Two Populations of X-ray Pulsars Produced by Two Types of Supernovae

Christian Knigge[1], Malcolm J. Coe[1] & Philipp Podsiadlowski[2]

[1]*University of Southampton, School of Physics and Astronomy, Southampton SO17 1BJ, UK*
[2]*University of Oxford, Department of Physics, Oxford OX1 3RH, UK*

Two types of supernova are thought to produce the overwhelming majority of neutron stars in the Universe[1]. The first type, iron-core collapse supernovae, occurs when a high-mass star develops a degenerate iron core that exceeds the Chandrasekhar limit[2]. The second type, electron-capture supernovae, is associated with the collapse of a lower-mass oxygen-neon-magnesium core as it loses pressure support owing to the sudden capture of electrons by neon and/or magnesium nuclei[3,4]. It has hitherto been impossible to identify the two distinct families of neutron stars produced in these formation channels. Here we report that a large, well-known class of neutron-star-hosting X-ray pulsars is actually composed of two distinct sub-populations with different characteristic spin periods, orbital periods and orbital eccentricities. This class, the Be/X-ray binaries, contains neutron stars that accrete material from a more massive companion star[5]. The two sub-populations are most probably associated with the two distinct types of neutron-star-forming supernovae, with electron-capture supernovae preferentially producing system with short spin period, short orbital periods and low eccentricity. Intriguingly, the split between the two sub-populations is clearest in the distribution of the logarithm of spin period, a result that had not been predicted and which still remains to be explained.

Be/X-ray binaries (BeXs) are strong X-ray sources because their neutron stars accrete material at a relatively high rate. Their mass-losing Be-type companions are fast-rotating $8M_\odot - 18M_\odot$ main-sequence stars that are surrounded by circumstellar "decretion disks". These disks are fueled by the injection of mass and angular momentum at the stellar surface[6]. Neutron star spin periods in BeXs are typically 1 - 1,000 s, and BeX orbits are usually elliptical, with orbital periods ranging from about 10 - 1,000 d. Most of the accretion takes place during periastron passages, when the neutron star passes close to, or even through, the Be-star decretion disk.

BeXs are exceptionally abundant in the Small Magellanic Cloud (SMC), where a burst of star formation $\simeq$60 Myr ago[7] seems to have produced a large population of these systems[8,9]. In fact, the SMC contains a comparable number of BeXs to the Milky Way (MW), even though the mass ratio of the two galaxies is about 1:100. By contrast, the number of BeXs in the Large Magellanic Cloud (LMC) is broadly in line with its stellar mass content when compared with the Milky Way.

In the context of studying neutron star formation channels, it is useful to focus on well-defined, simple and "clean" populations of neutron-star-hosting systems (that is, systems in which the orbital parameters have not yet evolved since the supernova) that nevertheless span a wide

range of properties. BeXs can provide this. This is not only because the neutron stars in BeXs all have the same type of companion, but also because the accretion process itself seeeems to be universal, with the neutron star spin in or near an equilibrium state in which the magnetospheric radius of the neutron star equals the Keplerian co-rotation radius[10–12]. This conclusion is suggested empirically by the location of BeXs in the $\log(P_{\mathrm{orb}})$-$\log(P_{\mathrm{spin}})$ plane (the Corbet diagram[13]; P_{orb} and P_{spin} are the orbital and period period, respectively), where they tend to lie along a line with slope $\alpha \simeq 2$ (Fig. 1).

The correlation in Fig. 1 between P_{spin} and P_{orb} among BeXs is highly significant, but the data have large scatter. Despite this scatter, however, the 1-D projections of the data onto the $\log P_{\mathrm{orb}}$ and $\log P_{\mathrm{spin}}$ axes both suggest that the BeX population might be bimodal. More specifically, the two sub-populations suggested by the data in Fig. 1 have characteristic periods of $P_{\mathrm{orb}} \simeq 40$ d and $P_{\mathrm{spin}} \simeq 10$ s (short-period mode), and $P_{\mathrm{orb}} \simeq 100$ d and $P_{\mathrm{spin}} \simeq 200$ s (long-period mode). The bimodality of the BeX population seems to be more prominent in $\log P_{\mathrm{spin}}$ than in $\log P_{\mathrm{orb}}$. This is helpful, because there are additional BeXs, not shown in Fig. 1, for which P_{spin} is known, but P_{orb} is not. For the purpose of analysing the P_{spin} data on their own, we can therefore add these systems to the list of confirmed and probable BeXs.

Such an analysis is shown in Fig. 2. It confirms that the $\log P_{\mathrm{spin}}$ distribution of BeXs contains two distinct sub-populations with characteristic spin periods of $P_{\mathrm{spin}} \simeq 10$ s and $P_{\mathrm{spin}} \simeq 200$ s and similar dispersions of about 0.4 dex. The short-P_{spin} and long-P_{spin} sub-populations contribute about 35 and 65% to the total number, respectively. The split into these two sub-populations is highly statistically significant for the full sample, and remains significant even if the data set is divided by host galaxy. The split also remains significant if we consider only spectroscopically confirmed BeXs. Finally, the evidence for two sub-populations even remains significant if we use non-parametric statistical tests (which are less powerful, but more robust than the KMM algorithm; see the *Supplementary Information* for details).

As shown explicitly in Fig 2 (bottom panel), the double-Gaussian decomposition of the independent SMC and MW+LMC samples are consistent with each other. This makes it highly unlikely that selection effects are responsible for the observed bimodality. In any case, it is hard to conceive of a selection bias that would select specifically against BeXs with intermediate P_{spin} and/or P_{orb}. We therefore believe that the two modes of the $\log(P_{\mathrm{spin}})$ distribution correspond to physically distinct BeX sub-populations.

In principle, there are at least three ways to account for the existence of these sub-populations. First, they could correspond to two distinct neutron star spin equilibria that are accessible at all orbital periods. However, even though the bimodality is stronger in P_{spin} than in P_{orb}, the existence of the $P_{\mathrm{spin}} - P_{\mathrm{orb}}$ correlation effectively rules out this possibility. Second, P_{orb} might be time-dependent for BeXs, with the two sub-populations representing two distinct long-lived evolutionary stages. However, the time scale for stellar-wind-driven changes in P_{orb} in the BeX phase,

$\tau_{P_{orb}} \sim 100 - 1000$ Myr (refs 14, 15), is substantially longer than the maximum duration of this phase, $\tau_{BeX,max} \sim 20$ Myr, the lifetime of an $8M_\odot$ star. Thus P_{orb} evolution also cannot account for the two observed sub-populations.

The third explanation that we consider is that the two sub-populations represent two distinct BeX formation channels. The most obvious possibility, with the farthest reaching implications, is that the two channels are associated with the two distinct types of supernova event noted above. More specifically, an iron-core-collapse supernova marks the end point of the evolution of any sufficiently massive star, whereas an electron-capture supernova can occur only under highly restrictive conditions. In particular, an electron-capture supernova requires that the core reaches the critical density for electron-capture to occur, $4.5 \times 10^9 \mathrm{g \, cm^3}$ (ref. 16). These conditions might be met in the late evolution of intermediate-mass stars[3,4] (those with initial masses satisfying $8M_\odot \lesssim M_{init} \lesssim 10M_\odot$), although the relevant mass range is uncertain and may be quite small[17,18]. However, it is much easier to meet the conditions for electron-capture supernovae naturally in binary systems[17].

The outcome of electron-capture supernovae differs from that of iron-core-collapse supernovae in two fundamental ways. First, electron-capture supernovae should produce somewhat less massive neutron stars ($\lesssim 1.3 \ M_\odot$) than iron-core-collapse supernova ($1.4 \ M_\odot$) (ref. 3). Second, electron-capture supernovae are expected to impart much smaller kicks to the neutron stars they produce (average kick velocity of $\lesssim 50 \ \mathrm{kms^{-1}}$) than are iron-core-collapse supernovae ($\gtrsim 200 \ \mathrm{kms^{-1}}$)(ref. 17). In binary systems, where kicks induce orbital eccentricity, these differences could naturally give rise to two distinct sub-populations. The more conventional iron-core-collapse channel would produce high-eccentricity binaries containing high-mass neutron stars, and the electron-capture channel would produce low-eccentricity binaries containing low-mass neutron stars[16,17,19,20].

If this is the correct explanation for the two BeX populations we have discovered, they should differ not only in P_{spin} and P_{orb}, but also in their characteristic orbital eccentricities, e. These have so far been measured for only about 20 BeXs. Figure 3 shows the distribution of these systems in the $\log P_{spin} - e$ plane. Even though there are only 8 such BeXs with $P_{spin} \gtrsim 40$ s, it seems that long-P_{spin} systems have preferentially higher eccentricities than short-P_{spin} systems. The figure also shows the cumulative eccentricity distributions of the slow ($P_{spin} < 40$ s) and fast ($P_{spin} > 40$ s) BeX pulsar sub-populations. A Kolmogorov-Smirnov test shows that, despite the small number of systems for which eccentricity has been measured, the maximum difference between these distributions is marginally significant *Supplementary Information*).

The two BeX populations that we have discovered are more clearly separated in P_{spin} than in P_{orb}. Given that P_{orb} does not evolve significantly within the BeX phase, whereas P_{spin} does, P_{orb} might be expected P_{orb} to be the more faithful tracer of the formation channels. However, the P_{orb} distribution after the supernova must depend strongly on the P_{orb} distribution before the

supernova, and both low- and high-velocity kicks can in principle produce a wide range of post-supernova orbital periods[17]. This may explain why the bimodality is only marginally observed in the P_{orb} distribution. By contrast, the equilibrium spin period is expected to depend on several system parameters other than P_{porb} (refs 10-12). If any of these parameters differ systematically between BeXs produced by the iron-core-collapse and electron-capture channels, P_{spin} may be a more reliable indicator of formation channel than P_{orb}.

Our results suggest numerous avenues for further research. First, it is important to expand the data base of BeXs with reliable measurements of P_{spin}, P_{orb} and e, to confirm and further quantify our findings. Second, if short-P_{spin} BeXs are formed by low-kick-velocity electron-capture supernovae, they should have systematically smaller space velocities than long-P_{spin} BeXs. This prediction might be testable[21, 22]. Third, short-P_{spin} systems should also have systematically lower neutron star masses than long-P_{spin} systems [16, 17, 19, 20]. This prediction might also be testable[23]. Fourth, although our discovery of two populations of Be/X-ray pulsars is robust, our suggested explanation for their origin is clearly speculative and demands a fuller investigation. Intriguingly, recent binary population synthesis calculations have shown that the electron-capture supernova channel may be very efficient at forming BeXs[24]. However, whereas the same population synthesis models also suggest that the electron-capture channel accounts for the overabundance of BeXs in the SMC[24], the two BeX populations we have discovered appear to have similar relative abundances in the SMC and the Milky Way.

1. Heger, A., Fryer, C. L., Woosley, S. E., Langer, N. & Hartmann, D. H. How Massive Single Stars End Their Life. *Astrophys. J.* **591**, 288–300 (2003).

2. Woosley, S. & Janka, T. The physics of core-collapse supernovae. *Nature Physics* **1**, 147–154 (2005).

3. Nomoto, K. Evolution of 8-10 solar mass stars toward electron capture supernovae. I - Formation of electron-degenerate O + NE + MG cores. *Astrophys. J.* **277**, 791–805 (1984).

4. Nomoto, K. Evolution of 8-10 solar mass stars toward electron capture supernovae. II - Collapse of an O + NE + MG core. *Astrophys. J.* **322**, 206–214 (1987).

5. Reig, P. Be/X-ray binaries. *Astrophys. Space Sci.* **332**, 1–29 (2011).

6. Lee, U., Osaki, Y. & Saio, H. Viscous excretion discs around Be stars. *Mon. Not. R. Astron. Soc.* **250**, 432–437 (1991).

7. Harris, J. & Zaritsky, D. The Star Formation History of the Small Magellanic Cloud. *Astronom. J.* **127**, 1531–1544 (2004).

8. Haberl, F. & Pietsch, W. X-ray observations of Be/X-ray binaries in the SMC. *Astron. & Astronphys.* **414**, 667–676 (2004).

9. Coe, M. J., Edge, W. R. T., Galache, J. L. & McBride, V. A. Optical properties of Small Magellanic Cloud X-ray binaries. *Mon. Not. R. Astron. Soc.* **356**, 502–514 (2005).

10. Waters, L. B. F. M. & van Kerkwijk, M. H. The relation between orbital and spin periods in massive X-ray binaries. *Astron. & Astronphys.* **223**, 196–206 (1989).

11. Li, X. & van den Heuvel, E. P. J. On the relation between spin and orbital periods in Be/X-ray binaries. *Astron. & Astronphys.* **314**, L13–L16 (1996).

12. Liu, Q. Z., Li, X. D. & Wei, D. M. The Relation between Spin and Orbital Periods in HMXBs. In X. D. Li, V. Trimble, & Z. R. Wang (ed.) *High Energy Processes and Phenomena in Astrophysics*, vol. 214 of *IAU Symposium*, 215–217 (2003).

13. Corbet, R. H. D. Be/neutron star binaries - A relationship between orbital period and neutron star spin period. *Astron. & Astronphys.* **141**, 91–93 (1984).

14. de Jager, C., Nieuwenhuijzen, H. & van der Hucht, K. A. Mass loss rates in the Hertzsprung-Russell diagram. *Astron. & Astronphys. Supp.* **72**, 259–289 (1988).

15. Tout, C. A. & Hall, D. S. Wind driven mass transfer in interacting binary systems. *Mon. Not. R. Astron. Soc.* **253**, 9–18 (1991).

16. Podsiadlowski, P. *et al.* The double pulsar J0737-3039: testing the neutron star equation of state. *Mon. Not. R. Astron. Soc.* **361**, 1243–1249 (2005).

17. Podsiadlowski, P. *et al.* The Effects of Binary Evolution on the Dynamics of Core Collapse and Neutron Star Kicks. *Astrophys. J.* **612**, 1044–1051 (2004).

18. Poelarends, A. J. T., Herwig, F., Langer, N. & Heger, A. The Supernova Channel of Super-AGB Stars. *Astrophys. J.* **675**, 614–625 (2008).

19. van den Heuvel, E. P. J. X-Ray Binaries and Their Descendants: Binary Radio Pulsars; Evidence for Three Classes of Neutron Stars? In V. Schoenfelder, G. Lichti, & C. Winkler (ed.) *5th INTEGRAL Workshop on the INTEGRAL Universe*, vol. 552 of *ESA Special Publication*, 185–194 (2004).

20. Schwab, J., Podsiadlowski, P. & Rappaport, S. Further Evidence for the Bimodal Distribution of Neutron-star Masses. *Astrophys. J.* **719**, 722–727 (2010).

21. Coe, M. J. An estimate of the supernova kick velocities for high-mass X-ray binaries in the Small Magellanic Cloud. *Mon. Not. R. Astron. Soc.* **358**, 1379–1382 (2005).

22. Antoniou, V., Zezas, A., Hatzidimitriou, D. & Kalogera, V. Star Formation History and X-ray Binary Populations: The Case of the Small Magellanic Cloud. *Astrophys. J. Lett.* **716**, L140–L145 (2010).

23. Coe, M. J., McBride, V. A. & Corbet, R. H. D. Exploring accretion theory with X-ray binaries in the Small Magellanic Cloud. *Mon. Not. R. Astron. Soc.* **401**, 252–256 (2010).

24. Linden, T., Sepinsky, J. F., Kalogera, V. & Belczynski, K. Probing Electron-Capture Supernovae: X-Ray Binaries in Starbursts. *Astrophys. J.* **699**, 1573–1577 (2009).

25. Ashman, K. M., Bird, C. M. & Zepf, S. E. Detecting bimodality in astronomical datasets. *Astronom. J.* **108**, 2348–2361 (1994).

Author Contributions C.K. carried out the statistical analysis for this project and wrote most of the text. M.J.C. compiled the high-mass X-ray binary data set that forms the basis for our analysis and collaborated with C.K. on all aspects of the project from its inception. P.P. was instrumental to the theoretical interpretation of the results and also contributed to the final text. All authors discussed the results and their presentation.
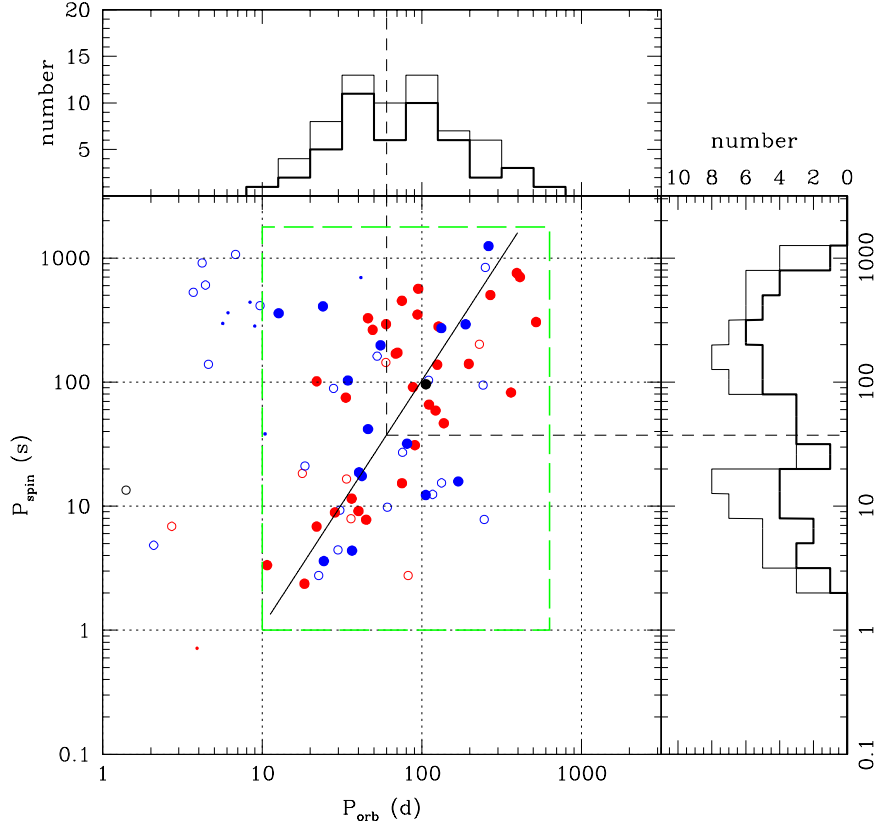
Figure 1 **The Corbet diagram for high-mass X-ray binaries.** The central panel shows $\log P_{\mathrm{orb}}$ versus $\log P_{\mathrm{spin}}$ for neutron-star-hosting high-mass X-ray binaries. Filled circles correspond to spectroscopically confirmed BeXs, small dots to confirmed non-BeX systems and open circles to candidate BeXs. There are additional confirmed and candidate BeXs for which only P_{orb} or P_{spin} is known, but these are not shown. The dashed green lines mark a selection box that conservatively includes all confirmed BeXs for which P_{orb} and P_{spin} have been measured. Candidate systems outside this box are excluded from our sample of probable BeXs. The spin and orbital periods of confirmed and probable BeX systems are correlated. The Spearman-rank correlation coefficient is $\rho = 0.49$ ($p = 3 \times 10^{-5}$; $N = 66$) for the full sample and $\rho = 0.49$ ($p = 4 \times 10^{-4}$; $N = 47$) for the confirmed systems (see *Supplementary Information* for a definition of p-values). The scatter around the correlation is $\sigma_{\log \mathrm{P_{spin}}} = 0.7$ dex relative to the best-fitting line with slope $\alpha = 2$ (solid black line). Different colours indicate different host galaxies: blue, Milky Way; red, MC; black, LMC. The histograms shown in the top and right-hand panels show the number of BeXs with spin and orbital periods in the respective ranges covered by the selection box. In each of these panels, the thick line corresponds to confirmed BeXs only, and the thin line corresponds to confirmed and probable BeXs. The vertical dashed line is drawn at $P_{\mathrm{orb}} = 60$ d, the location of the apparent dip in the $\log P_{\mathrm{orb}}$ distribution. This value of P_{orb} corresponds to $P_{\mathrm{spin}} \simeq 40$ s (horizontal dashed line), which marks a more pronounced dip in the $\log P_{\mathrm{spin}}$ distribution.
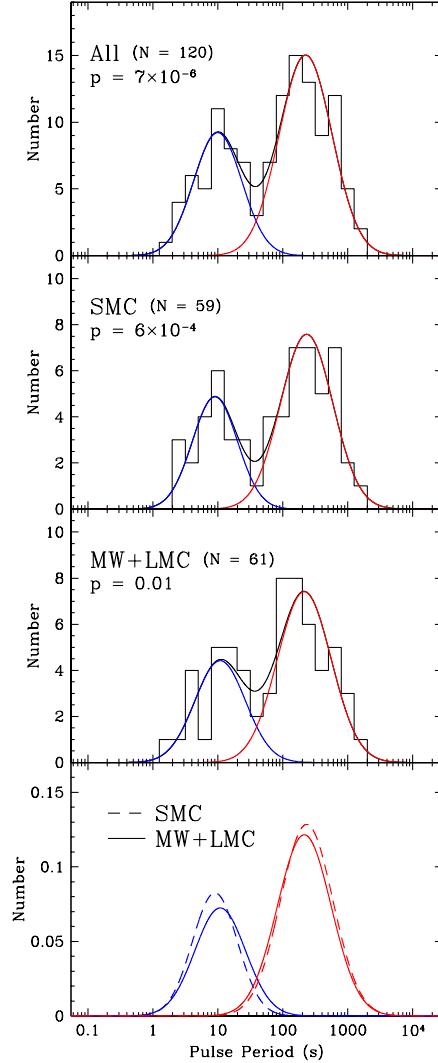
Figure 2 The $\log P_{\mathrm{spin}}$ **distribution of confirmed and probable BeXs. Top panel:** Distribution for all systems. **Middle panels:** Distribution broken down by host galaxy: SMC (second from top); Milky Way (MW)+LMC (third from top). All of these distributions are bimodal, and the double-Gaussian decomposition suggested by the KMM algorithm[25] is shown in each panel. The number of systems contributing to each observed distribution and the associated p-value provided by the algorithm are also shown. Applying the KMM test to the subset of spectroscopically confirmed systems (not shown) gives $p = 8 \times 10^{-3}$ (N = 64). **Bottom panel:** Direct comparison of the decompositions for the independent SMC and Milky Way+LMC populations, showing them to be mutually consistent. Additional details regarding the statistical evidence for the existence of distinct sub-populations in the P_{spin} data are given in *Supplementary Information*.
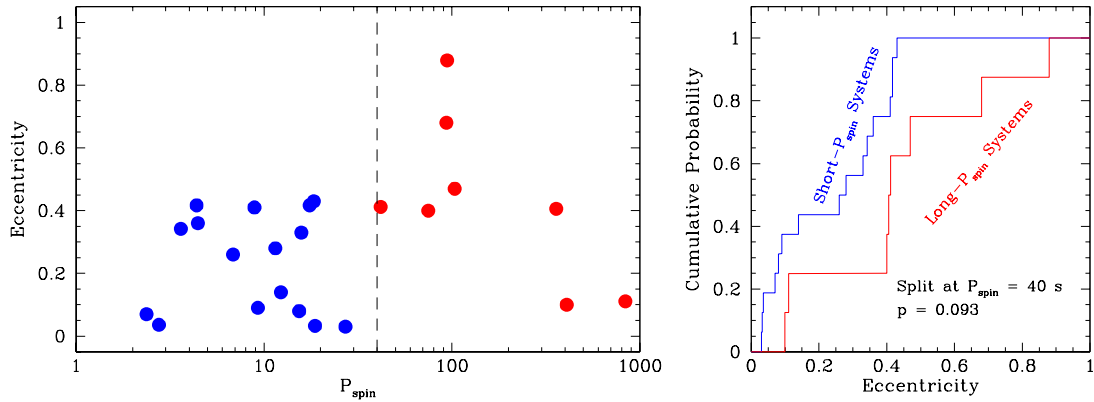
Figure 3 **The dependence of eccentricity on** P_{spin} **among BeXs. Left panel:** P_{spin} versus eccentricity for all confirmed and probable BeXs with measured spin periods and eccentricities. The vertical dashed line marks the approximate division between the short-P_{spin} and long-P_{spin} subpopulations (Figs 1 and 2). **Right panel:** Cumulative eccentricity distributions of these two subpopulations. A Kolmogorov-Smirnov test provides marginal evidence for a difference between these distributions ($p = 0.093$), with the short-P_{spin} population being characterized by lower eccentricities (see *Supplementary Information* for additional discussion).

Supplementary Information

1   Introduction

In the main paper, we present evidence for the existence of two distinct sub-populations of Be/X-ray binaries (BeXs), which are most clearly separated in the $\log P_{\rm spin}$ distribution. We argue that these two populations are most likely associated with different types of neutron star formation channels, with electron-capture supernovae primarily producing short-$P_{\rm spin}$ systems, and iron-core collapse supernovae primarily producing long-$P_{\rm spin}$ systems. This particular assignment is largely based on the eccentricity distribution of BeXs, which suggests that short-$P_{\rm spin}$ (long-$P_{\rm spin}$) systems are preferentially associated with low-e (high-e) binary orbits.

Below, we provide supplementary information for several aspects of our analysis. First, we add some historical context to our discovery of two BeX sub-populations with distinct characteristic spin periods. Second, we provide additional details regarding the significance with which these sub-populations are detected in our data. Third, we discuss some subtleties regarding the interpretation of the BeX eccentricity distribution that help to properly assess the strength of the evidence for a connection between e and $P_{\rm spin}$. We also show explicitly that there is no evidence for a connection between e and $P_{\rm orb}$.

2   Historical Context

The possible existence of a gap in the distribution between $P_{\rm spin} \simeq 10$ s and $P_{\rm spin} \simeq 100$ s was actually first suggested 30 years ago, although only $10 - 20$ systems were known at the time[26, 27]. This suggestion does not seem to have been followed-up, however, perhaps because some later discoveries fell into this gap. In any case, to the best of our knowledge, the evidence for multiple sub-populations in the spin-period distribution has not been seriously reconsidered until now.

3   The Statistical Evidence for Two Distinct Populations of BeXs

We show in the main text that the evidence for two sub-populations contributing to the $\log P_{\rm spin}$ distribution of BeXs is highly statistically significant according to the standard KMM test[25]. The test statistic used by KMM is the likelihood ratio between the best single-Gaussian representation and the best double-Gaussian representation of the data. For the latter, we always find that the variances of the two Gaussians are consistent with each other. In our application of KMM, we thus adopt equal variances in the double Gaussian representation and then calculate the p-values robustly from bootstrap simulations[28]. Here and throughout this paper, the quoted p-values have the usual statistical meaning, i.e. they represent the probability of obtaining a test statistic at least as extreme as the observed one, under the assumption that the null hypothesis is correct. In the

case of KMM, the null hypothesis is that the data are drawn from a single Gaussian.

Figure 2 in the main text shows that the best double-Gaussian representation actually provides quite a good match to the $\log P_{\rm spin}$ data. It is nevertheless interesting to ask if we can find evidence for multiple sub-populations in our data set even without the assumption of normality. This can be checked by non-parametric methods, such as the "bandwidth test"[29]. The standard bandwidth test is a non-parametric test for multi-modality based on bootstrapping from a kernel density estimate of the underlying probability distribution function (PDF). The statistic used in the test is the smallest kernel width that results in a unimodal estimate of the PDF. The standard bandwidth test is a strict test of modality, in the sense that the null hypothesis is only that the underlying PDF is unimodal, with no additional assumptions concerning its shape. However, a unimodal distribution with an obvious "shoulder", for example, may still be considered to provide strong evidence for multiple sub-populations. We can therefore also define an alternative test, in which the statistic used is the smallest kernel width that results in a kernel density estimate of the PDF that has only two inflection points[30]. Being entirely non-parametric, both of these tests are extremely robust. However, they are also much less powerful than KMM, mainly in the sense that their null hypotheses are so general that they are hard to rule out. The resulting p-values, which can be calibrated via Monte-Carlo simulations[31], are thus always larger than those provided by KMM.

We have applied both versions of the bandwidth test to the $P_{\rm spin}$ data for confirmed and probable BeXs. For the standard (modality) bandwidth test, we first apply a skew-minimising Box-Cox transformation to the $\log P_{\rm spin}$ data, because modes associated with distinct sub-populations can be either emphasised or suppressed by certain transformations[32]. A transformation to symmetry is expected to make any inherent bimodality (strictly speaking: "bimodalizability") easier to detect[33]. For our BeX data set, the classic bandwidth test then rejects the null hypothesis of unimodality with $p = 9 \times 10^{-3}$ (application of this test directly to the $\log P_{\rm spin}$ data would give $p = 0.04$). Applying the alternative (inflection point) bandwidth test directly to the $\log P_{\rm spin}$ data rejects the null hypothesis that there are only two inflection points in the underlying distribution with $p = 1 \times 10^{-3}$ (application of this test to the deskewed data would give $p = 5 \times 10^{-3}$). Given the robust non-parametric nature of these tests, we consider these results to be strong evidence for the existence of two distinct BeX sub-populations.

4 The Eccentricity Distribution of BeXs

Figure 3 in the main text shows that there is marginally significant evidence that the eccentricity distribution of long-$P_{\rm spin}$ BeXs is different from that of short-$P_{\rm spin}$ systems. Some care has to be taken in interpreting this result.

First, tidal circularization will always tend to reduce the eccentricity of close binary systems. However, the orbital periods within our BeX sample ($P_{\rm orb} > 10$ d) are too long for this mechanism to be effective[34]. Second, other things being equal, a given supernova kick velocity will induce

larger orbital eccentricities in wide, long-period, binaries, since such systems are more weakly bound. Given that P_{spin} is correlated with P_{orb} in BeXs, this could produce a $P_{\text{spin}} - e$ correlation in the observed sense.

Does this latter bias affect our analysis? Supplementary Figure 1 shows the P_{orb} versus e distribution for our BeX sample. Superposed on this, we also show the *predicted* location[35], under the assumption that they all receive the large kicks typically associated with conventional iron-core-collapse ($\langle v_{\text{kick}} \rangle \simeq 450$ km s^{-1}). Two key conclusions may be drawn immediately from this plot. First, the observed distribution is clearly inconsistent with being drawn from a single, high-kick population. In fact, the low eccentricities ($e \lesssim 0.4$) seen in many BeXs with $P_{\text{orb}} \gtrsim 20$ d cannot be explained by large kicks at all. A possible connection between such low-eccentricity HMXBs[36] and low-kick electron capture supernovae has indeed been suggested before[17,19]. Second, the $P_{\text{orb}} - e$ plane does not separate the two BeX sub-populations as cleanly as the $P_{\text{spin}} - e$ plane. Indeed, if we split the BeX sample into short-period ($P_{\text{orb}} < 60$ d) and long-period ($P_{\text{orb}} > 60$ d) systems, their cumulative eccentricity distributions do not differ significantly ($p = 0.35$). Both conclusions indicate that the link between spin period and eccentricity suggested in the main paper is not simply due to a P_{orb}-dependent bias.

Finally, there are two obvious outliers from the trend that long-P_{spin} systems have preferentially high eccentricities (see Figure 3 in the main text). These systems are X Per and 1A 1118-616, which actually have the longest spin periods ($P_{\text{spin}} = 837.7$ s and 407.7 s, respectively), but $e \simeq 0.1$. Both of these object are known to be highly unusual within the BeX class. For example, X Per does not show the transient X-ray bursts associated with periastron passages that are the hallmark of almost all BeXs, while 1A 1118-616 is one of the furthest systems from the standard $P_{\text{orb}} - P_{\text{spin}}$ correlation (Figure 1 of the main text). Indeed, in both systems, the dominant accretion mode probably differs from that which establishes the spin equilibrium in other BeXs[37,38]. This may explain why P_{spin} is *not* a valid tracer of formation channel in these two objects. Also, since 1A 1118-616 has a surprisingly short orbital period considering its slow spin ($P_{\text{orb}} = 24$ d[38]), the low eccentricity of this system is not actually in terrible conflict with a standard (high) kick scenario[35]. If X Per and 1A 1118-616 are excluded from the eccentricity distribution of long-P_{spin} systems, the difference to the short-P_{spin} systems becomes highly significant ($p = 0.006$). *A posteriori* outlier removal is, of course, a dangerous statistical practice, so this p-value cannot be taken quite at face value. Nevertheless, it is also interesting to note that all other high-P_{spin} systems in Figure 3 do, in fact, lie in the region predicted for systems produced via high-kick supernovae.
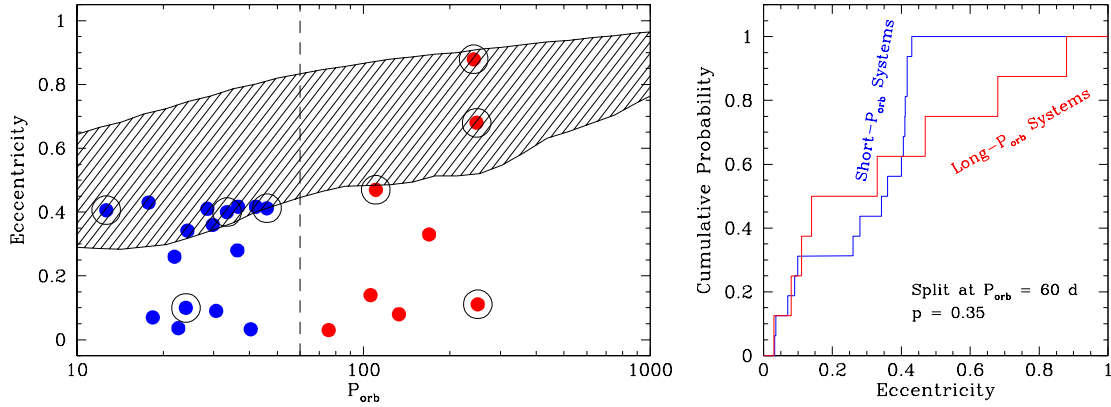
Supplementary Figure 1 The dependence of eccentricity on P_{orb} among BeXs. The left panel shows P_{orb} versus eccentricity for all confirmed and probable BeX systems with measured spin and orbital periods and eccentricities. Long-P_{spin} systems with $P_{spin} > 40$ s are additionally marked with open circles. The vertical dashed line marks the approximate division between the short-P_{orb} and long-P_{orb} sub-populations (see Figure 1 in the main text). The shaded area shows the region predicted to contain 60% of systems by population synthesis calculations adopting a kick velocity distribution appropriate for iron-core-collapse with mean $\langle v_{\mathrm{kick}} \rangle \simeq 450$ km s^{-1}[35]. The top right panel shows the cumulative eccentricity distributions of these two sub-samples. A Kolmogorov-Smirnov test provides no evidence for a significant difference between these distributions ($p = 0.35$).

26. Rappaport, S. & Joss, P. C. Binary X-ray pulsars. *Nature* **266**, 123–125 (1977).

27. Bhatt, H. C. The pulse-period distribution of binary X-ray pulsars. *Astrophys. Space Sci.* **81**, 379–385 (1982).

28. Muratov, A. L. & Gnedin, O. Y. Modeling the Metallicity Distribution of Globular Clusters. *Astrophys. J.* **718**, 1266–1288 (2010). `1002.1325`.

29. Silverman, B. W. Using kernel density estimates to investigate multimodality. *Journal of the Royal Statistical Society, Series B* **43**, 97–99 (1981).

30. González-Manteiga, W. & Cuevas, A. Data-driven smoothing based on convexity properties. In G. Roussas (ed.) *Nonparametric functional estimation and related topics*, vol. 335 of *NATO ASI Series C*, 225–240 (1991).

31. Hall, P. & York, M. On the calibration of Silverman's test for multimodality. *Statistica Sinica* **11**, 515–536 (2001).

32. Knigge, C., Scaringi, S., Goad, M. R. & Cottis, C. E. The intrinsic fraction of broad-absorption line quasars. *Mon. Not. R. Astron. Soc.* **386**, 1426–1435 (2008). `0802.3697`.

33. Wyszomirski, T. Detecting and displaying size bimodality: Kurtosis, skewness and bimodalizable distributions. *Journal of Theoretical Biology* **158**, 109–128 (1992).

34. North, P. & Zahn, J. Circularization in B-type eclipsing binariesin the Magellanic Clouds. *Astron. & Astronphys.* **405**, 677–684 (2003).

35. Brandt, N. & Podsiadlowski, P. The effects of high-velocity supernova kicks on the orbital properties and sky distributions of neutron-star binaries. *Mon. Not. R. Astron. Soc.* **274**, 461–484 (1995).

36. Pfahl, E., Rappaport, S., Podsiadlowski, P. & Spruit, H. A New Class of High-Mass X-Ray Binaries: Implications for Core Collapse and Neutron Star Recoil. *Astrophys. J.* **574**, 364–376 (2002).

37. Delgado-Martí, H., Levine, A. M., Pfahl, E. & Rappaport, S. A. The Orbit of X Persei and Its Neutron Star Companion. *Astrophys. J.* **546**, 455–468 (2001). `arXiv:astro-ph/0004258`.

38. Staubert, R. *et al.* Finding a 24-day orbital period for the X-ray binary 1A 1118-616. *Astron. & Astronphys.* **527**, A7+ (2011). `1012.2459`.