The identification procedure is similar to Igoshev & Perets (2019). Namely, we use the catalogue of HMXBs by Liu, van Paradijs & van den Heuvel (2006) and check the type of these systems in the Simbad database (Wenger et al. 2000). To proceed with identification, we request all stars from the *Gaia* EDR3 around the position of each HMXB in a radius of 15 arcsec. We select from these stars only ones with *Gaia*  $G$  magnitude in the range  $|G - V| < 1.5$  ( $V$  is visual magnitude available in Simbad) located at the smallest angular separation from the catalogue coordinate. We notice that the astrometry analysis procedure improved significantly since *Gaia* DR2 (Lindegren et al. 2018, 2021), and precision checks related to `astrometric_chi2_al` and `astrometric_n_good_obs_al` are already implemented. Therefore, we select only stars with five- and six-parameter astrometric solutions with values of renormalized unit weight error close to 1, namely  $0.6 < \text{ruwe} < 1.4$  (Lindegren et al. 2021). We additionally filter out all measurements with formal errors of parallax  $\varpi'/\sigma_\varpi < 3$ . After this procedure we end up with 45 systems. We summarize the BeXs identified in *Gaia* EDR3 in Table A1.

At the next stage, we inspect the catalogue and remove all systems with unknown spectral type or spectral type different from Be or Oe. In particular, we remove multiple wind-fed supergiant systems because the formation of these systems does not require an angular momentum transfer from primary to the secondary star, therefore these systems could have been formed via the dynamically unstable mass transfer or without any mass transfer at all. In Appendix A, we list some systems where we perform an additional literature search

to decide if they are to be included in the velocity analysis or not. Our final list contains 31 objects.

We plot location of these BeXs projected to the Galactic plane in Fig. 1. We also plot the cumulative distribution of nominal velocities in same Figure, where the nominal velocity in a particular direction (e.g. right ascension) is computed as

$$v'_\alpha = c \frac{\mu_\alpha}{\varpi'}, \quad (1)$$

where  $c = 4.74$  is a numerical coefficient to transform proper motion measurements  $\mu_\alpha$  [ $\text{mas yr}^{-1}$ ] and parallax measurements  $\varpi$  [ $\text{mas}$ ] to units [ $\text{km s}^{-1}$ ]. The transverse nominal velocities of BeXs are computed as the following:

$$v' = \sqrt{\left(c \frac{\mu_l}{\varpi'} - \Delta v_l(l, b, 1/\varpi')\right)^2 + \left(c \frac{\mu_b}{\varpi'} - \Delta v_b(l, b, 1/\varpi')\right)^2}. \quad (2)$$

Here,  $\Delta v_l(l, b, D)$  and  $\Delta v_b(l, b, D)$  are the corrections for the rotation of the Galaxy at location with Galactic coordinates  $l, b$  and distance  $D$  from the Earth. Exact equations for these corrections can be found in appendix B by Verbunt et al. (2017). It is easy to see in Fig. 1 (right-hand panel) that these transverse velocities are limited by  $80 \text{ km s}^{-1}$  and half of the velocities are below  $\approx 20 \text{ km s}^{-1}$ . Typical velocity dispersion inside regions where OB stars formed are below  $10 \text{ km s}^{-1}$  (de Bruijne 1999; Kiminki & Smith 2018). These star formation regions could have a velocity dispersion of  $\approx 5 \text{ km s}^{-1}$  in respect to regular Galactic rotation (Reid et al. 2014). We also analyse velocities of spectroscopic binaries with primary of B type (possible progenitors of BeXs). All details are summarized in Appendix E. We find that these stars have velocities, which can be described as a sum of two Maxwellians with  $w = 0.2$ ,  $\sigma = 3 \text{ km s}^{-1}$ , and  $\sigma_2 = 11 \text{ km s}^{-1}$  (see black dashed line in the right-hand panel of Fig. 1). Therefore, some BeXs have velocities that are clearly above the ones typical for their binary progenitors.

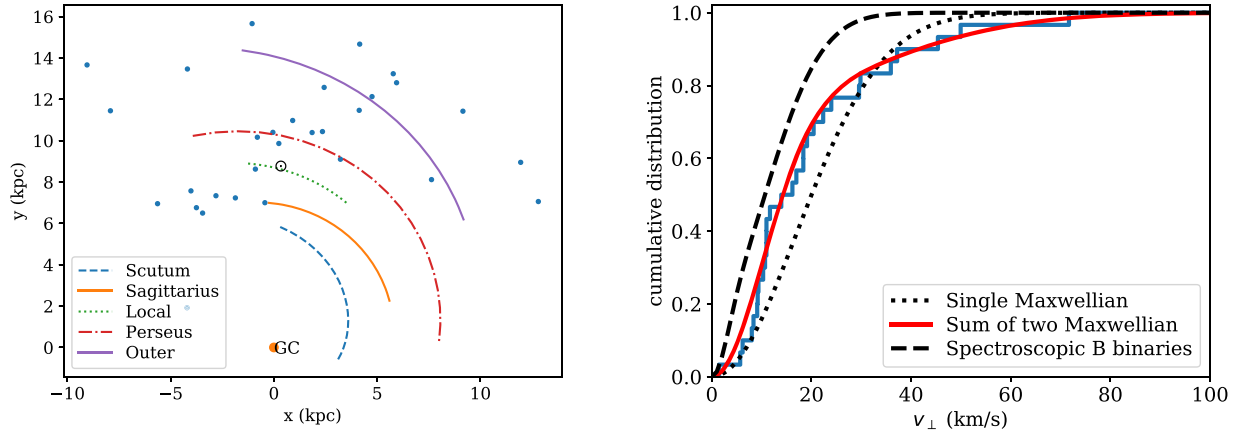
#### 3.2 Peculiar velocities for Be X-ray binaries

We analyse the parallaxes and proper motions of BeXs from Table A1 using the same maximum likelihood technique as was developed and described in detail by Verbunt & Cator (2017) and Verbunt et al. (2017). Brief summary of the method is available in Appendix B. The results of this analysis are summarized in Table 1, and we plot these velocity distributions in Fig. 1 (right-hand panel). A visual inspection confirms that the sum of two Maxwellians describes the peculiar velocities of BeXs significantly better than a single Maxwellian. The values for  $\sigma_1$  and  $\sigma_2$  which we found in our analysis are similar to one by Prišegren (2020).

For completeness, we perform the same analysis using newer measurements for solar rotation speed  $v_\odot = 233 \text{ km s}^{-1}$  and distance towards the Galactic Centre  $R_G = 8.122 \text{ kpc}$  (see e.g. Mróz et al. 2019). Results are also available in Table 1. Different values for Galactic rotation parameters could change the distribution for peculiar velocities of BeXs by a couple of  $\text{km s}^{-1}$ . These new values are within the confidence interval of values estimated using the old  $v_\odot$  and  $R_G$ .

Binary surviving the supernova explosion receives a peculiar velocity comparable to orbital speed of the exploded star. Therefore, the peculiar velocity of BeX is a complicated function of mass transfer, natal kick strength, and orientation, as well as orbital parameters at the moment of supernova explosion. Therefore, it is necessary to use a realistic binary population synthesis to meaningfully analyse Be X-ray peculiar velocities.

Renzo et al. (2019) computed peculiar velocities of binaries with NS. The authors did not model BeXs specifically, but looked at the



**Figure 1.** The distribution of BeXs identified in *Gaia* EDR3 in the Galaxy (left-hand panel) and distribution of their binary B spectroscopic stars' nominal peculiar velocities corrected for the Galactic rotation (right-hand panel). The parameters of spiral arms are estimated by Reid et al. (2014) and confirmed using the *Gaia* data (Xu et al. 2018).

**Table 1.** The parameters of the velocity distribution for BeXs in the Milky Way.

Fixed parameters	Model	Parameters	$2\Delta\mathcal{L}$
$v_{\odot} = 220 \text{ km s}^{-1}$ $R_{\odot} = 8.5 \text{ kpc}$	Single Maxwellian	$\sigma = 17 \pm 3 \text{ km s}^{-1}$	–
	Sum of two Maxwellians	$w = 0.74 \pm 0.15$ $\sigma_1 = 10 \pm 3 \text{ km s}^{-1}$ $\sigma_2 = 29 \pm 12 \text{ km s}^{-1}$	15.2
$v_{\odot} = 233 \text{ km s}^{-1}$ $R_{\odot} = 8.122 \text{ kpc}$	Single Maxwellian	$\sigma = 16 \pm 3 \text{ km s}^{-1}$	–
	Sum of two Maxwellians	$w = 0.82 \pm 0.13$ $\sigma_1 = 11 \pm 2 \text{ km s}^{-1}$ $\sigma_2 = 31 \pm 15 \text{ km s}^{-1}$	10.6

velocity distribution of all main-sequence NS binaries including ones which will never be seen as BeX. Therefore, a direct comparison of results by Renzo et al. (2019) with *Gaia* parallaxes and proper motions of BeXs is difficult. Thus, we perform our own binary population synthesis. Details of this synthesis are described in the following section.

## 4 MODEL EXPLORATION

### 4.1 Formation paths for Be X-ray binaries and isolated radio pulsars

In this article, we concentrate on isolated radio pulsar and BeXs formed via binary interactions in Milky Way (MW or Galaxy). We also examine BeXs population in SMC to allow for comparison with work by Vinciguerra et al. (2020). We do not consider objects formed via dynamical formation channel in crowded environments such as globular clusters. The main advantages of restricting analysis to isolated radio pulsars and BeXs are that formation of these binaries could only weakly depend on poorly understood common envelope stage (CE; Ivanova et al. 2013)<sup>2</sup> and stars in these systems are not expected to undergo ultrastripped SN (which could lead to weaker natal kicks).

<sup>2</sup>A small fraction of isolated radio pulsars could originate from massive stars which is a result of a merger after common envelope stage. However, that is not the dominant channel and we disregard such cases.

#### 4.1.1 Isolated radio pulsars

A significant fraction (up to 85 per cent) of NS progenitors are born in binaries or stellar multiples (Kobulnicky & Fryer 2007; Sana et al. 2012; Moe & Di Stefano 2017). A strong natal kick and fast mass-loss during the SN explosion often leads to binary disruption. A fraction of  $86_{-9}^{+11}$  per cent of NS born in binaries may end up as isolated free-floating NSs (Renzo et al. 2019) for natal kicks drawn from Maxwellian distribution with  $\sigma = 265 \text{ km s}^{-1}$  (Hobbs et al. 2005) and treatment of binary disruption by Tauris & Takens (1998). Thus, a high fraction of the observed isolated radio pulsars (Antoniadis 2021) does not contradict their expected binary origin.

NS could operate as an isolated radio pulsar up to Gyr depending on their initial rotational period, magnetic field strength, and evolution history. In practice, most NSs fade or stop emitting radio after 100 Myr (see e.g. Faucher-Giguère & Kaspi 2006). The radio emission produced by pulsars is strongly beamed, so we miss up to 90 per cent of pulsar population simply due to geometry (Tauris & Manchester 1998).

In our simulations, we assume that every NS is born as a radio pulsar which operates for 100 Myr. We do not integrate their motion in the Galactic gravitational potential and we do not take beaming and sensitivity of modern pulsars survey into account. Those factors become much more important when the whole population of radio pulsars is considered, in particular old pulsars, which could travel a significant distance within the Galaxy. Therefore, our synthetic pulsars can only be compared with young radio pulsars (spin-down ages  $\tau < 3 \text{ Myr}$ ). The velocities of our synthetic pulsars stay representative of the natal kicks, but get some correction due to the binary origin. A primary effect of binary origin is that low natal kick NSs stay more frequently bound in a binary and therefore are excluded from isolated radio pulsar population. There is no indication that beaming or radio luminosity correlates with natal kick velocity or orientation, so these factors are probably irrelevant in natal kick studies.

However, there is a bias that could affect the result of our research and is poorly understood at the moment. Pulsars with well-measured parallaxes and proper motions typically have high Galactic latitudes (Deller et al. 2019). The low natal kick pulsars should be concentrated towards the galactic plane and might be less represented in VLBI measurements. Future pulsar population synthesis is necessary to quantify the role of this effect on analysis of NS natal kicks.

	$M_1$	$M_2$	$t$ (Myr)	$P$
MSMS	17.59	10.8	0.0	32.9 ds
HGMS	17.17	10.79	10.0	7.8 ds
THRM	HGMS 17.17	10.79	10.0	7.8 ds
HdMS	4.55	17.09	10.0	11.9 ds
HgMS	3.88	17.06	12.0	12.7 ds
NSMS	1.48	17.06	12.0	110.8 ds
NSHG	1.48	16.88	18.0	56.9 ds
CE	NSHG 1.48	16.87	18.0	32.2 ds

**Figure 2.** The formation channel for Be X stars for mass accretion efficiency of 0.5 accreted mass.  $M_1$  and  $M_2$  are masses of primary and secondary star in solar masses, respectively,  $t$  is the age of the system, and  $P$  is the orbital period. The size of each circle indicates stellar radius. THRM stands for mass transfer on thermal time-scale and CE stands for common envelope. MSMS stands for two main-sequence stars, HG is Hertzsprung gap, Hd is a helium star, Hg is a helium giant, and NS is neutron star.

#### 4.1.2 Be X-ray binaries and binary stellar evolution before the NS formation

In this section, we first briefly discuss the most common evolutionary channels described in the literature (and found in our SEBA simulations) and then concentrate on the one producing BeXs.

If the binary is wide enough, its components never interact and evolve as single stars. In such cases, the secondary star is not spun up and does not form a decretion disc. Therefore, the NS has no material to accrete and the system does not appear as a BeX. Such systems could still be visible at later stages of evolution as supergiant wind-fed X-ray binaries. We further refer to this type of evolution as effectively single binary evolution channel.

If stars are born sufficiently close, the primary fills its Roche lobe and initiates a mass transfer at some point of its evolution. This could result both in stable and unstable mass transfer. Dynamically unstable mass transfer can lead to a merger, resulting in formation of a solitary rejuvenated star which could explode as SN and form isolated NS (see discussion of possible caveats in Section 7.5). Stable mass transfer at this stage is a necessary condition for the formation of a BeX (e.g. Portegies Zwart 1995; Shao & Li 2014; Vinciguerra et al. 2020). These previous studies found that the most common formation channel involves an episode of non-conservative, dynamically stable mass transfer initiated after the primary star depleted its core hydrogen. This channel is illustrated in Fig. 2. The post-mass transfer binary may get disrupted during the subsequent NS formation. If the binary is unbound due to the natal kick, we consider the NS as an isolated radio pulsar.

If instead of stable mass transfer the binary goes through a CE stage two things could happen: (1) CE is usually very short (see e.g. Igoshev, Perets & Michaely 2020) and only a small fraction of mass is transferred to the secondary, so a transfer of angular momentum is limited and no decretion disc is formed further in the evolution. (2) If the CE is initiated after one or multiple mass transfers, the CE could destroy the decretion disc.

In the interacting formation channel, the mass transfer is non-conservative, and the mass accretion efficiency controls how massive the secondary star (Be star) will be (Vinciguerra et al. 2020). A highly non-conservative mass transfer leads to the formation of BeXs with

Be star masses 3–5  $M_{\odot}$ . On the other hand, if 50–75 per cent of transferred mass is successfully accreted by the secondary, the peak of the Be star mass distribution shifts towards 8–10  $M_{\odot}$ , as it is expected from observations.

#### 4.2 Assumptions for the binary population synthesis

We briefly summarize the details of 14 evolutionary models which we consider in this paper in Table 2. We use the binary population synthesis code SEBA<sup>3</sup> (Portegies Zwart & Verbunt 1996; up-to-date description is available, Toonen, Nelemans & Portegies Zwart 2012). For our exploratory runs, we model the evolution of 100 000 zero-age main-sequence (ZAMS) binaries with primaries mass ranging from 6 to 40  $M_{\odot}$ . The primary masses are sampled from a power-law distribution with an exponent of  $-2.35$  (Salpeter 1955). We restrict the mass of the secondary to the range  $q \in [0.001, 1]$  where  $q = M_2/M_1$ . We draw mass ratio from the uniform distribution (Duchêne & Kraus 2013). For the stellar evolution simulations in the Galaxy, we use the solar metallicity 0.02; for the SMC, we use a metallicity of  $3.5 \times 10^{-3}$ . The semimajor axis distribution is chosen in the form of power law with exponent of  $-1$  ranging from Roche lobe radius to  $10^6 R_{\odot}$  (Abt 1983). The eccentricity distribution is thermal (Ambartsumian 1937; Heggie 1975). We assume 100 per cent binary fraction (see Section 7.1 for test of a smaller binary fraction).

We consider two mass accretion efficiency schemes: following the default SEBA implementation and a fixed semiconservative case. The default mass accretion efficiency implemented in SEBA is based on Pols & Marinus (1994). The accretion rate is limited by

$$\dot{m}_{\max} = \dot{m}_{\text{KH}} \left( 10^{-\beta} \log \frac{R_L}{R_{\text{ms}}} \right), \quad (3)$$

where all parameters correspond to accretor:  $R_L$  the Roche lobe radius,  $\dot{m}_{\text{KH}}$  is the Kelvin–Helmholtz time-scale accretion rate, and  $R_{\text{ms}}$  is the main-sequence radius. Numerical coefficient  $\beta$  is different for different types of stars and can be found in table 1 of Pols & Marinus (1994). This  $\dot{m}_{\max}$  corresponds to the condition that accretor stays within its Roche lobe. Equation (3) based on detailed simulations of stellar evolution for stars with masses below 17  $M_{\odot}$ .

Vinciguerra et al. (2020) analysed a range of fixed mass accretion efficiencies  $\alpha$  and found that values of  $\alpha = 0.5$  and  $\alpha = 0.75$  could explain the distribution of Be masses in BeXs. They noticed that a fixed value of  $\alpha = 0.5$  describes the observed population of BeXs in SMC best in terms of masses of Be stars and period distributions. There is no consensus on the unique mass accretion efficiency. For example, de Mink, Pols & Hilditch (2007) showed using detailed binary evolution simulations that there exists a spread in mass accretion efficiency. There are a few works where authors showed that the mass accretion efficiency should be less than 0.5 (e.g. Petrovic, Langer & van der Hucht 2005; Kruckow et al. 2018). Nevertheless, in our work we follow prescription by Vinciguerra et al. (2020). We implement a semiconservative mass transfer where the accretion rate is limited by

$$\dot{m}_{\max} = \frac{1}{2} \dot{m}, \quad (4)$$

where  $\dot{m}$  is the mass transfer rate set by donor expansion.

Another factor that could affect the final orbital parameters of binary is the relation between NS mass and the mass of its progenitor.

<sup>3</sup>The code is publicly available at <https://github.com/amusecode/SeBa>.

**Table 2.** List of models with relevant assumptions: Fryer and Antoniadis stand for a relation between progenitor and remnant masses based on Fryer et al. (2012) delayed explosion and Antoniadis et al. (2016), respectively. Hobbs corresponds to single Maxwellian natal kick with parameter  $\sigma = 265 \text{ km s}^{-1}$  (Hobbs et al. 2005); Igoshev corresponds to sum of two Maxwellians natal kick with  $w = 0.2$ , and parameters  $\sigma_1 = 56 \text{ km s}^{-1}$  and  $\sigma_2 = 336 \text{ km s}^{-1}$  (Igoshev 2020); Optimum stands for sum of two Maxwellians natal kick with  $w = 0.2$ , and parameters  $\sigma_1 = 45 \text{ km s}^{-1}$  and  $\sigma_2 = 336 \text{ km s}^{-1}$ . In model L core collapse SN receive natal kick in form of single Maxwellian with  $\sigma = 336 \text{ km s}^{-1}$  and ecSN receive single Maxwellian natal kick with  $\sigma = 45 \text{ km s}^{-1}$ . The asterisk (\*) means that ecSN occurs only in stars stripped of their hydrogen envelope.  $N_{\text{SMC}}$  is the total number of BeXs predicted by each model in the SMC.

Model	Natal kick	Remnant mass	Mass accretion efficiency	ecSN	Mass range for ecSN	$2\Delta\mathcal{L}$	$N_{\text{SMC}}$
A	Hobbs	Fryer	Default	No		99.98	200
B	Hobbs	Antoniadis	Default	No		107.8	210
C	Igoshev	Fryer	Default	No		50.44	320
D	Igoshev	Fryer	Semiconservative	No		1.3	180
E	No kick	Fryer	Semiconservative	No		20802.4	820
F	Hobbs	Fryer	Semiconservative	No		11.9	60
G0	Hobbs	Fryer	Semiconservative	Yes	1.83-2.25 (*)	12.4	90
G1	Hobbs	Fryer	Semiconservative	Yes	1.83-2.25	8.82	100
H0	Hobbs	Fryer	Semiconservative	Yes	1.63-2.45 (*)	10.9	170
H1	Hobbs	Fryer	Semiconservative	Yes	1.63-2.45	7.84	180
J0	Hobbs	Fryer	Semiconservative	Yes	1.6-2.25 (*)	10.66	90
J1	Hobbs	Fryer	Semiconservative	Yes	1.6-2.25	6.22	100
K	Optimum	Fryer	Semiconservative	No		-	140
L	New ecSN	Fryer	Semiconservative	Yes	1.63-2.45	5.18	140

Thus, in one of the variations we use the fitting equations by Fryer et al. (2012) delayed explosion. We also modify the maximum mass of a new born NS from 1.8 to 2.1  $M_{\odot}$ . In another variation we draw masses of NS from bimodal mass distribution found by Antoniadis et al. (2016) for millisecond radio pulsars, implemented as follows: for 42.5 per cent of all NSs we draw masses using normal distribution with mean 1.393  $M_{\odot}$  and standard deviation of  $\sigma_{m1} = 0.064$ , for remaining NS we draw masses from another normal distribution with mean 1.807  $M_{\odot}$  and standard deviation of  $\sigma_{m2} = 0.178$ . Note that in this procedure the NS mass is drawn randomly from the distribution, independent of its progenitor properties.

As for the natal kick, we consider seven different prescriptions. In our models A and B, we assign to each newly born NS a natal kick drawn from Maxwellian velocity distribution (see e.g. equation B2) with  $\sigma = 265 \text{ km s}^{-1}$  (Hobbs et al. 2005). In model E we assume NS receives no natal kick. Some isolated radio pulsars still receive a small velocity due to Blaauw effect (Blaauw 1961). In models C, D, F natal kick is chosen from two Maxwellians with  $\sigma_1 = 56 \text{ km s}^{-1}$  for 20 per cent of NS and with  $\sigma_2 = 336 \text{ km s}^{-1}$  for remaining NSs (Igoshev 2020).

In our model G0, H0, and J0 we select helium giants (evolved stars stripped of hydrogen envelope) and check if their initial helium core mass is within the ecSN mass range. We assume that such stars undergo an ecSN, receiving weak natal kicks drawn from the Maxwellian velocity distribution with  $\sigma = 30 \text{ km s}^{-1}$ . In these three models, we assume that ecSN occurs exclusively in stars which lost their hydrogen envelope due to mass transfer in line with original idea by Podsiadlowski et al. (2004), and more recently used by Willcox et al. (2021). In our models G1, H1, and J1 we also check if effectively isolated stars have helium core mass at the beginning of AGB in ranges as mentioned in Table 2. Such stars are also considered to undergo ecSN and receive a natal kick drawn from the Maxwellian velocity distribution with  $\sigma = 30 \text{ km s}^{-1}$ .

In our optimal (statistical) model K, we vary the fraction of weak natal kicks  $w$  and parameter of Maxwellian distribution  $\sigma_1$  for weak natal kicks while keeping the parameter  $\sigma_2$  for strong kicks fixed at value  $\sigma_2 = 336 \text{ km s}^{-1}$ . In our model L (physically motivated), we

use the prescription of the ecSN from H1 (model producing enough BeXs in the SMC) and assume that core-collapse SN receive single Maxwellian natal kick with parameter  $\sigma = 336 \text{ km s}^{-1}$  while ecSN receive natal kick drawn from the Maxwellian velocity distribution with  $\sigma = 45 \text{ km s}^{-1}$ .

To analyse the population synthesis simulation results, we proceed as follows. First, we keep track of the unbound NSs which formed in supernova explosion of the primary star. Primary star becomes isolated NS 17 times more frequently than the secondary star in our calculations, so we do not look at the velocities of secondaries which form NSs.<sup>4</sup> Second, we select binaries which, at some point of their evolution, have the NS primary and main sequence secondary more massive than 3  $M_{\odot}$  (i.e. the potential BeXs). We mark each of those binaries as

- (i) Be X if it experienced a mass transfer from the primary to the secondary star on thermal time-scale prior to SN explosion, see Fig. 2.
- (ii) CE if the binary experienced only CE stage prior to SN explosion.
- (iii) Effectively single if the binary experienced no mass transfer prior to SN explosion.
- (iv) Mixed if the binary experienced multiple mass transfer episodes of different nature prior to SN explosion.

In order to find the fraction of currently observable systems with particular orbital parameters we follow the procedure by Vinciguerra et al. (2020; see Appendix D). We use the star formation history by Rubele et al. (2015) to model the population of SMC and a constant star formation rate 1.65  $M_{\odot} \text{ yr}^{-1}$  for MW. We also keep track of the systemic velocities for BeXs. In our combined analysis, we use velocities of isolated NS and systemic velocities of BeXs weighted with the star formation rate for the MW.

<sup>4</sup>We examine the outcome of models D and H0 and noticed that addition of secondary NSs is hardly noticeable. We estimate that the histogram for isolated NS velocities changes at the level less than 0.5 per cent.

## 5 RESULTS: EXPLORATORY MODELS

With our exploratory model, we investigate the influence of four factors on formation of BeXs in SMC (see Table 2). These factors are as following: (1) natal kick distribution, (2) dependence between remnant mass and progenitor mass, (3) mass accretion efficiency during the mass transfer from primary to secondary, and (4) different natal kick for ecSN explosion.

### 5.1 Natal kicks

Our models A and C are identical, except for the natal kick prescription, which are described following Hobbs et al. (2005) or Igoshev (2020), respectively. Fig. 3 shows that the natal kick has little effect on the Be star mass distribution. However, it significantly affects the distribution of eccentricities (see Fig. 4). In case of model A, the peak of the distribution is shifted towards higher eccentricities ( $\sim 0.9$  compared to  $\sim 0.5$  in model C) and the low eccentricity systems – prevailing in the observed sample, are rare. However, the limited size of the observed BeX sample with measured eccentricities and potential biases (see Section 7.3) do not allow for any strong conclusions from this comparison. The natal kick prescription affects the total number of BeXs. In particular, in SMC, the prescription by Igoshev (2020) produces 1.6 times more BeXs than with the Hobbs et al. (2005) prescription. The natal kick prescription affects the distribution of measured velocities for both Galactic isolated radio pulsars and BeXs (see Fig. 5). The natal kick in form of Hobbs et al. (2005) produces multiple BeXs with systemic velocities above  $100 \text{ km s}^{-1}$  independently of assumptions about mass transfer.

We therefore restrict our further analysis to the discussion of the distributions of velocities of young radio pulsars and systemic velocities of BeXs in the Galaxy. Those distributions are very sensitive to the adopted natal kick prescription.

Comparing models D (natal kick distribution following Igoshev 2020) and E (no natal kick), one can notice that in the no-kick scenario the size of surviving BeX population is significantly increased (SMC could have up to 820 BeXs) and that BeXs with NS that formed with no kick have eccentricities below 0.3. However, if all NSs are formed with zero natal kick, the velocity distribution of radio pulsars is indistinguishable from the velocity distribution of massive stars (peaking at  $1\text{--}2 \text{ km s}^{-1}$  and showing a sharp decay at  $5 \text{ km s}^{-1}$ ), which is clearly incompatible with observations.

### 5.2 Dependence between remnant mass and progenitor mass

Our models A and B are identical, except for how we assign the NS mass depending on the mass of the progenitor. Figs 3 and 4 show that the distributions of Be masses and orbital elements are barely affected by the choice of prescription for the NS mass. The total number of BeXs in SMC is also comparable for both prescriptions. The exact choice of this dependence does not affect the distribution of isolated pulsar velocities and weakly affects the distribution of Be X-ray velocities (see Fig. 5). Therefore, we do not discuss the implications of alternative choice for relation between mass of the remnant and progenitor.

### 5.3 Mass accretion efficiency

Our models C and D are identical, except for the mass accretion efficiency. We confirm conclusions of Vinciguerra et al. (2020), showing that the mass accretion efficiency is a key variable influencing the Be star mass distribution of the simulated BeXs. Using the default SEBA

mass accretion efficiency prescription (model C), we find that the Be star mass distribution peaks at  $4\text{--}5 M_{\odot}$  – which is incompatible with observations. In model D, with (fixed) semiconservative mass accretion efficiency the distribution of Be star masses peaks at  $10\text{--}11 M_{\odot}$ , more in line with the observed distribution (see Fig. 3). Those conclusions are insensitive to the adopted natal kick prescription.

The mass accretion efficiency also controls the total number of BeXs in the SMC. For example, model F, which differs from model A only by prescription of this parameter, predicts 60 BeXs while model A predicts 200. Fig. 3 shows that fixed semiconservative mass accretion efficiency transforms systems, which could form BeX, to systems passing through a CE stage, especially in the Be mass range of  $3\text{--}7 M_{\odot}$ . That is why the number of BeXs in model F is significantly less than in model A.

The mass accretion efficiency weakly affects the distribution of synthetic binaries on  $P - e$  plot. This parameter is not important when comparing velocities of isolated radio pulsars. But the mass accretion efficiency is an important parameter for velocities of BeXs. The distribution of synthetic velocities for BeXs is very different for model C in comparison to model D. The model D is much more similar to the observed distribution. More quantitative comparison is available in Section 6.

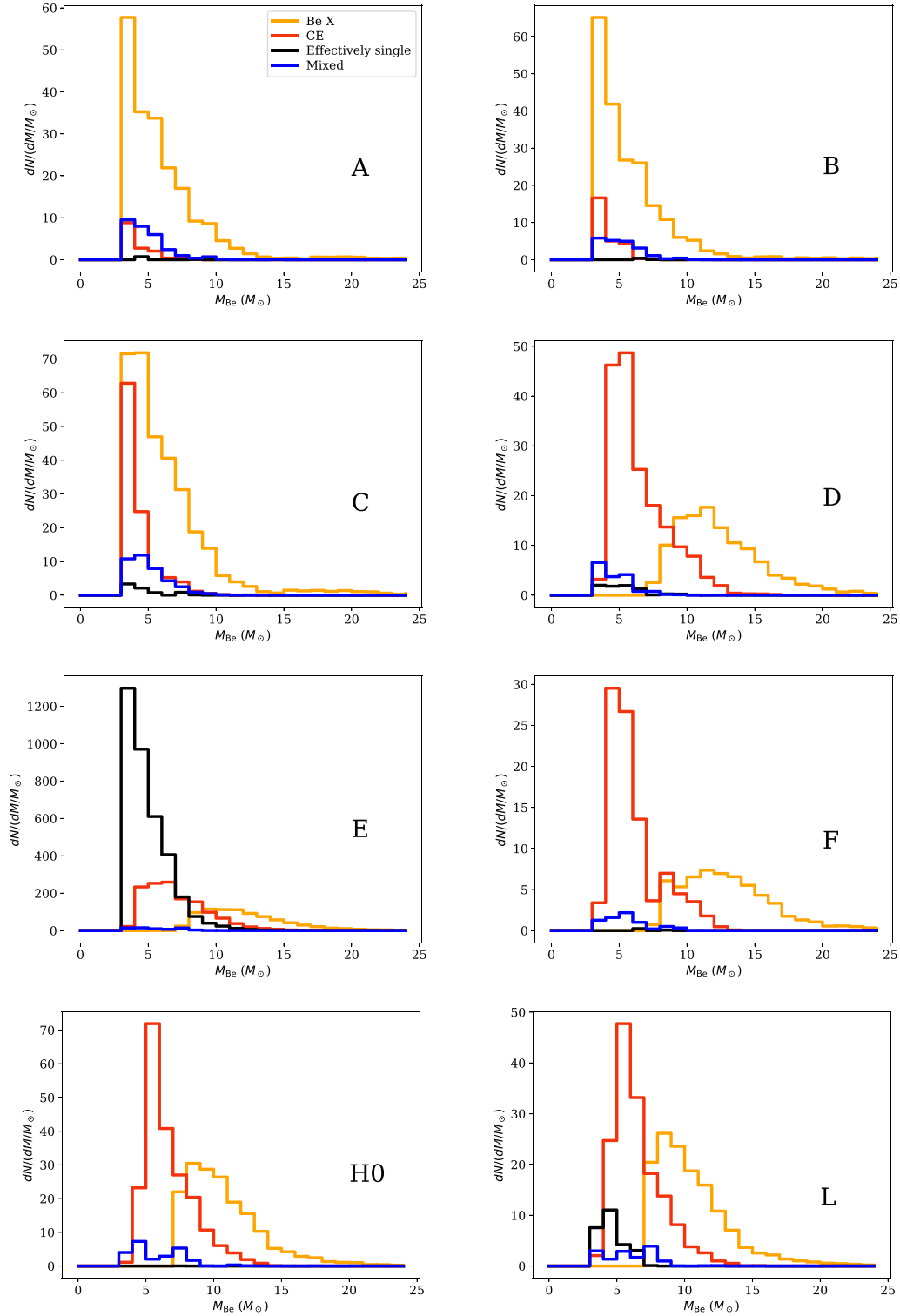
### 5.4 Electron capture supernova explosions

The model of electron capture explosions has multiple hidden dimensions. Even with our simplified prescription used in population synthesis to completely specify the model, we have to make the following assumptions: (1) choose a mass range for helium core needed to get ecSN, (2) choose if effectively isolated stars also produce ecSN, and (3) choose the magnitude of the natal kick, which is received by NS born in ecSN. It is not viable to restrict all these parameters in our study especially given the lack of observations for ecSN events. But it is worth to mention that current ecSN models are underconstrained and have to be further restricted by detailed simulations of supernova explosions and observations. In this section, we compare typical outcomes of assumptions (1) and (2) mentioned above.

We show the result of exploratory calculations for model H0 with ecSN explosions in Figs 3 and 4. A more detailed analysis is provided in Section 6.3. The model F has exactly the same assumptions as model H0 but no ecSN. Comparing these models, we see that allowing for ecSN increases the total number of BeXs. It also increases a fraction of binaries with lower eccentricities ( $e \approx 0.4$ ) which makes it more similar to observations. This comparison does not depend much on observational selection. The model F produces  $\approx 60$  BeXs which is not enough to explain all systems observed in the SMC. This model produces even fewer synthetic binaries in the region where orbital period and eccentricities are easiest to measure (orbital periods in range  $10\text{--}100 \text{ d}$  and  $e < 0.5$ ). For model H0, the Be mass distribution peaks close to the value found by Vinciguerra et al. (2020) for their preferred model computed with the COMPAS population synthesis code but with similar assumptions. The total expected number of BeXs in SMC in this model is  $\approx 170$ . In comparison to all other models that do not include any ecSN, the peak of the period distribution is shifted towards values of  $\sim 100$  to  $\sim 1000 \text{ d}$ .

Models where ecSN occurs only in stripped stars<sup>5</sup> (G0, J0, H0) produce synthetic radio pulsars with velocities very similar to the

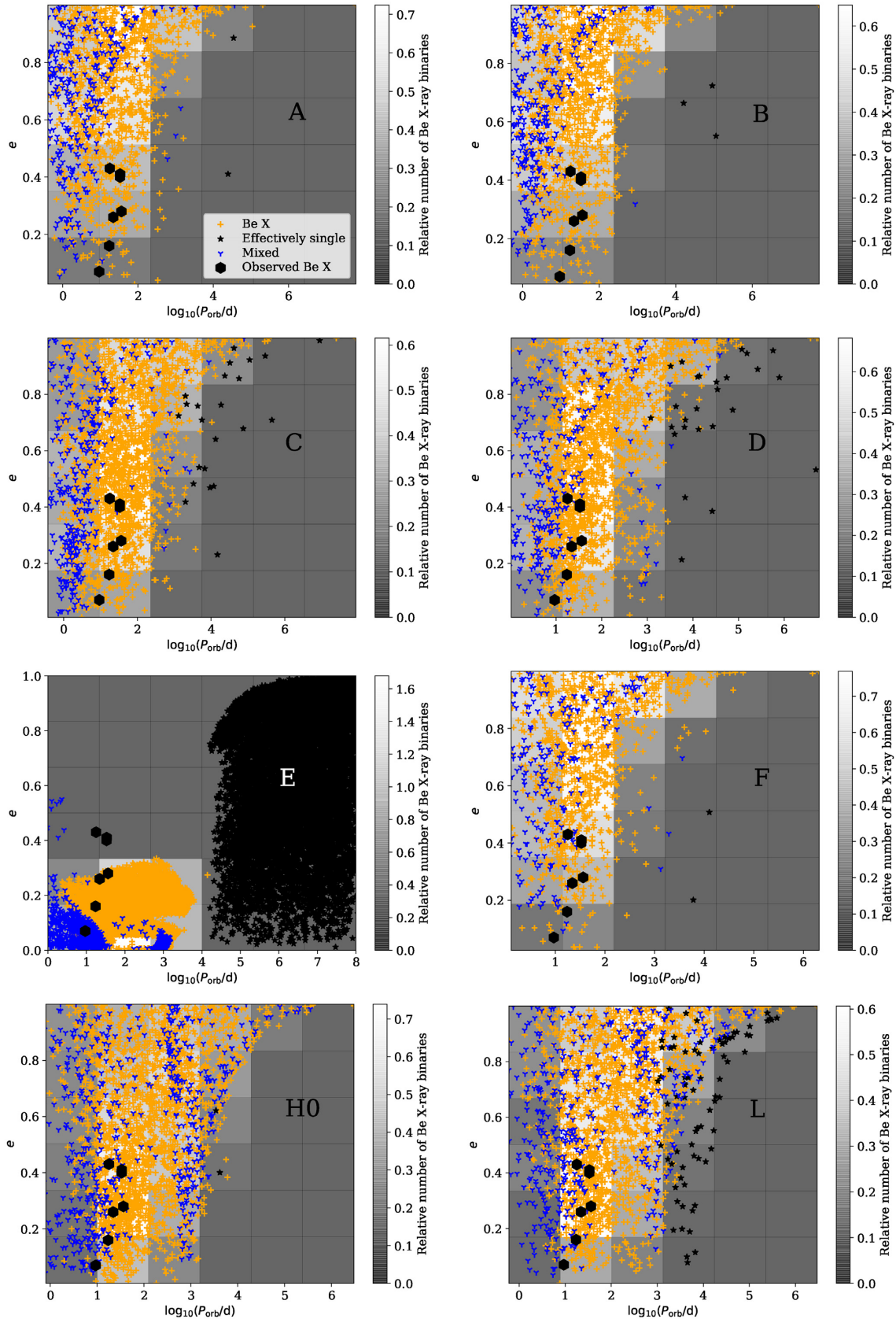
<sup>5</sup>Plots for all models are available in the online supplementary materials.



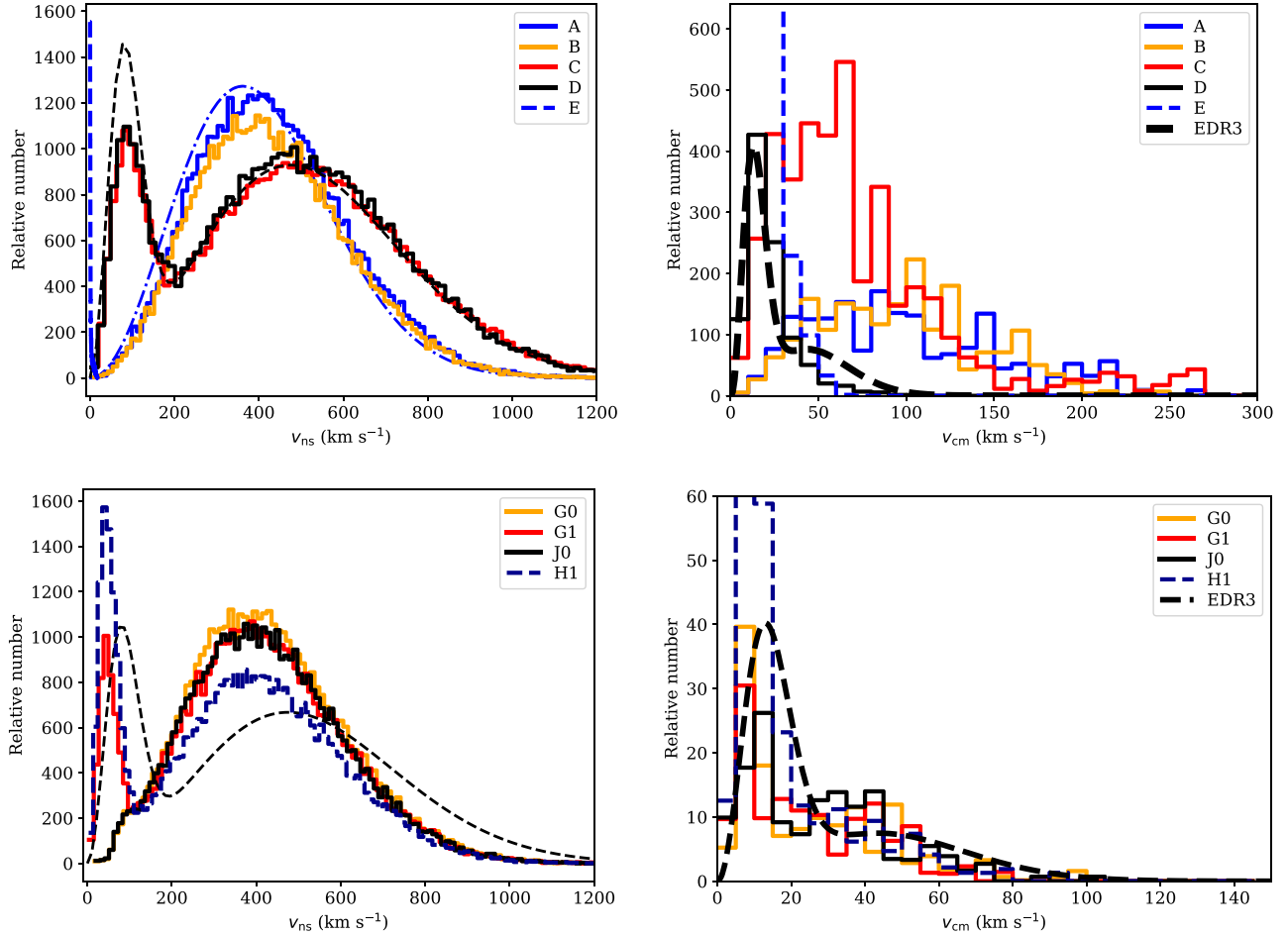
**Figure 3.** Histogram for mass of the secondary stars for simulated systems for different models. We perform the simulation using the SMC star formation history. Models A, B, C assume default SEBA mass transfer accretion efficiency prescription. Models D, E, F, H0, and L assume semiconservative mass transfer accretion efficiency and provide a better match to the mass distribution of Be stars in BeXs. Plots for the remaining models are available in the online supplementary materials.

Hobbs et al. (2005) distribution (see Fig. 5). All models with ecSN produce Galactic Be X-ray velocities that reasonably match observations. Models where ecSN occur also in effectively isolated stars (G1, J1, H1) produce many more isolated NS with velocities  $v$

$< 100 \text{ km s}^{-1}$ . The fraction of these NSs depend on the mass range assumed for ecSN progenitors: a wider mass range (H1) produces more slowly moving isolated radio pulsars and nearly two times more BeXs than a narrower range (G1).



**Figure 4.** Orbital periods and eccentricities for simulated binary systems for different models. Black hexagons are locations of observed BeXs in SMC by Coe & Kirk (2015). Histogram shows density of Be X-ray systems only. Plots for the remaining models are available in the online supplementary materials.



**Figure 5.** Left-hand panels: The velocity distribution of synthetic isolated NSs (histograms) compared with natal kick distributions by Hobbs et al. (2005) (dot and dashed blue line) and Igoshev (2020) (dashed black line). The low-velocity peak of the distribution Igoshev (2020) is noticeably higher than the same peak of the histogram for model D because a large fraction of these NSs stay bound in binaries. The SFH corresponds to MW. Only a fraction of 5–7 per cent of these systems could be ever discovered in radio surveys because of the beaming. Right-hand panels: The peculiar velocity distribution of BeXs obtained in simulations and restored based on Gaia EDR3 observations (black dashed line; normalization is chosen arbitrarily). Top panels: Models without ecSN. Bottom panels: Models with ecSN. It is worth noticing that model G1 (ecSN occurs also in effectively single stars) produces significantly more low-velocity isolated NSs in the velocity range of 10–100 km s<sup>−1</sup> in comparison to model G0 (ecSN occurs only in stripped stars). Plots for the remaining models are available in the online supplementary materials.

## 6 COMBINED ANALYSIS

In this section, we introduce our combined analysis that allows for a comparison between the observed velocities of young isolated radio pulsars, peculiar velocities of BeXs, and the binary population synthesis models. This kind of analysis allows us to examine multiple formation channels (with different natal kick prescriptions) simultaneously.

### 6.1 Structure of the combined analysis

The best representation of velocity distribution for pulsars and BeXs, which one can obtain using the binary population synthesis code is the histogram of three-dimensional velocities (see example in Fig. 5). Therefore, it is important to develop a technique that allows to quantitatively compare these histograms with observations taking into account observational uncertainties. In our view, such a comparison can be achieved by using binned velocity distribution. We define the binned velocity distribution as a function of

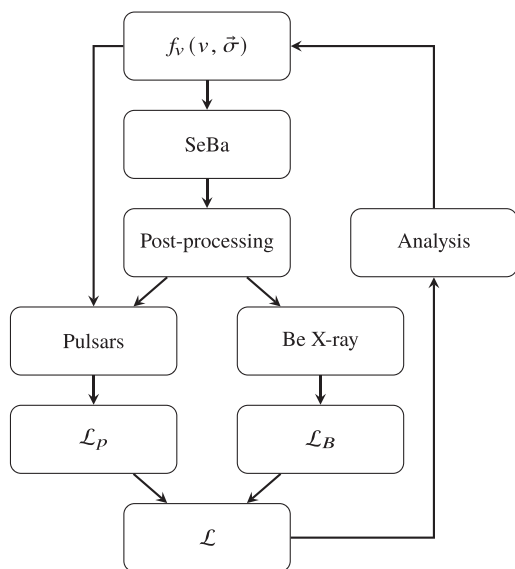
parameters  $\vec{m}$ :

$$f_v(v, \vec{m}) = \sum_{i=1}^N m_i v^2 v_h(v), \quad (5)$$

where function  $v_h(v)$  is a rectangular function (piece-wise constant) and it is equal to 1 in the range of velocities  $[i\Delta v, (i+1)\Delta v)$  and is equal to 0 otherwise,  $\Delta v$  is the size of velocity bin. The factor of  $v^2$  is present in  $f_v(\vec{m})$  because it is a part of Jacobian for the spherical coordinate system. The vector of parameters  $\vec{m}$  is normalized in such a way that

$$\int_0^\infty f_v(\vec{m}) dv = \sum_{i=1}^N m_i \Delta v^3 \left( \frac{(i+1)^3 - i^3}{3} \right) = 1. \quad (6)$$

This velocity distribution is also part of the joint probability for parallaxes and proper motion measurements (see Appendix C). Therefore, given measured parallaxes and proper motions for BeXs and isolated radio pulsars we can compute a likelihood for a



**Figure 6.** Schematic representation of the algorithm used to test various model natal kick distributions using the observational constraints based on the parallaxes and proper motions of the observed isolated radio pulsars and BeXs.  $f_v(v, \vec{\sigma})$  is the natal kick distribution, SEBA is the binary population synthesis code used in this work,  $\mathcal{L}_P$  and  $\mathcal{L}_B$  correspond to a stage when the likelihood is computed separately for isolated radio pulsars and BeXs.

distribution of peculiar velocities obtained from binary population synthesis simulations.

We present the algorithm for our combined analysis schematically in Fig. 6. This algorithm is the following: (1) we choose a natal kick distribution  $f_v(v, \vec{\sigma})$  and fix its parameters  $\vec{\sigma}$ , (2) we run SEBA binary population synthesis code for 20 000 ZAMS binaries using the natal kick distribution chosen in the previous step, and (3) analyse its output selecting isolated radio pulsars and BeXs (see details in Section 4.2). We prepare a histogram for velocities and convert number of systems in individual velocity bins  $h_k$  to parameters of the binned velocity distribution  $m_k$  using equation (C15). At the next step (4), we compute likelihood that the observed parallaxes and proper motions for isolated radio pulsars and BeXs are drawn from this model (see equation C14). Then (5), we combine these likelihoods. Finally, we select a next model or change the parameters and repeat the algorithm.

We can perform model selection using the Akaike information criterion (AIC). The AIC becomes the likelihood ratio test if model has no estimated parameters as in our case (besides model K):

$$\Delta\text{AIC} = 2\Delta\mathcal{L} = 2 \log \left( \frac{L_1}{L_2} \right), \quad (7)$$

where  $L_1$  and  $L_2$  are the likelihoods of two compared models. We provide  $2\Delta\mathcal{L}$  in our Table 2 comparing all models with model K. It is possible to quantify how probable model 1 is in comparison to model 2 as

$$p \sim \exp(-\mathcal{L}). \quad (8)$$

For example, the model E (no natal kick) is the most improbable model in our comparison. While model D is very similar to optimal model ( $p = 1.3$ ).<sup>6</sup> The AIC =  $2\mathcal{L}$  value of more than 9.2 means that model is hardly compatible with measurements  $p < 0.01$ .

<sup>6</sup>Strictly speaking, when we compare models with K we should subtract 4 from the difference because model K has two estimated parameters. But it

Our algorithm has four significant advantages over natal kick studies that were performed in the past (e.g. Hobbs et al. 2005; Verbunt et al. 2017; Igoshev 2020): (1) it allows to include effects related to binary evolution, (2) it allows to rigorously test any velocity distribution that can be implemented in the population synthesis code using exactly the same form of the likelihood function, (3) it can be easily extended to include more observational information for any given binary formation channel, and (4) it is straightforward to extend this analysis and include Markov chain Monte Carlo sampler to study correlations between parameters, if necessary.

## 6.2 Results of the combined analysis

Our exploratory model D (see Figs 5 and 4) seems to fit well all available observational constraints: (1) a peak of Be masses is located around  $11 M_\odot$ , (2) it produces  $\approx 180$  BeXs in SMC (compatible with estimates of 120 HMXBs by Maravelias et al. 2019 and Haberl & Sturm 2016), (3) BeXs with measured periods and eccentricities coincide with maximum density of simulated systems (could be affected by observational selection), (4) it leads to natal kick distribution of isolated radio pulsars that is the most compatible with the one found in the recent work by Igoshev (2020) (left-hand panel Fig. 5), and (5) the velocity distribution of BeXs provides a reasonable match to the observation-based distribution of the systemic velocities of Galactic BeXs estimated in Section 3 (right-hand panel Fig. 5).

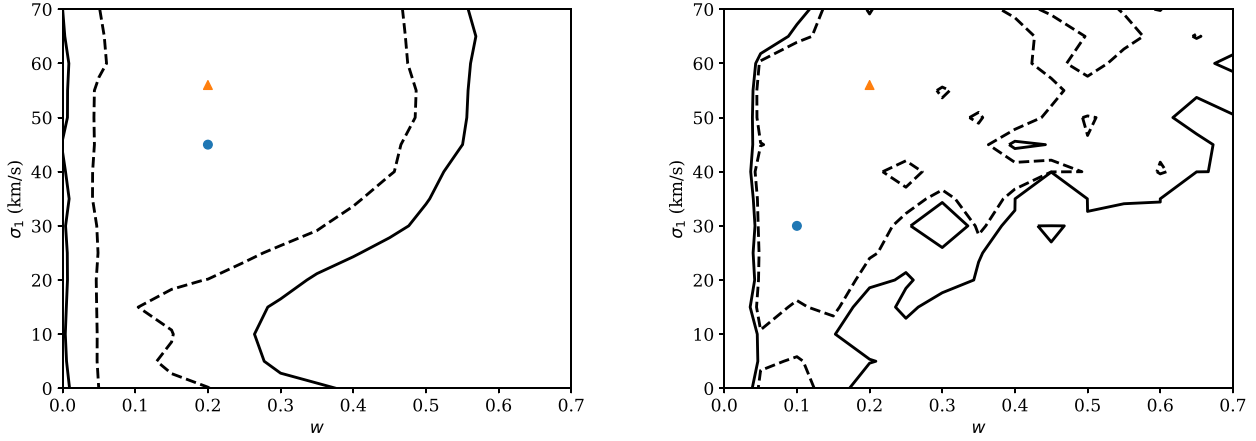
Further in this section, we use model D as our base model. We only vary the parameters that describe the low velocity part of the natal kick distribution ( $w$  and  $\sigma_1$ ). We do not vary parameter  $\sigma_2$  because: (1) only the isolated pulsar population, whose velocity distribution is relatively well constrained (Igoshev 2020), is sensitive to large velocities and (2) it is numerically expensive to study three parameters instead of two. We assume  $\sigma_2 = 336 \text{ km s}^{-1}$ , as found by Igoshev (2020).

We show the results of the optimization procedure in Figs 7 and 8. The acceptable values are concentrated in the top left corners of the corresponding figures. The confidence regions are somewhat noisy due to the stochastic nature of our models. Nevertheless, we can determine the maximum likelihood value from the combined analysis as  $w = 0.2 \pm 0.1$  and  $\sigma_1 = 45_{-15}^{+25} \text{ km s}^{-1}$  and fixed high-velocity component at  $\sigma_2 = 336 \text{ km s}^{-1}$ . Based on our combined analysis we can completely exclude  $w = 0$ , i.e. the situation in which only the high natal kick velocity component of the distribution is present. We also rule out cases with  $w > 0.5$ , i.e. when more than half of the NS population is born with weak natal kicks. The default parameters of model D are compatible with the optimal model but are slightly less preferred.

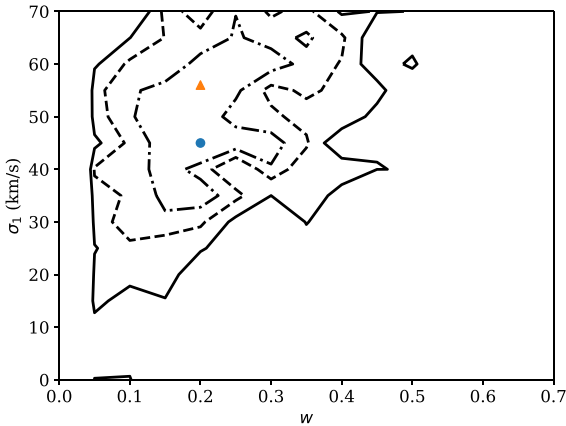
## 6.3 Electron capture supernova explosions

In this section, we test a hypothesis that slowly moving radio pulsars ( $|v| < 100 \text{ km s}^{-1}$ ) are predominantly produced via ecSN. It is clear that describing the NS formation in the ecSN based on the fixed (either ZAMS or  $M_{C, \text{BAGB}}$ ) mass limits is a severe simplification and the uncertainties in the models are considerable. None the less, given the fact that such prescription is widely used in population synthesis, it is instructive to discuss whether our analysis can rule out such a scenario.

is possible to compare models with D subtracting only 1.3 from the two log-likelihood difference because this model has no estimated parameters.



**Figure 7.** Likelihood profiles obtained for parallaxes and proper motions of isolated radio pulsars (left-hand panel) and BeXs (right-hand panel). Blue dot corresponds to the maximum likelihood  $w = 0.2$ ,  $\sigma_1 = 45$  (left-hand panel),  $w = 0.1$ ,  $\sigma_1 = 30$  (right-hand panel); orange triangle corresponds to parameters of model D. Dashed line corresponds to 95 per cent confidence interval and solid lines correspond to 99 per cent confidence interval.



**Figure 8.** Likelihood profiles obtained in combined analysis of parallaxes and proper motions of isolated radio pulsars and BeXs. Blue dot corresponds to the maximum likelihood  $w = 0.2$ ,  $\sigma_1 = 45$  (left-hand panel), orange triangle corresponds to parameters of model D. Dash-dotted line corresponds to 68 per cent confidence interval, dashed line corresponds to 95 per cent confidence interval, and solid line corresponds to 99 per cent confidence interval.

In order to rigorously test the origin of weak natal kicks via ecSN, we run six binary stellar evolution synthesis models (H0, H1, G0, G1, J0, and J1 models in Table 2). In these models, core collapse supernova explosions lead to natal kicks described by the Maxwellian velocity distribution with  $\sigma = 265 \text{ km s}^{-1}$  and ecSN lead to natal kicks with  $\sigma = 30 \text{ km s}^{-1}$ . Our basic mass range is  $[1.83, 2.25] M_{\odot}$  (models G0, G1). Chruslinska et al. (2018) consider an ad hoc, ‘optimistic’ model variation that allows to form a larger fraction of NS with low natal kicks in the ecSN scenario, i.e. assuming that stars with  $M_{\text{C, BAGB}}$  in a wider range of  $[1.63, 2.45] M_{\odot}$  can lead to ecSN (models H0 and H1). We also consider such model variation in this study to speculate whether the highly efficient ecSN scenario could explain the low velocity natal kick component of the radio pulsar velocity distribution. Additionally, we try the mass range  $[1.6, 2.25]$  used by Willcox et al. (2021) (models J0, J1).

In all cases we find that our statistical optimal model (K) describes the observations better. All models where only stripped stars produce ecSN, (G0, H0, J0) have  $2\Delta\mathcal{L} > 9.2$ , therefore their probability is

below 1 per cent. It happens because these models do not produce enough low-velocity isolated radio pulsars (Fig. 5). On the other hand, models G1, H1, and J1 could be accepted at some level ( $p = 0.012$ ,  $p = 0.02$ , and  $p = 0.04$ , respectively). These models clearly produce a significant fraction of low-velocity isolated radio pulsars (see Fig. 5 and the online supplementary materials).

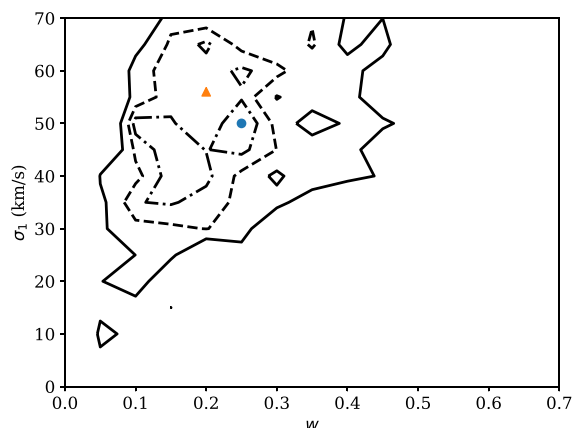
It is possible to improve these models and produce a model that explains observations and includes ecSN. It is our model L. In this new ecSN model core-collapse SNe impart Maxwellian natal kicks with  $\sigma = 336 \text{ km s}^{-1}$  and ecSNe (mass range  $1.63\text{--}2.45 M_{\odot}$ ; ecSN occurs also in effectively single stars) impart Maxwellian natal kicks with  $\sigma = 45 \text{ km s}^{-1}$ . This model has  $2\Delta\mathcal{L} = 5.18$  which means  $p = 0.075$ . This model produces 140 BeXs in the SMC. Therefore, this model describes both observations of isolated radio pulsars and BeXs well. It is worth mentioning that it is unclear if ecSN could physically occur in effectively isolated stars (i.e. stars that have not passed through the stage of stable mass transfer).

In comparison to Willcox et al. (2021), we do require ecSN to occur in effectively isolated stars. The main reason for this is that Willcox et al. (2021) used a non-parametric description for natal kick of NSs. This prescription already describes the natal kicks of isolated radio pulsars (this is how it was obtained in first place). Therefore, it is necessary for Willcox et al. (2021) to suppress the production of slowly moving isolated NS from binaries, which is done by prohibiting ecSN events in effectively isolated stars. In our case, the core collapse NS have natal kicks drawn from the Maxwellian velocity distribution with  $\sigma = 336 \text{ km s}^{-1}$  instead of drawing natal kicks from observed parallaxes and proper motions. It means that a very small fraction of NSs have velocities below  $50\text{--}100 \text{ km s}^{-1}$ . To add NS in this velocity range we do require ecSN in effectively isolated stars.

## 7 DISCUSSION

### 7.1 Consequence of binary fraction 85 per cent

In our basic simulations, we assume 100 per cent binary origin for progenitors of isolated radio pulsars. In reality, a fraction of NS could be formed from truly isolated progenitors. To check how sensitive our simulations are to this assumption, we perform additional simulations modifying model K by assuming that 15 per cent of isolated radio



**Figure 9.** Likelihood profiles obtained in combined analysis of parallaxes and proper motions of isolated radio pulsars and BeXs in the case when 15 per cent of isolated radio pulsars are formed from isolated progenitors. Blue dot corresponds to the maximum likelihood  $w = 0.25$ ,  $\sigma_1 = 50$ ; orange triangle corresponds to parameters of model D. Dash-dotted line corresponds to 68 per cent confidence interval, dashed line corresponds to 95 per cent confidence interval, and solid line corresponds to 99 per cent confidence interval.

pulsars originate from truly isolated progenitors. These systems receive exactly the same natal kick as other NS born in binaries.<sup>7</sup>

We show the likelihood profiles obtained from our new combined analysis in Figs 9. The resulting profiles resemble the ones obtained for our original model (see Figs 7 and 8). The overall maximum likelihood resulting from the combined analysis is shifted towards larger fraction of slower objects  $w = 0.25$  with slightly increased  $\sigma_1 = 50 \text{ km s}^{-1}$ . Nevertheless, this value is within the confidence interval of our original model. Thus, we conclude that the contribution of pulsars with truly isolated origin to the low kick velocity component of the distribution is minimal.

## 7.2 Initial velocity dispersion of SN progenitors

It is important to check how reliably BeX velocities trace binary peculiar motion. ZAMS binaries might have an initial velocity dispersion of  $\approx 10 \text{ km s}^{-1}$  (de Bruijne 1999; Kiminki & Smith 2018). On top of that, their observed velocity may contain a contribution from the velocity dispersion of their parent association of the order of  $\approx 5 \text{ km s}^{-1}$  (Reid et al. 2014). In our analysis of binaries with B stars we find  $\sigma = 11 \text{ km s}^{-1}$  (see Appendix E).

To quantify the effect of the velocity dispersion on our results, we perform an additional combined analysis based on model D where we add to each simulated pulsar and BeX a random velocity drawn from a Maxwellian distribution with  $\sigma = 11 \text{ km s}^{-1}$ . We show the results of our simulations in Fig. 10.

In the case of isolated radio pulsars,  $\sigma_1$  is lowered by  $10 \text{ km s}^{-1}$ . On the other hand, the response of BeXs is not that trivial. The solution with  $w = 0.2$  and  $\sigma_1 = 45 \text{ km s}^{-1}$  is nearly excluded and the preferable solution is  $w = 0.15$   $\sigma_1 = 10 \text{ km s}^{-1}$ . A range of solutions exists that are acceptable both by isolated radio pulsars and BeXs. Optimization of the combined model suggests that the minimum is located around  $w = 0.25$  and  $\sigma_1 = 45 \text{ km s}^{-1}$  (see

<sup>7</sup>Note that the observed velocity of NS born in binaries is the sum of natal kick and orbital velocity, while observed velocities of NSs born from truly isolated stars consist of the natal kick only.

Fig. 11), which is within the 68 per cent confidence interval of our original optimization.

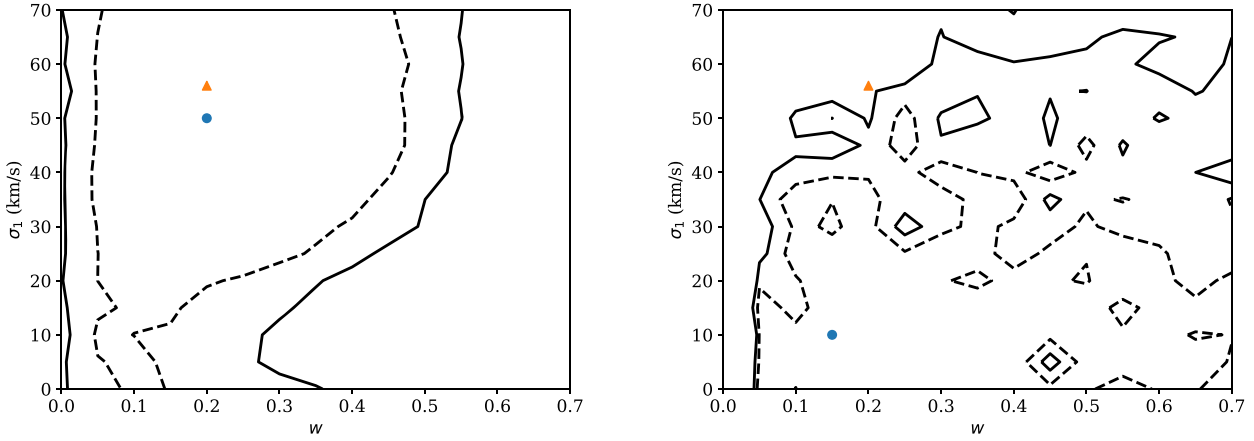
## 7.3 Effects of observational selection

Although our model D seems to satisfy a large number of observational constraints, it is worth mentioning that some of these constraints are a matter of strong observational selection. For example, there are only seven measurement of orbital periods and eccentricities for BeXs in the SMC (see e.g. Fig. 4). That is because eccentricity and orbital periods are difficult to obtain from X-ray observations. Many BeXs are only seen close to their periastron passages for a short period of time. These observations are sufficient to identify presence of the BeX and measure NS spin period (if it is less than  $\approx 500 \text{ s}$ ; Laycock et al. 2010), but do not allow to constrain the orbital period and especially eccentricity. In the case of SMC, many BeXs were discovered in RXTE scans. This technique is biased against the long-period binaries (Laycock et al. 2010; Vinciguerra et al. 2020).

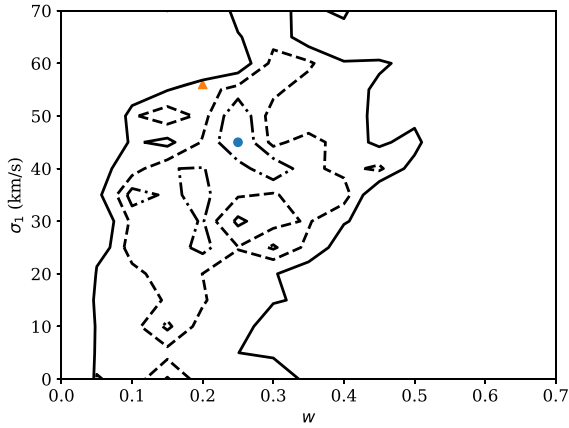
Only two BeXs with spin period longer than  $1000 \text{ s}$  are known within the Galactic sample, while our simplified simulations contain systems with much longer periods. Such systems might not be seen as Be X-ray pulsars because NS does not come close enough to interact with the decretion disc. Thus, some of those systems might be seen as single Be stars. The effect of this selection is the following: systemic velocity correlates with orbital and spin period. In Fig. 12, we show the natal kick of NS as the function of the final orbital period and eccentricity for BeXs. The Galactic BeXs have measured orbital periods ranging  $10\text{--}250 \text{ d}$  (see e.g. Raguzova & Popov 2005). In this range of orbital periods, we see binaries that receive a natal kick ranging from extremely small values  $\approx 10\text{--}15 \text{ km s}^{-1}$  to large values  $> 500 \text{ km s}^{-1}$ . Therefore, our analysis should be able to probe the full range of kick values. It seems that due to these selection effects we may preferentially miss the low kick systems in the observed sample (there are hardly any high kick binaries in the long orbital period part of the population). But if we limit the model sample to the range of orbital periods probed observationally, their velocity distribution would be different. We check how important it is for the result of the optimization procedure below.

We perform our combined analysis once again, including only synthetic BeXs with orbital periods in range  $10\text{--}250 \text{ d}$  to make it more comparable with observations. We show results in Fig. 13. This result is more noisy because fewer BeXs are included in the analysis. The result of optimization is  $w = 0.15$  and  $\sigma_1 = 40 \text{ km s}^{-1}$ . These values are well within the  $1\sigma$  confidence interval of our original optimal model. Therefore, simple exclusion of binaries with very short and very long orbital periods does not lead to a significantly different result.

It is possible to argue that BeXs with large eccentricities are absent in the observed sample based on Fig. 4. Thus, the observed sample of BeXs (see Fig. 12) is biased towards smaller natal kicks ( $< 100 \text{ km s}^{-1}$ ). This is not entirely true, because we include in our sample multiple BeXs with unknown eccentricity. In practice, it is much easier to discover BeX with large eccentricity than to measure its eccentricity. In principle, for system with eccentricity  $e > 0.5$  periastron is located at much closer distance to Be star, thus such NS could cause more disruption to the decretion disc and accrete more as a result. In the MW we know system J1845–024 with  $e = 0.88$  and  $P_{\text{orb}} = 242.18 \text{ d}$  (Raguzova & Popov 2005; Liu et al. 2006). We simulate the extreme selection by removing all systems with  $e > 0.5$  and plot the result in Fig. 14. The main effect of such selection is that the number of BeXs with systemic velocities above  $150 \text{ km s}^{-1}$  is decreased in the case of the Hobbs et al. (2005) natal kick distribution.



**Figure 10.** Likelihood profiles obtained for parallaxes and proper motions of isolated radio pulsars (left-hand panel) and BeXs (right-hand panel) in the case when initial velocity dispersion is added to simulations. Blue dot corresponds to the maximum likelihood:  $w = 0.2$  and  $\sigma_1 = 50$  km s<sup>-1</sup> (left-hand panel), and  $w = 0.15$  and  $\sigma_1 = 10$  km s<sup>-1</sup> (right-hand panel). Orange triangle corresponds to parameters of model D. Dashed line corresponds to 95 per cent confidence interval and solid lines correspond to 99 per cent confidence interval.

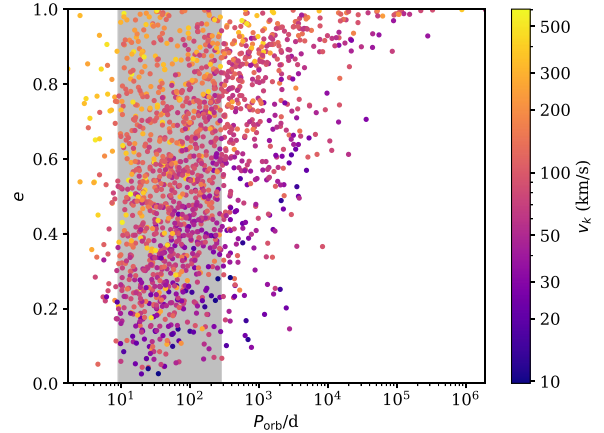


**Figure 11.** Likelihood profiles obtained for parallaxes and proper motions of isolated radio pulsars and BeXs in the case when initial velocity dispersion is added to simulations. Blue dot corresponds to the maximum likelihood  $w = 0.25$  and  $\sigma_1 = 45$  km s<sup>-1</sup>. Orange triangle corresponds to parameters of model D. Dashed line corresponds to 95 per cent confidence interval and solid lines correspond to 99 per cent confidence interval.

In the case of natal kick distribution by Igoshev (2020), the shape of systemic velocity distribution is conserved. Overall, even after applying this severe (and quite unreasonable) observational selection, our models A and D predict very different velocities for BeXs. Therefore, the effect of real observational selection is probably much less significant.

#### 7.4 Applicability of our results

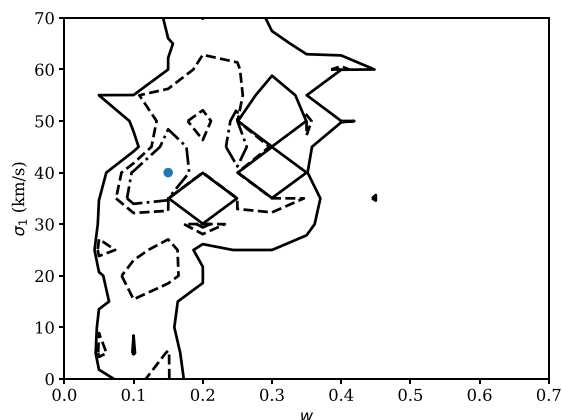
Our result can be applied to model binary stellar evolution especially concerning the first supernova explosion. If the readers are concerned with modelling of isolated pulsar population, it is much better to use results by Igoshev (2020) because the velocity distribution of isolated radio pulsars prefers slightly larger velocities than what results from the full natal kick distribution (NS with weaker kicks is preferentially bound in binaries).



**Figure 12.** Distribution of natal kicks for synthetic BeXs (colour) as a function of their orbital period and eccentricity. We show here the results of simulations D for MW metallicity. Grey region shows a range of orbital periods measured for Galactic BeXs based on Raguzova & Popov (2005).

One should be cautious when using our natal kick distribution to model the formation of DNS systems. Formation of such systems contains additional phases of interaction that may affect the formation of the second NS (and the related natal kick) and that cannot be constrained with BeX observations. In particular, the secondary star in the progenitor of a merging DNS is thought to be almost completely stripped of its envelope before the SN (i.e. it undergoes the so-called ultrastripped SN). Such stars might receive a very different natal kick. It is not feasible to test this hypothesis using the isolated radio pulsars and BeX populations because the required extreme stripping is thought to occur during the late phases of mass transfer in compact binaries, where the secondary star is stripped by the first-born compact object. The natal kicks of DNSs have to be tested using the observational information on millisecond radio pulsar and DNS populations.

Note that our best natal kick velocity distribution resembles those found by Fryer et al. (1998). The authors find that a fraction of  $\approx 30$  per cent of NS should receive almost no kick while the high-velocity pulsars form with natal kicks of the order of 600–700 km s<sup>-1</sup>.



**Figure 13.** Likelihood profiles obtained in combined analysis of parallaxes and proper motions of isolated radio pulsars and BeXs in the case when only BeXs with orbital periods in range 10–250 d are included. Blue dot corresponds to the maximum likelihood  $w = 0.15$  and  $\sigma_1 = 40 \text{ km s}^{-1}$ . Dash-dotted line corresponds to 68 per cent confidence interval, dashed line corresponds to 95 per cent confidence interval, and solid line corresponds to 99 per cent confidence interval.

Fryer et al. (1998) analyse multiple channels simultaneously including isolated radio pulsars, high- and low-mass X-ray binaries, and DNSs. They only concentrate on the production rates but do not consider orbital parameters and peculiar velocities of BeXs. So, it might be the case that our natal kick model is acceptable for DNSs, but future studies are required.

### 7.5 Caveat

In our analysis we treat merged stars as effectively isolated. Merger occurs in  $\approx 22_{-9}^{+26}$  per cent of binaries according to simulations by Renzo et al. (2019). Some of these stars could explode as supernova and form isolated radio pulsars which will effectively increase a number of radio pulsars with isolated progenitors. We assume that natal kicks of these NSs will be drawn from the same distribution as natal kicks of NS born from initially isolated progenitors. This

assumption has a caveat that these merger products might have a different structure of the core and might result in a different supernova explosion with somewhat different natal kick distribution for produced NSs. This hypothesis is worth of future investigation.

## 8 CONCLUSIONS

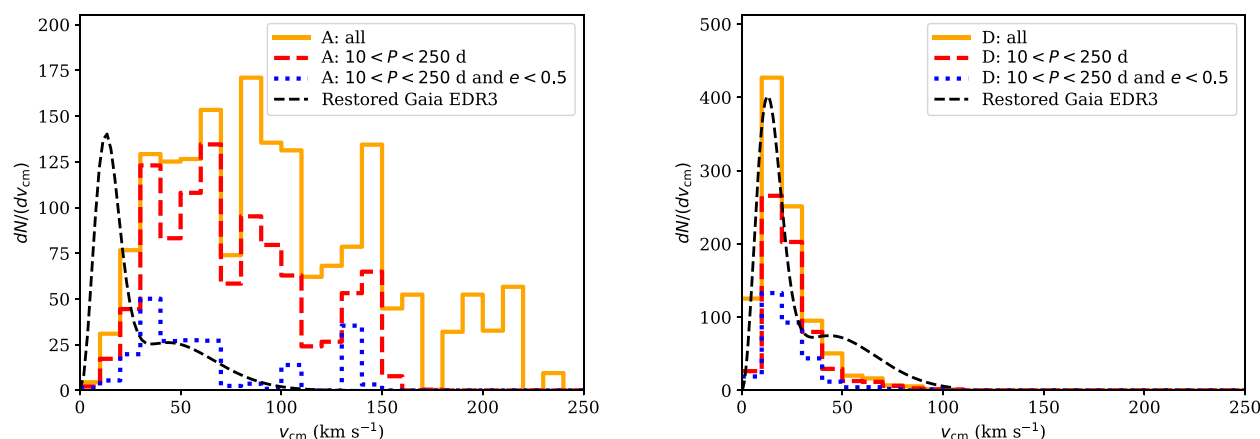
In this work, we revisit the natal kick velocity distribution of NSs in light of the available constraints of the observed populations of isolated radio pulsars and BeXs.

We introduce a framework that allows us to test any natal kick velocity distribution by confronting the velocity distributions of the mock isolated radio pulsars and BeX populations with the distributions obtained from observations. We modelled BeXs self-consistently to account for the effects of binary interactions.

Observational distributions of velocities are obtained using the parallaxes and proper motion measurements from radio interferometry for isolated radio pulsars (similar to the earlier study of isolated pulsars velocities by Verbunt et al. 2017 and Igoshev 2020), and peculiar velocities of Galactic BeXs from *Gaia* EDR3.

As a starting point for our analysis, we simulate the population of BeXs in the SMC and identify evolutionary models that can reasonably reproduce the demographics of the population: the size of the population and the mass distribution of Be stars. We also discuss the resulting distributions of orbital parameters (periods and eccentricities), however, in that case the potential severe observational selection effects limit the meaningful conclusions from the comparison with observations. Similarly as Vinciguerra et al. (2020), we find that the properties of the population are strongly affected by the assumptions about the mass transfer phase prior to NS formation. We find that the properties of the SMC BeXs are best reproduced if we assume that their progenitors evolve through a phase of stable, semiconservative mass transfer, consistent with the results of Vinciguerra et al. (2020). Therefore, in our default model for the formation of (Galactic and SMC) BeXs we require an episode of the stable, semiconservative mass transfer.

We identify the following comparison metrics: (1) total number of SMC BeXs, (2) dependence of eccentricity on orbital period, (3) mass of Be stars, and (4) velocities of Galactic radio pulsars and BeXs. We find that the natal kick distributions significantly affect



**Figure 14.** Possible influence of the observational selection on the peculiar velocity distribution of the model BeX population (different panels and colours) compared with the distribution based on the *Gaia* EDR3 observations (black dashed line; height is chosen arbitrary). Left-hand panel: model A. Right-hand panel: model D. Different types of lines correspond to different possible selection effects discussed in the text. The solid orange line shows all synthetic binaries (no observational selection). The red dashed line corresponds to the case when we can observe only BeXs with orbital periods in range 10–250 d. The blue dotted line corresponds to the case when we can only observe binaries with orbital periods in range 10–250 d and eccentricities smaller than 0.5.

the total number of produced BeXs in the SMC, the eccentricities of BeXs, and the velocities of both Galactic radio pulsars and BeXs. The exact dependence between remnant and progenitor mass hardly affects any of our metrics. The mass accretion efficiency affects the total number of BeXs, Be mass, and distribution of velocities for Galactic BeXs.

We find that using the bimodal Maxwellian natal kick distribution from Igoshev (2020) (based on the isolated radio pulsar data) we can obtain a reasonable match of the model with the observation-based distribution of Galactic systems. We further apply our framework to find the optimal parameters of the low-velocity component of this distribution in light of the additional constraints from BeXs. The optimization procedure results in a fraction of NS forming with kicks in the low-velocity component of  $w = 0.2 \pm 0.1$  and the Maxwellian parameter  $\sigma_1 = 45^{+25}_{-15} \text{ km s}^{-1}$  and we fix  $\sigma_2 = 336 \text{ km s}^{-1}$ .

Finally, we show that the natal kick prescription commonly used in population synthesis studies: combining the distribution from Hobbs et al. (2005) and allowing for low-velocity NS formation in ecSN with a Maxwellian distribution with  $\sigma = 30 \text{ km s}^{-1}$  is incompatible with observations of isolated radio pulsars if ecSN occurs only in stripped stars. If ecSN occurs also in effectively isolated stars, this distribution is marginally compatible with observations. It is possible to improve this model and suggest that core collapse supernova explosions impart a natal kick with a Maxwellian distribution with  $\sigma = 336 \text{ km s}^{-1}$  while ecSN occurs in both effectively isolated and stripped stars with helium core mass in the range  $1.63\text{--}2.25 M_{\odot}$  at the base of asymptotic giant branch. These NS receive Maxwellian natal kick with  $\sigma = 45 \text{ km s}^{-1}$ . This model is slightly less preferable than our statistical model K, but it is still compatible with all available observations. This model can be improved further when a larger sample of low-velocity pulsars becomes available. Whether it is physically possible to form an NS in ecSN explosions in this broad mass range is unclear at the moment. Further modelling is required to resolve this issue.

## ACKNOWLEDGEMENTS

AIP is grateful to Prof. Gijs Nelemans, Eva Laplace, Dr Serena Vinciguerra, Anastasia Frantsuzova, and Prof. Sergei Popov, and all participants of SEBA and binary stellar evolution meeting for multiple fruitful discussions. AI was supported by STFC grant no. ST/S000275/1. ST acknowledge support from the Netherlands Research Council NWO (VENI 639.041.645 grants). This work was undertaken on ARC4, part of the High Performance Computing facilities at the University of Leeds, UK. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. MC acknowledges support from the Netherlands Organisation for Scientific Research (NWO).

## DATA AVAILABILITY STATEMENT

The data underlying this article will be shared on reasonable request to the corresponding author.

## REFERENCES

Abt H. A., 1983, *ARA&A*, 21, 343  
Ambartsumian V. A., 1937, *Astron. Zh.*, 14, 207

- Andrews J. J., Farr W. M., Kalogera V., Willems B., 2015, *ApJ*, 801, 32  
Antoniadis J., 2020, *Res. Notes Am. Astron. Soc.*, 4, 223  
Antoniadis J., 2021, *MNRAS*, 501, 1116  
Antoniadis J., Tauris T. M., Ozel F., Barr E., Champion D. J., Freire P. C. C., 2016, preprint ([arXiv:1605.01665](https://arxiv.org/abs/1605.01665))  
Beniamini P., Piran T., 2016, *MNRAS*, 456, 4089  
Bildsten L. et al., 1997, *ApJS*, 113, 367  
Blaauw A., 1961, *Bull. Astron. Inst. Neth.*, 15, 265  
Bodaghee A. et al., 2021, *ApJ*, 919, 81  
Bozzo E., Pavan L., Ferrigno C., Falanga M., Campana S., Paltani S., Stella L., Walter R., 2012, *A&A*, 544, A118  
Brandt N., Podsiadlowski P., 1995, *MNRAS*, 274, 461  
Briskin W. F., Benson J. M., Goss W. M., Thorsett S. E., 2002, *ApJ*, 571, 906  
Briskin W. F., Fruchter A. S., Goss W. M., Herrnstein R. M., Thorsett S. E., 2003, *AJ*, 126, 3090  
Casares J., Ribó M., Ribas I., Paredes J. M., Martí J., Herrero A., 2005, *MNRAS*, 364, 899  
Chatterjee S., Cordes J. M., Lazio T. J. W., Goss W. M., Fomalont E. B., Benson J. M., 2001, *ApJ*, 550, 287  
Chatterjee S., Cordes J. M., Vlemmings W. H. T., Arzoumanian Z., Goss W. M., Lazio T. J. W., 2004, *ApJ*, 604, 339  
Chatterjee S. et al., 2009, *ApJ*, 698, 250  
Chevalier C., Ilovaisky S. A., 1998, *A&A*, 330, 201  
Chruslinska M., Belczynski K., Klencki J., Benacquista M., 2018, *MNRAS*, 474, 2937  
Coe M. J., Kirk J., 2015, *MNRAS*, 452, 969  
Cox N. L. J., Kaper L., Mokiem M. R., 2005, *A&A*, 436, 661  
de Bruijne J. H. J., 1999, *MNRAS*, 310, 585  
de Mink S. E., Pols O. R., Hilditch R. W., 2007, *A&A*, 467, 1181  
Deller A. T., Tingay S. J., Bailes M., Reynolds J. E., 2009, *ApJ*, 701, 1243  
Deller A. T. et al., 2019, *ApJ*, 875, 100  
Dewey R. J., Cordes J. M., 1987, *ApJ*, 321, 780  
Dewi J. D. M., Podsiadlowski P., Pols O. R., 2005, *MNRAS*, 363, L71  
Doherty C. L., Gil-Pons P., Siess L., Lattanzio J. C., 2017, *Publ. Astron. Soc. Aust.*, 34, e056  
Duchêne G., Kraus A., 2013, *ARA&A*, 51, 269  
Edwards R. T., Hobbs G. B., Manchester R. N., 2006, *MNRAS*, 372, 1549  
Eldridge J. J., Tout C. A., 2004, *MNRAS*, 353, 87  
Faucher-Giguère C.-A., Kaspi V. M., 2006, *ApJ*, 643, 332  
Fryer C., Burrows A., Benz W., 1998, *ApJ*, 496, 333  
Fryer C. L., Belczynski K., Wiktorowicz G., Dominik M., Kalogera V., Holz D. E., 2012, *ApJ*, 749, 91  
Gaia Collaboration, 2016, *A&A*, 595, A1  
Gaia Collaboration, 2021, *A&A*, 649, A1  
Gamen R., Barbà R. H., Walborn N. R., Morrell N. I., Arias J. I., Maíz Apellániz J., Sota A., Alfaro E. J., 2015, *A&A*, 583, L4  
González-Galán A., Negueruela I., Castro N., Simón-Díaz S., Lorenzo J., Vilardell F., 2014, *A&A*, 566, A131  
Haberl F., Sturm R., 2016, *A&A*, 586, A81  
Heggie D. C., 1975, *MNRAS*, 173, 729  
Hills J. G., 1983, *ApJ*, 267, 322  
Hobbs G., Lorimer D. R., Lyne A. G., Kramer M., 2005, *MNRAS*, 360, 974  
Hurley J. R., Pols O. R., Tout C. A., 2000, *MNRAS*, 315, 543  
Igoshev A. P., 2020, *MNRAS*, 494, 3663  
Igoshev A. P., Perets H. B., 2019, *MNRAS*, 486, 4098  
Igoshev A. P., Perets H. B., Michaely E., 2020, *MNRAS*, 494, 1448  
Ivanova N. et al., 2013, *A&AR*, 21, 59  
Janka H.-T., 2017, *ApJ*, 837, 84  
Jones S., Röpke F. K., Pakmor R., Seitenzahl I. R., Ohlmann S. T., Edelmann P. V. F., 2016, *A&A*, 593, A72  
Kalogera V., 1996, *ApJ*, 471, 352  
Kiminki M. M., Smith N., 2018, *MNRAS*, 477, 2068  
Kirsten F., Vlemmings W., Campbell R. M., Kramer M., Chatterjee S., 2015, *A&A*, 577, A111  
Kobulnicky H. A., Fryer C. L., 2007, *ApJ*, 670, 747  
Kriz S., Harmanec P., 1975, *Bull. Astron. Inst. Czech.*, 26, 65  
Kruckow M. U., Tauris T. M., Langer N., Kramer M., Izzard R. G., 2018, *MNRAS*, 481, 1908

- Kudritzki R. P., Simon K. P., 1978, *A&A*, 70, 653
- Kuranov A. G., Popov S. B., Postnov K. A., 2009, *MNRAS*, 395, 2087
- Laycock S., Zezas A., Hong J., Drake J. J., Antoniou V., 2010, *ApJ*, 716, 1217
- Lindgren L. et al., 2018, *A&A*, 616, A2
- Lindgren L. et al., 2021, *A&A*, 649, A2
- Lipunov V. M., Postnov K. A., Prokhorov M. E., 1997, *MNRAS*, 288, 245
- Liu Q. Z., van Paradijs J., van den Heuvel E. P. J., 2006, *A&A*, 455, 1165
- Lutovinov A. A., Revnivtsev M. G., Tsygankov S. S., Krivonos R. A., 2013, *MNRAS*, 431, 327
- Lyne A. G., Lorimer D. R., 1994, *Nature*, 369, 127
- Manchester R. N., Hobbs G. B., Teoh A., Hobbs M., 2005, *AJ*, 129, 1993
- Maravelias G., Zezas A., Antoniou V., Hatzidimitriou D., Haberl F., 2019, *Proc. IAU Symp.* 346, High-mass X-ray Binaries: Illuminating the Passage from Massive Binaries to Merging Compact Objects. Kluwer, Dordrecht, p. 350
- Margon B., 1984, *ARA&A*, 22, 507
- Masetti N. et al., 2009, *A&A*, 495, 121
- Miyaji S., Nomoto K., Yokoi K., Sugimoto D., 1980, *PASJ*, 32, 303
- Moe M., Di Stefano R., 2017, *ApJS*, 230, 15
- Mróz P. et al., 2019, *ApJ*, 870, L10
- Negueruela I., Torrejón J. M., McBride V., 2007, *Astron. Telegram*, 1239, 1
- Nomoto K., 1984, *ApJ*, 277, 791
- Nomoto K., 1987, *ApJ*, 322, 206
- Nomoto K., Kondo Y., 1991, *ApJ*, 367, L19
- Orosz J. A., McClintock J. E., Aufdenberg J. P., Remillard R. A., Reid M. J., Narayan R., Gou L., 2011, *ApJ*, 742, 84
- Pellizza L. J., Chaty S., Negueruela I., 2006, *A&A*, 455, 653
- Petrovic J., Langer N., van der Hucht K. A., 2005, *A&A*, 435, 1013
- Podsiadlowski P., Langer N., Poelarends A. J. T., Rappaport S., Heger A., Pfahl E., 2004, *ApJ*, 612, 1044
- Poelarends A. J. T., Herwig F., Langer N., Heger A., 2008, *ApJ*, 675, 614
- Poelarends A. J. T., Wurtz S., Tarka J., Cole Adams L., Hills S. T., 2017, *ApJ*, 850, 197
- Pols O. R., Marinus M., 1994, *A&A*, 288, 475
- Portegies Zwart S. F., 1995, *A&A*, 296, 691
- Portegies Zwart S. F., Verbunt F., 1996, *A&A*, 309, 179
- Porter J. M., 1996, *MNRAS*, 280, L31
- Pourbaix D. et al., 2004, *A&A*, 424, 727
- Prišegen M., 2020, *A&A*, 640, A86
- Raguzova N. V., Popov S. B., 2005, *Astron. Astrophys. Trans.*, 24, 151
- Rappaport S., van den Heuvel E. P. J., 1982, in Jaschek M., Groth H. G., eds, *Proc. IAU Symp.* 98, Be Stars. Kluwer, Dordrecht, p. 327
- Reid M. J. et al., 2014, *ApJ*, 783, 130
- Reig P., 2011, *Ap&SS*, 332, 1
- Reig P., Zezas A., 2014, *A&A*, 561, A137
- Renzo M. et al., 2019, *A&A*, 624, A66
- Rubele S. et al., 2015, *MNRAS*, 449, 639
- Sakamaki A., Negoro H., 2017, in Serino M., Shidatsu M., Iwakiri W., Mihara T., eds, 7 years of MAXI: Monitoring X-ray Transients. RIKEN, Wako, Saitama, Japan, p. 167
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Sana H. et al., 2012, *Science*, 337, 444
- Sguera V. et al., 2007, *A&A*, 462, 695
- Shao Y., Li X.-D., 2014, *ApJ*, 796, 37
- Siess L., Lebreuilly U., 2018, *A&A*, 614, A99
- Slettebak A., 1982, *ApJS*, 50, 55
- Smith M. A., Lopes de Oliveira R., Motch C., 2016, *Adv. Space Res.*, 58, 782
- Tauris T. M., Manchester R. N., 1998, *MNRAS*, 298, 625
- Tauris T. M., Takens R. J., 1998, *A&A*, 330, 1047
- Tauris T. M., Langer N., Moriya T. J., Podsiadlowski P., Yoon S.-C., Blinnikov S. I., 2013, *ApJ*, 778, L23
- Tauris T. M. et al., 2017, *ApJ*, 846, 170
- Toonen S., Nelemans G., Portegies Zwart S., 2012, *A&A*, 546, A70
- Townsend R. H. D., Owocki S. P., Howarth I. D., 2004, *MNRAS*, 350, 189
- Verbunt F., Cator E., 2017, *J. Astrophys. Astron.*, 38, 40
- Verbunt F., Igoshev A., Cator E., 2017, *A&A*, 608, A57
- Vinciguerra S. et al., 2020, *MNRAS*, 498, 4705
- Wenger M. et al., 2000, *A&AS*, 143, 9
- Willcox R., Mandel I., Thrane E., Deller A., Stevenson S., Vigna-Gómez A., 2021, preprint ([arXiv:2107.04251](https://arxiv.org/abs/2107.04251))
- Woosley S. E., Heger A., 2015, *ApJ*, 810, 34
- Xu Y. et al., 2018, *A&A*, 616, L15
- Yudin R. V., 2001, *A&A*, 368, 912

## SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

**Figure 3.** Histogram for mass of the secondary stars for simulated systems for different models.

**Figure 4.** Orbital periods and eccentricities for simulated binary systems for different models.

**Figure 5.** Left-hand panels: The velocity distribution of synthetic isolated NSs (histograms) compared with natal kick distributions by Hobbs et al. (2005) (dot and dashed blue line) and Igoshev (2020) (dashed black line).

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

## APPENDIX A: CATALOGUE OF GALACTIC BE X-RAY BINARIES

We found the following information about individual HMXBs in the literature.

[KRL2007b] 335 also known as IGR J18450–0435 is a supergiant system (Sguera et al. 2007). This system is excluded from analysis.

2MASS J16193220–4944305 also known as IGR J16195–4945 is a supergiant system (Lutovinov et al. 2013). This system is excluded from analysis.

2MASS J21342037+4738002 also known as IGR J21343+4738 is a B1Ve star (Reig & Zezas 2014). This system is included in the analysis.

UCAC2 4813819 also known as IGR J11435–6109 is B0Ve or B2Ve (Negueruela, Torrejón & McBride 2007; Masetti et al. 2009). This system is included in the analysis.

AX J1739.1–3020 is probably supergiant system (Sakamaki & Negoro 2017). This system is excluded from analysis.

V\* GP Vel also known as Vela X-1 is a supergiant system (see e.g. Bildsten et al. 1997). This system is excluded from analysis.

4U 1907+09 is a supergiant system (Cox, Kaper & Mokiem 2005). This system is excluded from analysis.

SS 188 also known as IGR J08262–3736 is a supergiant system (Bozzo et al. 2012). This system is excluded from analysis.

V\* V479 Sct is a possible counterpart of LS 5039 which seems to be a black hole (Casares et al. 2005). This system is excluded from analysis.

BD+60 73 also known as IGR J00370+6122 is a supergiant system (González-Galán et al. 2014). This system is excluded from analysis.

EM\* AS 14 also known as TYC 3681-695-1 is B1-2 III-Ve star (Smith, Lopes de Oliveira & Motch 2016). This system is included in the analysis.

HD 74194 also known as LM Vel is a supergiant system (Gamen et al. 2015). This system is excluded from analysis.

**Table A1.** List of HMXBs with optical counterpart in the *Gaia* EDR3 with well measured parallax ( $\varpi'/\sigma_\varpi > 3$ ). *Y* in the third column indicates that object is included in the velocity analysis.

HMXB	Sp. type	Inc	<i>Gaia</i> counterpart Gaia EDR3 ID	<i>V</i> (mag)	<i>g</i> (mag)	$\varpi \pm \sigma_\varpi$ (mag)	$\mu_\alpha \pm \sigma_\alpha$ (mas yr <sup>-1</sup> )	$\mu_\delta \pm \sigma_\delta$ (mas yr <sup>-1</sup> )
[KRL2007b] 335	O9Ia	N	4258160560148155648	14.06	12.76	0.164 ± 0.024	-1.366 ± 0.024	-5.595 ± 0.022
GRO J2058+42	O9.5-B0IV-Ve	Y	2065653598916388352	14.74	14.13	0.078 ± 0.015	-2.21 ± 0.015	-3.351 ± 0.017
2MASS J16193220-4944305	B1sg	N	5935509395659726592	16.8	16.37	0.359 ± 0.051	-0.184 ± 0.062	-0.545 ± 0.044
2MASS J21342037+4738002	BII Ve	Y	1978365123143522176	14.16	14.0	0.084 ± 0.014	-2.212 ± 0.015	-2.558 ± 0.015
LS 1698	B0III/V:e	Y	5352018121173519488	11.48	11.24	0.171 ± 0.016	-6.305 ± 0.021	3.01 ± 0.018
UCAC2 39636510	B0.5Ve	Y	3423526544838563328	12.21	12.17	0.138 ± 0.018	0.573 ± 0.02	-0.608 ± 0.014
LS I +61 303	B0Ve	Y	465645515129855872	10.75	10.4	0.378 ± 0.013	-0.423 ± 0.011	-0.256 ± 0.012
UCAC2 4813819	B2III/BOV	Y	5335021599905643264	13.4	14.73	0.242 ± 0.021	-3.448 ± 0.021	1.165 ± 0.021
<i>Ginga</i> 0834-430	B0-2III-Ve	Y	5523448274762133632	20.4	19.17	1.105 ± 0.217	-3.235 ± 0.215	3.755 ± 0.266
[KRL2007b] 84	B0e	Y	5258414192353423360	15.27	13.88	0.243 ± 0.013	-4.702 ± 0.016	3.559 ± 0.14
2E 1752	O9.7Ve	Y	3052677318793446016	10.9	11.99	0.154 ± 0.015	-0.638 ± 0.015	1.256 ± 0.014
GSC 03588-00834	B0Ve	Y	2162805896614571904	14.2	13.77	0.131 ± 0.013	-3.505 ± 0.014	-3.16 ± 0.013
V* V635 Cas	B0.2Ve	Y	524677469790488960	15.19	14.3	0.136 ± 0.016	-1.684 ± 0.013	0.504 ± 0.017
V* BQ Cam	O8.5Ve	Y	444752973131169664	15.42	14.2	0.134 ± 0.02	-0.268 ± 0.02	0.44 ± 0.02
HD 34921	B0IVpe	Y	184497471323752064	7.48	7.23	0.721 ± 0.03	1.305 ± 0.041	-3.999 ± 0.028
2MASS J17002524-4219003	B2e	Y	5966213219190201856	9.15	8.71	0.641 ± 0.023	1.181 ± 0.03	-1.47 ± 0.023
V* V441 Pup	O5Ve	Y	5613494119551805184	11.83	11.6	0.096 ± 0.017	-0.881 ± 0.012	1.785 ± 0.018
AX J1739.1-3020	O8.5Iab(f)	N	4056922100878037120	14.8	16.22	0.68 ± 0.053	2.954 ± 0.062	1.821 ± 0.041
BD+53 2790	O9.5Vep	Y	2005653524280214400	9.84	9.74	0.305 ± 0.014	-4.173 ± 0.015	-3.317 ± 0.014
V* V572 Pup	B0.2IVe	Y	5548261400354128768	12.74	12.42	0.118 ± 0.012	-1.455 ± 0.011	2.146 ± 0.016
V* GP Vel	B0.5Ia	N	5620657678322625920	6.87	6.74	0.496 ± 0.015	-4.822 ± 0.015	9.282 ± 0.016
4U 1907+09	O8.5Iab	N	4309225217336733824	16.35	16.82	0.233 ± 0.073	-2.749 ± 0.074	-2.455 ± 0.076
LS V +44 17	B0.2Ve	Y	252878401557369088	10.73	10.4	0.379 ± 0.015	0.101 ± 0.016	-1.186 ± 0.014
SS 188	OB	N	5541793213959987968	12.33	12.16	0.178 ± 0.01	-2.367 ± 0.009	3.177 ± 0.013
V* V479 Sct	ON6V((f)z	N	4104196427943626624	11.27	10.8	0.49 ± 0.015	7.425 ± 0.014	-8.151 ± 0.012
BD+60 73	B1Ib	N	427234969757165952	9.66	9.45	0.272 ± 0.012	-1.796 ± 0.011	-0.525 ± 0.014
EM* AS 14	B1-2 III-Ve	N	414196617287885312	11.36	11.41	0.339 ± 0.018	-2.463 ± 0.015	-0.546 ± 0.017
4U 2238+60	Be	Y	2201091578667140352	14.8	14.1	0.104 ± 0.014	-2.344 ± 0.015	-1.015 ± 0.014
EM* GGA 104	B1IIIe	Y	511220031584305536	11.42	11.22	0.328 ± 0.022	-1.029 ± 0.016	-0.082 ± 0.017
V* V662 Cas	B1Iae	Y	524924310153249920	11.14	10.52	0.196 ± 0.011	-1.243 ± 0.009	0.761 ± 0.012
2MASS J01581848+6713234	B2IVe+	Y	518990967445248256	14.43	13.69	0.133 ± 0.013	-1.198 ± 0.011	0.3 ± 0.013
HD 74194	O8.5Ib-II(f)p	N	5522306019626566528	7.55	7.45	0.443 ± 0.017	-7.465 ± 0.02	6.1 ± 0.019
HD 63666	B7IV/V	N	5489434710755238400	7.6	7.54	1.536 ± 0.021	-4.572 ± 0.027	8.53 ± 0.028
WRAY 15-793	O9.5III/Ve	Y	5336957010898124160	12.12	11.59	0.329 ± 0.011	-5.421 ± 0.012	1.37 ± 0.012
V* QV Nor	B0.2Ia:e	Y	5886085557746480000	14.5	13.16	0.128 ± 0.015	-6.711 ± 0.015	-4.111 ± 0.014
2MASS J01354986+6612433	B1Ve	Y	519352324516039680	13.31	12.46	0.167 ± 0.011	-1.626 ± 0.009	-0.027 ± 0.011
HD 49798	sdO6	N	5562023884304074240	8.287	8.22	1.92 ± 0.05	-4.162 ± 0.066	5.926 ± 0.058
HD 259440	B0pe	Y	3131822364779745536	9.12	8.88	0.54 ± 0.023	-0.026 ± 0.02	-0.428 ± 0.016
V* CI Cam	B0/2I[e]	Y	276644757710014976	11.77	10.77	0.21 ± 0.015	-0.474 ± 0.018	-0.51 ± 0.013
V* BP Cru	B1.5Iaeq	Y	6054569565614460800	10.66	9.75	0.251 ± 0.016	-5.227 ± 0.016	-2.071 ± 0.019
HD 245770	O9/B0III/Ve	Y	3441207615229815040	9.39	8.6	0.525 ± 0.023	-0.59 ± 0.031	-2.88 ± 0.016
IGR J17544-2619	O9Ib	N	4063908810076415872	12.94	11.66	0.396 ± 0.027	-0.506 ± 0.029	-0.668 ± 0.018
SS 433	A7Ib:	N	4293406612283985024	13.0	12.6	0.118 ± 0.023	-3.027 ± 0.024	-4.777 ± 0.024
HD 226868	O9.7Iabpvar	N	2059383668236814720	8.91	8.54	0.444 ± 0.015	-3.812 ± 0.015	-6.31 ± 0.017
CPD-63 2495	O9.5Ve	Y	5862299960127967488	9.98	9.63	0.443 ± 0.013	-7.093 ± 0.012	-0.342 ± 0.014

HD 63666 is a dwarf star, so it cannot be HMXB (Chevalier & Ilovaisky 1998). This system is excluded from analysis.

HD 49798 is a subdwarf system (Kudritzki & Simon 1978). This system is excluded from analysis.

IGR J17544-2619 is a supergiant system (Pellizza, Chaty & Negueruela 2006). This system is excluded from analysis.

SS 433 is a microquasar with accretion disc around the compact object (Margon 1984). The spectral type might be A and accretion does not proceed through the decretion disc. This system is excluded from analysis.

HD 226868 also known as Cyg X-1 is probably HMXB with a black hole (Orosz et al. 2011). This system is excluded from analysis.

## APPENDIX B: MAXIMUM LIKELIHOOD ANALYSIS OF THE VELOCITY DISTRIBUTION

The maximum likelihood technique was described in complete detail by Verbunt et al. (2017). Here, we briefly summarize the most important details. The joint probability to measure parallax and proper motion for a star given velocity distribution  $f_v(\vec{v}, \sigma)$  with

parameters  $\sigma$  can be written as

$$\begin{aligned}
 P(\varpi'_i, \mu'_{\alpha*i}, \mu'_{\delta i}, D, v_\alpha, v_\delta, v_r) &\propto f_D(D) f_v(\vec{v}, \sigma) \\
 &\times \exp \left[ -\frac{(1/D - \varpi'_i)^2}{2\sigma_D^2} \right] \\
 &\times \exp \left[ -\frac{(\mu_{\alpha*i}(D) + cv_\alpha/D - \mu'_{\alpha*i})^2}{2\sigma_\alpha^2} \right] \\
 &\times \exp \left[ -\frac{(\mu_{\delta i}(D) + cv_\delta/D - \mu'_{\delta i})^2}{2\sigma_\delta^2} \right]. \quad (\text{B1})
 \end{aligned}$$

Here,  $f_D(D)$  is the initial distribution for actual distance  $D$  which depends on location in the sky. Values  $\varpi'_i$ ,  $\mu'_{\alpha*i}$ , and  $\mu'_{\delta i}$  correspond to measured parallax and proper motion components for star with index  $i$  in catalogue. Values of  $\mu_{\alpha*i}(D)$  and  $\mu_{\delta i}(D)$  stay for values of proper motions which appears due to the rotation of the Galaxy at distance  $D$ . They depend on location in the sky. Coefficients  $A_\alpha$  and  $A_\delta$  allow a transformation of velocity  $v$  to component of proper motion along  $\alpha$  and  $\delta$  directions. Values  $\sigma_D$ ,  $\sigma_\alpha$ , and  $\sigma_\delta$  correspond to measurement errors for parallax and proper motion components. Coefficient  $c = 4.74$  allows a transformation of measurements to units of km s<sup>-1</sup>.

The Maxwellian velocity distribution for individual velocity components  $v_\alpha$ ,  $v_\delta$ , and  $v_r$  consists of normal distributions:

$$f_v^M(\vec{v}, \sigma) = \frac{1}{2\pi\sqrt{2\pi}\sigma^3} \exp\left(-\frac{(v_\alpha^2 + v_\delta^2 + v_r^2)}{2\sigma^2}\right). \quad (\text{B2})$$

Sum of two Maxwellians has the following form:

$$f_v^{2M}(\vec{v}, \vec{\sigma}) = w f_v^M(\vec{v}, \sigma_1) + (1-w) f_v^M(\vec{v}, \sigma_2). \quad (\text{B3})$$

The actual velocity components  $v_\alpha$ ,  $v_\delta$ ,  $v_r$  and distances  $D$  are not known. Therefore, we integrate the joint probability equation (B1) over all unknown variables:

$$P(\varpi'_i, \mu'_{\alpha*}, \mu'_{\delta i}) = \iint P(\varpi'_i, \mu'_{\alpha*}, \mu'_{\delta i}, D, \vec{v}) d^3\vec{v} dD. \quad (\text{B4})$$

Some of these integrations are performed analytically and some numerically. Practical equations used for calculations of this joint probability can be found in section 4 by Verbunt et al. (2017). As soon as exact model for velocity distribution is chosen and measured values are fixed this joint probability becomes likelihood for a model parameter:

$$L_i(\sigma) = P(\varpi'_i, \mu'_{\alpha*}, \mu'_{\delta i} | \sigma). \quad (\text{B5})$$

Working with multiple independent measurements we introduce log-likelihood:

$$\mathcal{L}(\sigma) = -2 \sum_{i=1}^N \log L_i(\sigma). \quad (\text{B6})$$

This log-likelihood is maximized further to find the best parameter  $\sigma$  or a vector of parameters  $\vec{\sigma}$ .

### APPENDIX C: BINNED VELOCITY DISTRIBUTION

First, we introduce a local polar coordinate system to describe the velocity vector:

$$v_\alpha = v \sin \psi_1 \cos \psi_2 = A_\alpha v \quad (\text{C1})$$

$$v_\delta = v \sin \psi_1 \sin \psi_2 = A_\delta v \quad (\text{C2})$$

$$v_r = v \cos \psi_1. \quad (\text{C3})$$

Our binned velocity distribution is based on individual velocity components and defined as following:

$$f_v^B(\vec{v}, \vec{m}) = m_j v^2, \quad (\text{C4})$$

if  $\Delta v j < \sqrt{v_\alpha^2 + v_\delta^2 + v_r^2} \leq \Delta v(j+1)$ . We substitute the binned velocity distribution in equation (B1). It is possible to analytically compute the most inner integral over absolute value of velocity:

$$I_{1,i,j} = m_i \int_{\Delta v i}^{\Delta v(i+1)} v^2 \exp\left[-\frac{1}{2} \left\{ \frac{A_\alpha v + D(\mu_{\alpha*}(D) - \mu'_{\alpha*})^2}{D^2 \sigma_\alpha^2} + \frac{A_\alpha v + D(\mu_{\alpha*}(D) - \mu'_{\alpha*})^2}{D^2 \sigma_\alpha^2} \right\}\right] dv. \quad (\text{C5})$$

This integral has following form:

$$I_v = m_i e^{-A_3} \int_{\Delta v i}^{\Delta v(i+1)} v^2 \exp(-A_1 v^2 - A_2 v) dv, \quad (\text{C6})$$

with the following values:

$$A_1 = \frac{c^2 A_\alpha^2}{2\sigma_\alpha^2 D^2} + \frac{c^2 A_\delta^2}{2\sigma_\delta^2 D^2} \quad (\text{C7})$$

$$A_2 = \frac{c A_\alpha (\mu_{\alpha*}(D) - \mu'_{\alpha*})}{D \sigma_\alpha^2} + \frac{c A_\delta (\mu_\delta(D) - \mu'_\delta)}{D \sigma_\delta^2} \quad (\text{C8})$$

$$A_3 = \frac{(\mu_{\alpha*}(D) - \mu'_{\alpha*})^2}{2\sigma_\alpha^2} + \frac{(\mu_\delta(D) - \mu'_\delta)^2}{2\sigma_\delta^2}. \quad (\text{C9})$$

The result of integration for equation (C5) is

$$I_v = m_i \frac{\exp(E^2 - A_3)}{A_1 \sqrt{A_1}} \left( \frac{\sqrt{\pi}}{4} [1 + 2E^2] \operatorname{erf}(\sqrt{A_1} v + E) - \frac{1}{2} [2E - 1] \exp(-(\sqrt{A_1} v + E)^2) \right) \Big|_{\Delta v i}^{\Delta v(i+1)}, \quad (\text{C10})$$

where  $E = A_2 / (2\sqrt{A_1})$ .

$$P(\varpi'_i, \mu'_{\alpha*}, \mu'_{\delta i} | m_i) = m_i \int_0^{D_{\max}} f_D(D) \times \iint I_v(D, \psi_1, \psi_2, \sigma) d\Omega dD \quad (\text{C11})$$

These integrals are computed numerically following the same procedure as described by Verbunt et al. (2017).

For each individual pulsar and for each velocity bin we can compute the likelihood as

$$L_i^{v_k, v_{k+1}} = P(\varpi'_i, \mu'_{\alpha*}, \mu'_{\delta i} | 1), \quad (\text{C12})$$

because model is described using a constant parameter  $m_k$ , we can move this parameter out of all integrations. In this case, the likelihood for a particular model but for a single pulsar can be computed as

$$L_i = \sum_{k=1}^N m_k L_i^{v_k, v_{k+1}}. \quad (\text{C13})$$

And the complete log-likelihood is computed then as

$$\mathcal{L}(\vec{m}) = -2 \sum_{i=1}^N \log \left( \sum_{k=1}^N m_k L_i^{v_k, v_{k+1}} \right). \quad (\text{C14})$$

This equation has an interesting property. Given the data it is possible to compute numerically expensive part  $L_i^{v_k, v_{k+1}}$  just once and substitute this matrix into equation (C14) together with  $m_k$  obtained from a model to get the log-likelihood.

The most natural usage of the log-likelihood equation (C14) is simply to minimize it. This exercise is, however, quite useless at the moment because a number of young pulsars and BeXs is quite small, therefore not enough pulsars are seen in each velocity bin to really constrain the velocity distribution. However, if a binary evolution model provides a model for velocity distribution of isolated radio pulsars or BeXs, equation (C14) provides a natural way to test such a model against real measurements. Given a fact that velocity bins have exactly the same extent in model and  $L_i^{v_k, v_{k+1}}$ , the relation between  $m_k$  and relative number of objects in velocity bin  $h_k$  is as following:

$$m_k = \frac{3h_k}{\Delta v^3((i+1)^3 - i^3)}. \quad (\text{C15})$$

Therefore, after we separate BeXs and isolated NS from the results of the population synthesis we prepare a histogram with a step in velocity of  $33.4 \text{ km s}^{-1}$  for normal radio pulsars and  $10 \text{ km s}^{-1}$  for BeXs and compute  $m_k$  which is further used in calculations of likelihood according to equation (C14).

#### APPENDIX D: CALCULATION OF DISTRIBUTIONS FOR SECONDARY MASS, ORBITAL PERIODS ETC

Here, we use same approach as Vinciguerra et al. (2020). Namely we assume that orbital properties do not change during the BeX stage. In this case each individual binary contributes a weight of

$$w = \frac{1}{M_{\text{tot}}} \int \text{SFH}(-t) dt, \quad (\text{D1})$$

where  $M_{\text{tot}}$  is the total stellar mass required to produce 100 000 high-mass binaries. We use the star formation history (SFH) for SMC by Rubele et al. (2015). It is a piece-wise function, so we can replace integral with a sum:

$$w = \frac{1}{M_{\text{tot}}} \sum_i \text{SFH}(-t_i) \Delta t_i. \quad (\text{D2})$$

Further we plot a histogram for each individual orbital parameter using these weights.

#### APPENDIX E: VELOCITIES OF SPECTROSCOPIC B BINARIES

It is important to check if the velocity dispersion of BeX progenitors is indeed small and comparable to  $\approx 10\text{--}15 \text{ km s}^{-1}$ . We estimate the velocity dispersion using our code (see Appendix B), parallaxes, and proper motions measured by *Gaia*. In order to perform this analysis, we use the catalogue of spectroscopic binaries by Pourbaix

et al. (2004). Our choice for the catalogue can be explained as the following: isolated B stars could be formed as a result of binary stellar evolution (walk away and ran away stars) which changes their velocities significantly. As for the spectroscopic binaries, we select those where primary did not evolve of the main sequence yet, thus these binaries probably did not interact yet. Although a small fraction of these stars could move fast because of triple and multiple stars disruption or due to OB association dissolution, this sample should provide us some information about velocity dispersion of B stars before any interaction occurs.

From this catalogue, we select only binaries where the primary is main-sequence B star without any chemical peculiarity. Chemical peculiarity could appear as a result of mass transfer or supernova explosion which could change the velocity of the system. Further we find optical counterpart for each star in the *Gaia* EDR3. After choosing only stars with good quality astrometry (relative error of parallax is less than 0.33), we have a list of 117 systems.

At the next step we run our analysis for velocities. For model with single Maxwellian distribution we obtain  $\sigma = 11 \text{ km s}^{-1}$ . For model which is a sum of two Maxwellians, we obtain a fraction of low-velocity objects  $w = 20 \pm 8$  with  $\sigma_1 = 3 \text{ km s}^{-1}$ , remaining binaries have parameter  $\sigma_2 = 11 \text{ km s}^{-1}$ . The model with two Maxwellian is more probable. Overall, we confirm that velocity dispersion of progenitors for BeXs is  $\sigma = 11 \text{ km s}^{-1}$ .

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