Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **530,** 2972–2993 (2024) https://doi.org/10.1093/mnras/stae711
Advance Access publication 2024 March 8

The epoch of the Milky Way's bar formation: dynamical modelling of Mira variables in the nuclear stellar disc

Jason L. Sanders [1]★ Daisuke Kawata [2] Noriyuki Matsunaga,[3] Mattia C. Sormani [4] Leigh C. Smith,[5] Dante Minniti [6,7,8] and Ortwin Gerhard [9]

[1]*Department of Physics and Astronomy, University College London, London WC1E 6BT, UK*
[2]*Mullard Space Science Laboratory, University College London, Holmbury St. Mary, Dorking, Surrey RH5 6NT, UK*
[3]*Department of Astronomy, School of Science, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*
[4]*School of Mathematics & Physics, University of Surrey, Guildford GU2 7XH, UK*
[5]*Institute of Astronomy, University of Cambridge, Madingley Rise, Cambridge CB3 0HA, UK*
[6]*Departamento de Ciencias Físicas, Facultad de Ciencias Exactas, Universidad Andrés Bello, Fernández Concha 700, 7591538 Las Condes, Santiago, Chile*
[7]*Vatican Observatory, Vatican City State, V-00120, Italy*
[8]*Department of Physics, Universidade Federal de Santa Catarina, Florianopolis, 88040-900, Brazil*
[9]*Max-Planck Institut für extraterrestrische Physik, Giessenbachstraße, D-85748 Garching, Germany*

Accepted 2024 March 6. Received 2024 February 15; in original form 2023 October 23

ABSTRACT

A key event in the history of the Milky Way is the formation of the bar. This event affects the subsequent structural and dynamical evolution of the entire Galaxy. When the bar formed, gas was likely rapidly funnelled to the centre of the Galaxy settling in a star-forming nuclear disc. The Milky Way bar formation can then be dated by considering the age distribution of the oldest stars in the formed nuclear stellar disc. In this highly obscured and crowded region, reliable age tracers are limited, but bright, high-amplitude Mira variables make useful age indicators as they follow a period–age relation. We fit dynamical models to the proper motions of a sample of Mira variables in the Milky Way's nuclear stellar disc region. Weak evidence for inside-out growth and both radial and vertical dynamical heating with time of the nuclear stellar disc is presented, suggesting that the nuclear stellar disc is dynamically well-mixed. Furthermore, for Mira variables around a ∼350-d period, there is a clear transition from nuclear stellar disc-dominated kinematics to background bar-bulge-dominated kinematics. Using a Mira variable period–age relation calibrated in the solar neighbourhood, this suggests the nuclear stellar disc formed in a significant burst in star formation (8 ± 1) Gyr ago, although the data are also weakly consistent with a more gradual formation of the nuclear stellar disc at even earlier epochs. This implies a relatively early formation time for the Milky Way bar ($\gtrsim 8$ Gyr), which has implications for the growth and state of the young Milky Way and its subsequent history.

Key words: stars: AGB – stars: variables: general – Galaxy: evolution – Galaxy: formation – Galaxy: kinematics and dynamics – Galaxy: nucleus.

1 INTRODUCTION

With the advance of large astrometric, photometric, and spectroscopic stellar surveys, a detailed picture and understanding of the formation and evolution of the Milky Way is being built up. About 13 Gyr ago, the proto-disc of the Milky Way likely formed from an early turbulent, merger-dominated phase in which it was rapidly metal-enriched and spun up (Arentsen et al. 2020; Belokurov & Kravtsov 2022; Rix et al. 2022). Since then the Milky Way has undergone a series of merger events: about 10 Gyr ago (Belokurov et al. 2020; Bonaca et al. 2020), the Gaia-Sausage-Enceladus galaxy (Belokurov et al. 2018; Helmi et al. 2018) merged with the Milky Way, and today the Sagittarius dwarf galaxy and Magellanic system are merging with the Milky Way in processes that have likely produced the non-equilibrium structure and features we see throughout the disc and the halo (Antoja et al. 2018; Petersen & Peñarrubia 2021; Drimmel et al. 2023; Koposov et al. 2023).

A further significant event must be placed within this chronology of the Galaxy: the formation epoch of the Galactic bar. It is generally accepted that the Milky Way has an X-shaped/boxy-peanut bar-bulge that evolved from a pre-existing disc (Blitz & Spergel 1991; Wegg & Gerhard 2013). This likely occurred in two distinct stages: a flattish in-plane bar was initially produced through an in-plane disc instability (Hohl 1971; Ostriker & Peebles 1973; Sparke & Sellwood 1987) that may have been triggered by an external perturbation from a satellite galaxy (Noguchi 1987; Gerin, Combes & Athanassoula 1990; Łokas et al. 2014; Łokas 2021), and then the central parts fattened into the observed X-shaped/boxy-peanut bar-bulge either through a violent buckling event (Raha et al. 1991), or a more gradual resonant thickening process (Combes & Sanders 1981; Combes et al. 1990; Quillen et al. 2014; Sellwood & Gerhard 2020).

★ E-mail: jason.sanders@ucl.ac.uk

Since the epoch of bar formation, the Milky Way bar has likely had a significant dynamical and evolutionary impact on the Galaxy in several ways:

(i) The barred potential has likely led to significant restructuring of the Galactic disc. This is most notable through the presence of moving groups in the solar neighbourhood that have been related to various resonances in the bar (Dehnen 2000; Antoja et al. 2014; Monari et al. 2017; Pérez-Villegas et al. 2017; Hunt & Bovy 2018; Monari et al. 2019; Chiba, Friske & Schönrich 2021), but it is also likely that the resonant overlap between the bar and spirals has led to enhanced radial migration (Quillen 2003; Minchev & Famaey 2010) that has shuffled stars radially in the disc (Frankel et al. 2020);

(ii) The dark matter halo has likely exchanged angular momentum with the Galactic bar through a dynamical friction process causing the bar to slow (Tremaine & Weinberg 1984; Athanassoula 2003). This is an excellent probe of the fundamental properties of dark matter and is suggested by the morphology and metallicity distribution of the Hercules moving group in the solar neighbourhood (Chiba & Schönrich 2021). However, it is also known that both gas and tidal interactions can cause the bar to spin up again (e.g. Łokas et al. 2014; Beane et al. 2022);

(iii) Finally, the presence of a bar significantly alters the structure of gas orbits within a Galaxy (Binney et al. 1991; Fux 1999; Sormani, Binney & Magorrian 2015a, b; Li et al. 2022), and in particular can cause gas to be directly funnelled towards the central regions of the Galaxy leading to the formation of nuclear discs (Bournaud & Combes 2002; Sormani & Barnes 2019; Hatchfield et al. 2021). It is therefore likely that bar formation is somehow linked to active galactic nucleus activity and black hole growth within the Milky Way but this is currently unclear (Shlosman, Frank & Begelman 1989; Sellwood 2014; Emsellem et al. 2015).

In addition to these evolutionary effects, whether a bar forms or not is affected by a galaxy's baryon dominance, gas fraction, and stellar/gas velocity dispersion (Ostriker & Peebles 1973; Hohl 1976; Athanassoula & Misiriotis 2002; Athanassoula 2008; Fragkoudi et al. 2021; Bland-Hawthorn et al. 2023; Romeo, Agertz & Renaud 2023). In this way, a measurement of the bar age gives a direct probe of the morphological properties of the Milky Way and its relative dark matter/baryon fraction at a fixed epoch, allowing for detailed testing of the cosmological picture (e.g. Sheth et al. 2012).

1.1 When do bars form in galaxies?

To understand these phenomena and effects further, it is crucial to date the formation of the Milky Way bar, making it a key goal of Galactic archaeology. However, this task is not simple, as, importantly, the dynamical age of a galaxy's bar can be different from the age of the constituent stars: stars can be born in a precursor disc from which the bar forms, stars can form in the bar region or a bar can capture stars that formed after bar formation as it grows and evolves. From observing galaxies across redshift, constraints can be placed on the occurrence of bars in galaxies and the maximum redshift at which galaxies host bars. This information gives the typical age of bars in galaxies other than the Milky Way. Locally, approximately two-thirds of spiral galaxies are observed to host bars (Eskridge et al. 2000; Erwin 2018). Studies using *Hubble Space Telescope* (*HST*) data (Sheth et al. 2008; Melvin et al. 2014; Simmons et al. 2014) have found that the fraction of *strongly* barred galaxies decreases with redshift. However, the rate of this decrease may in part arise from the detectability of bars at high redshift with *HST* as recent *JWST* analyses (Chen et al. 2022; Ferreira et al. 2022; Guo et al. 2023; Jacobs et al. 2023) have found several examples of barred galaxies at look-back times between 8 and 11 Gyr and a recent population-level analysis (Le Conte et al. 2023) finds a factor 3–4 higher bar fractions at $z \gtrsim 1$ than the earlier *HST* results. Such observations suggest that the conditions are right for bars to form quite early on in the typical history of galaxies (Bland-Hawthorn et al. 2023), as indicated by the analysis of the bar fractions in cosmological simulations (Fragkoudi et al. 2020; Rosas-Guevara et al. 2022). However, it should be noted that the results on bars from cosmological simulations are likely still in a state of flux as there are tensions with the data in the distribution of bar lengths and/or pattern speeds they predict (Zhao et al. 2020; Fragkoudi et al. 2021; Roshan et al. 2021; Frankel et al. 2022). Once formed, simulations indicate that bars are generically long-lived (Kraljic, Bournaud & Martig 2012; Sellwood 2014) but can be weakened or even destroyed by mergers (Ghosh et al. 2021; Bi, Shlosman & Romano-Díaz 2022; Cavanagh et al. 2022).

There are several estimates of the epoch of bar formation in the Milky Way. The Galactic bar contains predominantly old stars of around ~ 10 Gyr (Bernard et al. 2018; Bovy et al. 2019; Surot et al. 2019; Grady, Belokurov & Evans 2020; Hasselquist et al. 2020; Savino et al. 2020) approximately supporting the idea that the Milky Way bar is old. However, there is the suggestion that the bar is not solely composed of old stars: using microlensed dwarf stars towards the Galactic bar-bulge, Bensby et al. (2013, 2017) argued that there is ∼20 per cent of stars younger than ~ 5 Gyr old, which is approximately corroborated by the tail towards younger ages observed in *HST* colour–magnitude diagram modelling of the Galactic bar-bulge (Bernard et al. 2018). This minority of stars may be linked with in-bar star formation (e.g. Anderson et al. 2020) or capture. Cole & Weinberg (2002) photometrically identified a significant population of carbon stars in the Galactic bar-bulge using 2MASS. As intrinsic carbon stars only form from low-metallicity and/or high-mass progenitors, Cole & Weinberg (2002) argued the bar might be as young as 3–6 Gyr old. However, the photometric selection employed by Catchpole et al. 2016; Matsunaga et al. 2017), possibly pushing this bar age estimate higher. In fact, it seems that the low number of genuine carbon-rich AGB stars in the bar-bulge is consistent with their extrinsic production through binary evolution (Sanders & Matsunaga 2023) which would push the bar age estimate from Cole & Weinberg (2002) higher still. Recently, Tepper-Garcia et al. (2021) presented a simulation of a Milky Way-like galaxy that they argued matches the Milky Way well in several aspects and then supports the idea that the bar formed 3–4 Gyr ago, but similarly Buck et al. (2018) presented a Milky Way analogue simulation with a bar formation time of ~ 8 Gyr. Using the potential dynamical impact of the bar, Khoperskov et al. (2019) presented a scenario for the formation of the 'Gaia snail', the local spiral in the vertical displacement versus vertical velocity of the stars (Antoja et al. 2018), in which the perturbation arises from the buckling of the bar. Although there are alternatives to their presented scenario (such as a satellite perturbation), this suggests a bar buckling event within the last ~ 3 Gyr, and if bars buckle promptly, possibly then a more recent bar formation. Wylie, Clarke & Gerhard (2022) used APOGEE data to identify an inner $\sim$ 4–6 kpc ring of metal-rich stars within the bar's corotation radius. Such rings form after the bar, and analysing similar barred galaxies from the Auriga cosmological simulations, Fragkoudi et al. (2020) argued that the time at which the star formation in the bar starts quenching, while star formation in the ring continues, can be used to estimate a lower limit to the age of the bar. Using the inferred age distributions of the APOGEE bar and ring stars, Wylie et al. (2022) thus estimated a lower limit on the Milky Way's bar formation epoch of 7 Gyr ago. Finally, Nepal et al.

(2024) have recently argued from the age–metallicity distribution of turn-off/subgiant stars in the solar neighbourhood that a burst of star formation around 3 Gyr ago accompanied by a subsequent decline in migration of younger metal-rich stars to the solar neighbourhood indicate the Milky Way bar is possibly 3–4 Gyr old.

1.2 Bar formation and nuclear stellar discs

Another very different approach to dating the epoch of bar formation is to not look at the stars in the bar at all but instead to look at stars in structures that probably only started forming when bars have formed: nuclear stellar discs (NSDs, Erwin & Sparke 2002; Pizzella et al. 2002). Controlled simulations of disc galaxies (e.g. Athanassoula 1992; Cole et al. 2014; Seo et al. 2019; Baba & Kawata 2020) have demonstrated that when a bar forms, gas funnels along bar lanes towards the centre of the galaxy where it can settle in a nuclear ring (Sormani, Sobacchi & Sanders 2023) and begin forming stars that constitute a nuclear disc. Based on these theoretical ideas, the MUSE-TIMER project (Gadotti et al. 2015, 2018, 2020; De Sá-Freitas et al. 2023a, b) has measured the age and abundance distributions of the nuclear regions of barred galaxies, finding that nuclear discs typically have lower velocity dispersions, are more metal-rich and are younger than the surrounding bar, with the indication of inside-out growth of the NSDs (Bittner et al. 2020).

This approach to measuring the formation age of a bar was highlighted specifically for dating the Milky Way bar formation epoch by Baba & Kawata (2020). They demonstrated from controlled simulations that within $\sim$ 1 Gyr of a bar forming there is an associated $\sim$ 1 Gyr long central star formation burst forming an NSD and highlighted the importance of proper motion data for isolating the NSD population. The Milky Way is known to host an NSD of around $10^9\ M_\odot$ with a scale length of $\sim$ 75 pc and a scale height of $\sim$ 25 pc that rotates at approximately 100 km s^{-1} (Catchpole, Whitelock & Glass 1990; Lindqvist, Habing & Winnberg 1992; Launhardt, Zylka & Mezger 2002; Nishiyama et al. 2013; Matsunaga et al. 2015; Schönrich, Aumer & Sale 2015; Gallego-Cano et al. 2020; Schultheis et al. 2021; Shahzamanian et al. 2022; Sormani et al. 2022a). The nuclear star cluster (NSC, Neumayer, Seth & Böker 2020) sits at the centre of the NSD and has an effective radius of $\sim$ 5 pc. Recently, Sormani et al. (2022a) fitted self-consistent axisymmetric dynamical models to a combination of spectroscopic and proper motion data (Smith et al. 2018; Fritz et al. 2021) for $\sim$3500 giant stars across the NSD. The assumption of axisymmetry gives very good fits to the data, although detecting whether the NSD is genuinely an axisymmetric disc or a nuclear bar is challenging (Alard 2001; Gerhard & Martinez-Valpuesta 2012; Gonzalez et al. 2012). There have been several studies of the Milky Way NSD star formation history, beginning with Figer et al. (2004) who found a relatively continuous star formation history using Hubble Space Telescope. More recently, the GALACTICNUCLEUS survey (Nogueras-Lara et al. 2019) has provided deep JHK_s photometry from which the morphology of the NSD giant branch luminosity function can be analysed. These data present a richer picture suggesting that the star formation history of the central NSD is more bursty, predominantly forming $\gtrsim$ 7–8 Gyr ago (consistent with the presence of RR Lyrae stars in this region, Minniti et al. 2016) with a more recent star formation burst in the last $\sim$ 1 Gyr (Matsunaga et al. 2015; Nogueras-Lara et al. 2020; Schödel et al. 2023). However, the outer parts of the NSD have evidence of more significant intermediate-age populations of 2–7 Gyr (Nogueras-Lara, Schödel & Neumayer 2022; Nogueras-Lara et al. 2023b), which supports an inside-out growth of the NSD in the Milky Way.

1.3 Mira variables as age tracers of the nuclear stellar disc

One useful *resolved* age tracer for the NSD region is Mira variable stars. These are high-amplitude (typically $\Delta K_s \gtrsim 0.4$) asymptotic giant branch stars with periods of 80–1000 d thermally pulsing in the fundamental mode (Wood 2015) driven through some convective mechanism (Freytag, Liljegren & Höfner 2017; Xiong, Deng & Zhang 2018). Primarily from observations of the LMC, they are known to follow period–luminosity relations (Glass & Evans 1981; Feast et al. 1989; Groenewegen 2004; Ita et al. 2004; Fraser, Hawley & Cook 2008; Riebel et al. 2010; Ita & Matsunaga 2011; Yuan et al. 2017a, b; Bhardwaj et al. 2019; Iwanek, Soszyński & Kozłowski 2021; Sanders 2023) as stars of a given mass only begin pulsating in the fundamental mode at a narrow range of radii (Trabucchi et al. 2019). For a similar reason, there is a theoretical expectation that Mira variables follow a period–age relation (Wyatt & Cahn 1983; Feast & Whitelock 1987; Eggen 1998; Trabucchi & Mowlavi 2022). Such a relation has been known observationally for some time from the related correlations between the periods of the Mira variables and their kinematics (Merrill 1923) and between the periods and the scale-heights of the populations (Feast 1963). As hotter kinematics are typically associated with older populations, these correlations are typically interpreted as manifestations of a period–age relation, which is further observationally validated through the limited number of cluster Mira variables (Grady, Belokurov & Evans 2019; Marigo et al. 2022; Zhang & Sanders 2023). The solar neighbourhood correlations led to a number of empirical calibrations (Feast & Whitelock 1987, 2000, 2014; Feast, Whitelock & Menzies 2006; Catchpole et al. 2016; López-Corredoira 2017; Grady et al. 2020; Nikzat et al. 2022). Recently, Zhang & Sanders (2023) fitted dynamical models to Mira variables from *Gaia* DR3 and by relating the radial and vertical dispersions with period to the dispersions as a function of age from the analysis of main-sequence turn-off and sub-giant stars from Yu & Liu (2018), they derived a period–age relation that agreed well with both previous analyses and the results from the cluster members.

The calibrated period–age relations demonstrate that Mira variables span ages from $\sim$ 1–10 Gyr making them excellent probes of intermediate age populations and so they have found significant use as an archaeological probe of the Milky Way (Catchpole et al. 2016; Grady et al. 2019, 2020; Semczuk et al. 2022; Zhang & Sanders 2023). In addition to this, the brightness and low contamination of Mira variables also make them ideal age tracers of the NSD. Recently, Sanders et al. (2022b) presented a sample of 1782 Mira variables in the NSD region extracted from the Vista Variables in Via Lactea (VVV) survey. This catalogue built on previous work searching for Mira variables by targeting OH/IR maser stars (Blommaert et al. 1998; Wood, Habing & McGregor 1998, which are biased towards longer period sources) and the broader variable star searches of Glass et al. (2001) and Matsunaga et al. (2009) that only targeted smaller on-sky areas close to the Galactic Centre. In addition to providing a more panoramic view of the entire NSD region (the inner 3×3 deg^2), the catalogue of Sanders et al. (2022b) contained proper motions for all Mira variable candidates from the VIRAC reduction of VVV (Smith et al. 2018). Although NSD membership can be assessed using Mira variable period–luminosity relations, these rely on a solid understanding of the extinction (both interstellar and circumstellar) and population effects, making kinematic membership significantly more reliable. With this sample, we are in a position to kinematically explore and identify the onset of star formation in the NSD and hence the epoch of bar formation in the Milky Way.

In this work, we present dynamical models of the Mira variable population in the NSD region from Sanders et al. (2022b). As contamination from the foreground disc and bar is significant (Sormani et al. 2022a), the models consist of two components: an axisymmetric model for the NSD population and a barred 'background' *N*-body model from Portail et al. (2017). The NSD models and their relative weight compared to the background are parametrized by the period/age of the Mira variable population. Furthermore, we provide a detailed discussion and implementation of the selection function of the sample, which preferentially biases us towards observing fainter background objects. The paper is structured as follows: in Section 2 we describe the data set used, whilst in Section 3 we lay out the components of the model providing a detailed discussion of the selection function of the sample in Appendix A. The results of our modelling are presented in Section 4 and our conclusions and a discussion of the implications of our results are described in Section 5.

2 DATA AND INITIAL MODELLING CONSIDERATIONS

The Mira variable sample is taken from Sanders et al. (2022b). This sample was extracted from an intermediate version of the VIRAC2 reduction of the VVV data (Smith et al. 2018). VIRAC2 is the second iteration of the VVV (Minniti et al. 2010; Saito et al. 2012) InfraRed Astrometric Catalogue constructed from fitting five-parameter astrometric solutions to point-source-function catalogues of the VVV and VVVX K_s epoch data over a baseline of around 10 yr that was photometrically calibrated to 2MASS data and astrometrically calibrated to the *Gaia* EDR3 reference frame (e.g. Clarke et al. 2019; Sanders et al. 2019). In this way, absolute proper motions are available despite the lack of background astrometric calibrators in this highly crowded and extincted region. Likely Mira variables were first identified in the inner $3 \times 3\,\mathrm{deg}^2$ of the Galaxy using basic cuts on the K_s amplitude inspired by the study of Matsunaga et al. (2009), and then more detailed light curve modelling of the selected stars was performed to measure periods allowing selection using the period – amplitude diagram and some period – Wesenheit index planes. Any Mira variable from Matsunaga et al. (2009) missed by the pipeline but still in the VIRAC catalogue is also included. There are 1782 Mira variable candidates in the sample, all of which have proper motion measurements. For some of the Mira variable candidates, radial velocities are available from maser observations (Engels & Bunzel 2015 for OH masers and Messineo et al. 2002, 2004, Deguchi et al. 2004, and Fujii et al. 2006 for SiO masers).

The median proper motion uncertainty in each component is $\sim 0.5\,\mathrm{mas\,yr^{-1}}$ corresponding to $\sim 20\,\mathrm{km\,s^{-1}}$ at the Galactic centre distance of 8.275 kpc (Gravity Collaboration 2021). The full distribution of the uncertainties is shown in Fig. 1, which demonstrates the deterioration of the astrometry for the brighter stars due to saturation effects. The unit weight error calculated from the residuals to the astrometric fit increases from <1 for stars with $K_s \gtrsim 11.5$ to ∼2 for stars with $K_s \sim 10.5$. This suggests that the astrometric centroid errors are underestimated for stars in the saturated regime and possibly the resulting proper motion uncertainties are underestimated. In addition to this, Luna et al. (2023) compared the VIRAC2 and *HST* proper motions in a few available fields finding that possibly VIRAC2 proper motion uncertainties are underestimated by ∼10 per cent. From the model fits of Sormani et al. (2022a) the expected vertical velocity dispersion of the NSD is around $50\,\mathrm{km\,s^{-1}}$ (although note it decreases with radius and the velocity distributions are cuspier than a Gaussian

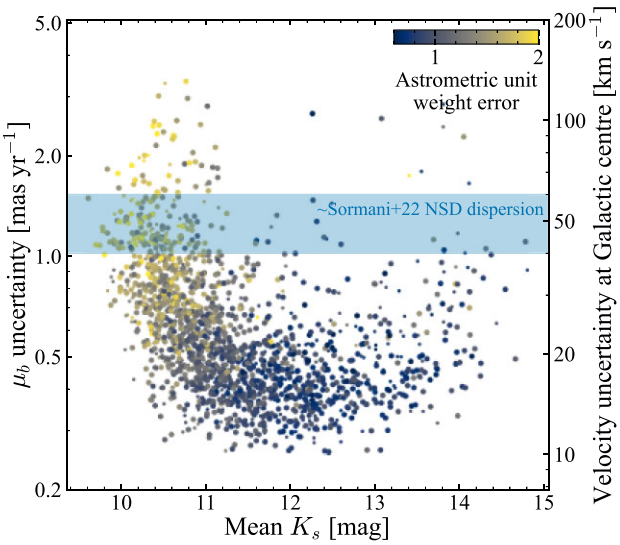

Figure 1. Astrometric data quality: Galactic latitudinal proper motion uncertainty (μ_b) against VIRAC K_s averaged over observed epochs coloured by astrometric unit weight error. The smaller points are unreliable or low amplitude. The band is the approximate vertical velocity dispersion of the NSD from Sormani et al. (2022a).

such that the standard deviation does not capture the full distribution) which the proper motion uncertainties of our sample are small enough to resolve.

The parent sample from Sanders et al. (2022b) is further cleaned and filtered to isolate higher confidence Mira variables that probe the NSD region:

(i) 91 Mira variable candidates have been flagged as unreliable from a visual inspection of the light curves (Sanders et al. 2022b);
(ii) 342 Mira variable candidates have K_s amplitudes <0.4 making their Mira variable classification more suspect (Matsunaga et al. 2009);
(iii) 18 stars have periods outside the 100 to 1000-d range typically used to define Mira variables;
(iv) 23 stars have proper motions more than 3 times the standard deviation away from the median suggesting they have unreliable astrometry or are foreground contaminants.

After removing these stars, we further restrict our analyses to stars at low Galactic latitude $|b| < 0.4$ deg (58 pc projected height) as these best probe the NSD. These selections result in a sample of 1163 stars. The 0.5th and 99.5 percentile of the period distribution of this sample are 105 and 845 d respectively. We do not remove possible NSC contaminants: within a projected radius of $r_{\rm eff} \approx 5$ pc from the Galactic Centre, there are 21 stars and 51 within $2r_{\rm eff}$ so any contamination is likely low.

We display the data sample in Fig. 2. The concentration of Mira variables towards the plane is a combination of the NSD flattening but perhaps more predominantly selection effects (see next section). We have fitted two simple models to the Galactic longitudinal and latitudinal proper motion distributions of the reliable, high-amplitude, low-latitude sample: (i) a Gaussian model with a flexible smoothed cubic spline as a function of period (Campagne et al. 2023, with 10 equally spaced knots in $\log_{10}P$ between the.5th and 99.5th percentile of the data) for the mean and standard deviation, and (ii) a two-component Gaussian mixture model. Both models account for the formal uncertainties in the proper motions. The models are

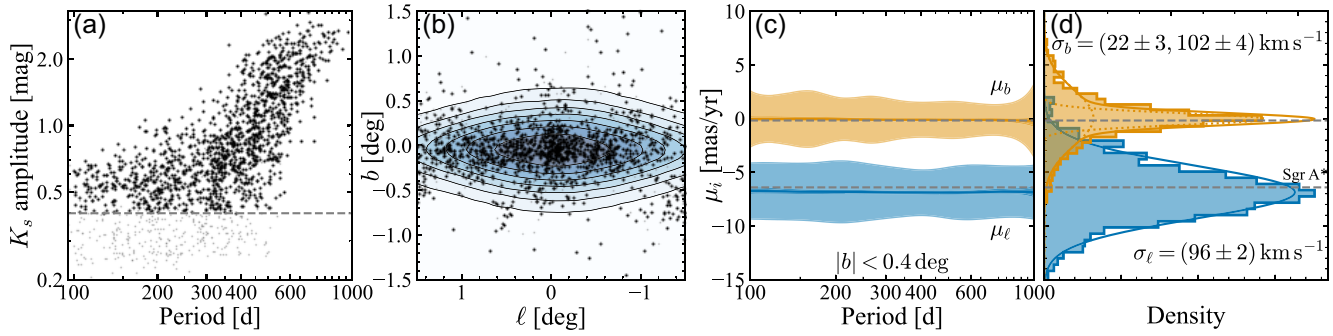

Figure 2. Summary of NSD Mira variable sample: panel (a) shows the period–amplitude distribution (unreliable and low-amplitude Mira variable candidates are shown with smaller grey points); panel (b) shows the distribution of reliable, high-amplitude Mira variable candidates on the sky with the background contours showing the ratio of the NSD surface density to the total (NSD + bar-bulge) surface density (excluding the NSC) from the models of Sormani et al. (2022a) (the contours are equally spaced every 10 per cent with the outermost contour corresponding to 10 per cent); panel (c) shows the mean proper motion, its uncertainty (darker bracket), and the standard deviation (lighter bracket) for the reliable, high-amplitude Mira variables with $|b| < 0.4$ deg; panel (d) shows histograms and two-component Gaussian mixture models (solid lines are the full models and dotted lines each component) for the proper motions of the same sample. Note the μ_b model distribution is narrower than the data as the plotted models are not convolved with uncertainties. The proper motion of Sgr A∗ is marked by grey dashed lines in panels (c) and (d).

implemented in JAX (Bradbury et al. 2018) and NUMPYRO (Bingham et al. 2019; Phan, Pradhan & Jankowiak 2019), and sampled using the NUTS sampler (Hoffman, Gelman et al. 2014).[1] The results are shown in the right panels of Fig. 2. The mean proper motions are in accord with the motion of Sgr A∗. The standard deviations are relatively constant with period but a slight narrowing is visible towards longer periods (particularly for μ_b; note the broadening at long period occurs beyond the 99.5th percentile of the sample). As assessed by the Bayesian information criterion, two components are not necessary for the μ_ℓ distribution which is well represented by a single Gaussian with $\sim 100\,\mathrm{km\,s^{-1}}$ dispersion (assuming all stars are situated at the Galactic centre; note Shahzamanian et al. 2022, suggest μ_ℓ is well modelled by three Gaussian components which might improve the fit in the wings of panel (d) of Fig. 2 but we have not considered such complex models here). However, the μ_b distribution consists of two Gaussian components (with approximately equal weights): a colder core with a dispersion of $(22 \pm 3)\,\mathrm{km\,s^{-1}}$ and a hotter component with $(102 \pm 4)\,\mathrm{km\,s^{-1}}$. Using giant stars in a central field of $|\ell| \lesssim 0.25$ deg and $|b| \lesssim 0.1$ deg, (Shahzamanian et al. 2022) also find two Gaussian fits are appropriate with the hotter component approximately consistent with our findings. However, the colder component from Shahzamanian et al. (2022) has dispersion $\sim 60\,\mathrm{km\,s^{-1}}$. This is broader than our findings here possibly due to the decay of the NSD vertical dispersion with radius, the non-Gaussianity of the NSD velocity distribution meaning the balance of bar and NSD with location impacts the results or the impact of background disc stars on our sample (Sormani et al. 2022a). The mean velocity of the colder component is $(-2.0 \pm 1.7)\,\mathrm{km\,s^{-1}}$ in relatively good agreement with the expected reflex motion from the Sun's vertical velocity of $7.25\,\mathrm{km\,s^{-1}}$ (Schönrich, Binney & Dehnen 2010) giving good evidence that the colder component is part of the NSD.

We investigate further the narrowing of the dispersion with period by fitting Gaussian mixture models as a function of period. We fit two-component Gaussian mixture models where the standard deviation of the first colder Gaussian that represents the NSD component is a cubic spline with period whilst the second hotter Gaussian that represents the background bar-bulge component has a period-independent dispersion. These choices are to keep the models simple whilst capturing the features in the data. We set the Gaussian means at the reflex solar velocity. The results are shown in Fig. 3 for two variants: one for fixed mixture weight with period and one with a cubic spline variation of the mixture weight with period (all splines are set to have 10 knots logarithmically spaced between 120 and 900 d). We note how in both cases the dispersion is around $20\,\mathrm{km\,s^{-1}}$ at long period and then there is a transition around 300-d period below which the dispersion rises towards that of the second hotter component and the expected dispersion of the bar-bulge (around $90\,\mathrm{km\,s^{-1}}$, Sanders et al. 2019).

The drawbacks of this simple approach are that (i) it does not capture spatial trends in kinematics, (ii) the velocity distributions are not Gaussian (Sormani et al. 2022a) and (iii) no selection function or line-of-sight distance distribution is considered and we effectively consider all of the stars at the Galactic centre distance. However, these initial considerations have clearly demonstrated that (i) some part of the data suggest NSD kinematics, (ii) there is a transition around a period of ∼300 d with longer period stars showing colder NSD kinematics and shorter period stars exhibiting more bar-bulge-like kinematics, and (iii) the contamination from the surrounding bar-bulge is significant, we will now turn to a more sophisticated modelling approach to address the highlighted drawbacks and elucidate some of these features of the data further.

3 MULTICOMPONENT DYNAMICAL MODELLING FRAMEWORK

The aim is to model the transverse kinematics (proper motions, $\boldsymbol{\mu} = \boldsymbol{v}_\mathrm{t}/s$ for transverse velocity $\boldsymbol{v}_\mathrm{t}$ and distance s) of the Mira variable stars as a function of their on-sky locations (Galactic coordinates: ℓ, b) and their periods, P. We are therefore building a likelihood from $p(\boldsymbol{\mu}|\ell, b, P)$. This is similar in spirit to Zhang & Sanders (2023). We begin by writing

$$p(\boldsymbol{\mu}|\ell, b, P) = \int \mathrm{d}^2\tilde{\boldsymbol{\mu}}\,\mathrm{d}s\,\mathcal{N}(\boldsymbol{\mu}|\tilde{\boldsymbol{\mu}}, \boldsymbol{\Sigma}_{\boldsymbol{\mu}})S(\ell, b, s, P)p(\tilde{\boldsymbol{\mu}}, s|\ell, b, P), \quad (1)$$

where $\mathcal{N}(\boldsymbol{\mu}|\tilde{\boldsymbol{\mu}}, \boldsymbol{\Sigma}_{\boldsymbol{\mu}})$ is a multivariate Gaussian accounting for proper motion uncertainties and $\tilde{\boldsymbol{\mu}}$ the true proper motions. We neglect

[1] See https://adrian.pw/blog/flexible-density-model-jax/ for a useful description of using JAX and NUMPYRO with splines for stream modelling.

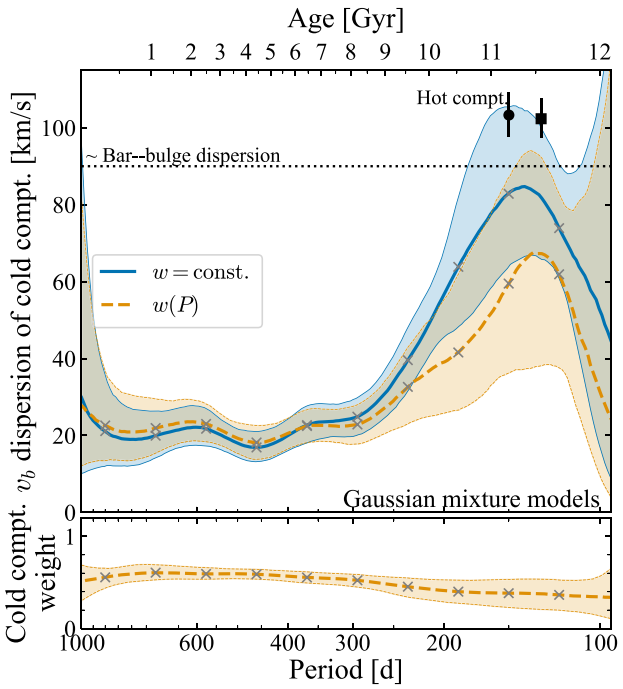

Figure 3. Flexible Gaussian mixture model fitting results to the latitudinal proper motion distributions, μ_b. We display the dispersion of the colder component (converted into a vertical velocity assuming all stars are at the Galactic Centre distance of 8.275 kpc) with crosses showing the spline knots. The blue solid line is for a model with a fixed cold component weight with period and the orange dashed line is for a period-dependent cold component weight (as displayed in the lower panel). The brackets are ±1 standard deviation. The hot component has dispersion given by the error bars (circle for the fixed weight model and square for the variable; positioned at an arbitrary x-axis location). The approximate dispersion of the bar-bulge is given by the dotted line.

uncertainties in the period – typically the posterior distribution from Lomb-Scargle/Fourier fitting methods is narrow or significantly multimodal due to a series of alias peaks (see fig. C1 of Sanders et al. 2022b, for tests of the period recovery quality of the sample). $S(\ell, b, s, P)$ is the selection function which gives the fraction of stars at each Galactic coordinate (ℓ, b), distance s and period P that enters the sample. Appendix A discusses our approach towards modelling the selection function. We shall show how its impact on our results is minimal.

For the kinematic model, $p(\tilde{\boldsymbol{\mu}}, s|\ell, b, P)$, we adopt a mixture model that is a combination of an axisymmetric NSD model and a 'background' bar/disc model (Sormani et al. 2022a, hereafter S22) – labelled 'bar'. We write

$$p(\tilde{\boldsymbol{\mu}}, s|\ell, b, P) = \frac{p(\ell, b, \tilde{\boldsymbol{\mu}}, s|P)}{p(\ell, b|P)} = \frac{w(P)p_{\rm NSD}(\ell, b, s, \tilde{\boldsymbol{\mu}}|P) + p_{\rm bar}(\ell, b, s, \tilde{\boldsymbol{\mu}})}{w(P)p_{\rm NSD}(\ell, b|P) + p_{\rm bar}(\ell, b|P)}. \quad (2)$$

$p_{\rm NSD}(\ell, b, s, \boldsymbol{\mu}|P)$ is the NSD model for each population labelled by period P with $p_{\rm NSD}(\ell, b|P)$ its marginalization over distance and proper motion weighted by the selection function, S,

$$p_{\rm NSD}(\ell, b|P) \equiv \int {\rm d}^2\boldsymbol{\mu}\, {\rm d}s\, S(\ell, b, s, P)p_{\rm NSD}(\ell, b, s, \boldsymbol{\mu}|P), \quad (3)$$

and likewise for $p_{\rm bar}(\ell, b|P)$ (note that although the bar model is independent of period, the marginalized distribution is conditioned on the period as the selection function depends on period). $w(P)$ is a period-dependent weight function that gives the total mass ratio of the NSD stars to the bar stars at fixed period. In later modelling, we either keep w fixed independent of period or allow it to be a flexible interpolated cubic spline with period.

Our model is not sensitive to the relative number of stars at each (ℓ, b) which is dependent on the extinction and detectability of the Mira variables (amongst other things), but is sensitive to the relative fraction, $F(\ell, b, P)$, of stars in each component at each on-sky location (ℓ, b) and with a period, P:

$$F(\ell, b, P) \equiv \frac{p_{\rm bar}(\ell, b|P)}{w(P)p_{\rm NSD}(\ell, b|P)}. \quad (4)$$

Furthermore, our model is not sensitive to the total normalization (or mass of the model). We opt to normalize $p_{\rm NSD}$ by the mass of the NSD found in S22, $M_{\rm NSD}$. In this way, $w(P)$ is the total mass ratio of the NSD to the bar at fixed period relative to the mass ratio found in S22 i.e. $w = 1$ corresponds to the relative NSD/bar weight found in S22. We now describe the specific model components in more detail.

3.1 Nuclear stellar disc model

$p_{\rm NSD}$ is modelled using action-based distribution functions (DFs):

$$p_{\rm NSD}(\ell, b, s, \tilde{\boldsymbol{\mu}}|P) = \int {\rm d}v_{||}\, s^4 \cos b\, f(\boldsymbol{J}|P). \quad (5)$$

Here, $\boldsymbol{J}$ are action coordinates computed from a set of observables $(\ell, b, s, \boldsymbol{\mu}, v_{||})$ with $v_{||}$ the line-of-sight velocity in a choice of axisymmetric potential, Φ (here fixed). The actions $\boldsymbol{J}$ are a triplet of integrals of motion, (J_r, J_ϕ, J_z), that approximately give the amplitude of radial oscillation, the degree of circulation, and the amplitude of the vertical oscillation of each orbit, respectively. We use the AGAMA (Vasiliev 2019) implementation of the 'Stäckel fudge' (Binney 2012; Sanders & Binney 2016) for the action computation. $s^4\cos b$ is a Jacobian factor between the observable coordinates and the actions. The potential is modelled by combining the NSC model from Chatzopoulos et al. (2015), the best-fitting NSD model of S22 and an axisymmetrized version of the Portail et al. (2017) potential (including the dark matter halo) with the central nuclear component removed as parametrized by Sormani et al. (2022b). The spherical enclosed mass for this potential is shown in Fig. 4 alongside mass measurements from McGinn et al. (1989), Genzel et al. (1996), Lindqvist et al. (1992), Burton & Liszt (1978), and Portail et al. (2017). We work with an axisymmetrized potential as algorithms for the computation of actions in rotating barred potentials are not available. Correspondingly, the action-based models we use here are axisymmetric. Whilst the known non-axisymmetric contribution from the Milky Way bar-bulge impacts the dynamics of the gas in the NSD region, it likely has a much weaker impact on the orbit-averaged dynamics of warmer components such as the stars. As discussed in the Introduction, it is possible that the NSD is a non-axisymmetric nuclear bar (Alard 2001, e.g.) as observed in other galaxies (Erwin 2024), but S22 have demonstrated axisymmetric models provide good fits to NSD stellar data giving confidence that the assumption of axisymmetry will only weakly impact our results.

$f(\boldsymbol{J}|P)$ is the action-based DF parametrized by period. We choose the 'quasi-isothermal' class of disc DFs introduced in Binney (2010) and Binney & McMillan (2011) given by the functional form

$$f(\boldsymbol{J}) = \frac{\Omega}{2\pi^2\kappa^2}\tilde{\Sigma}(R_c)f_R(J_R)f_z(J_z)\iota(J_\phi) \quad (6)$$

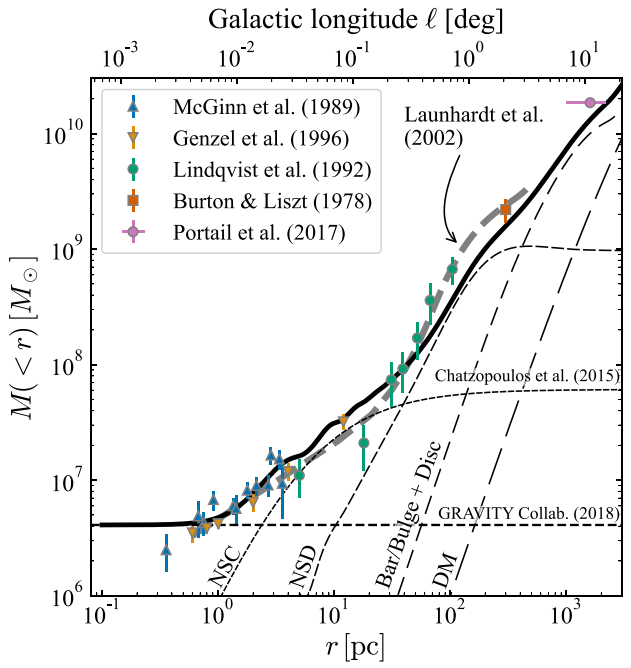

Figure 4. Spherical mass enclosed for the adopted gravitational potential: the solid black line shows the total with the breakdown by different components shown by the series of labelled dashed lines. The thick dashed grey line shows the model of Launhardt et al. (2002), and the coloured data points measurements from the literature (the measurement from Portail et al. 2017 is given within a cuboid so the depicted error bar shows the minimum and maximum dimensions of that cuboid). Note Sgr A∗ is not included in the potential for the modelling in Section 3.

where

$$\begin{aligned}\tilde{\Sigma}(R_c) &= \Sigma_0 \exp(-R_c/R_{\rm NSD}),\\ f_R(J_R) &= \frac{\kappa}{\tilde{\sigma}_R^2}\exp(-\kappa J_R/\tilde{\sigma}_R^2),\\ f_z(J_z) &= \frac{\nu}{\tilde{\sigma}_z^2}\exp(-\nu J_z/\tilde{\sigma}_z^2),\end{aligned} \tag{7}$$

and $t(J_\phi)$ is a taper function given by

$$t(J_\phi) = \begin{cases} 1, & J_\phi \geq 0,\\ \exp\left(\frac{2\Omega J_\phi}{\tilde{\sigma}_r^2}\right), & J_\phi < 0.\end{cases} \tag{8}$$

All epicyclic frequencies (κ, ν, Ω) are evaluated at the radius R_c of a circular orbit with angular momentum

$$\tilde{J} = \sqrt{(|J_\phi| + k_R J_R + k_z J_z)^2 + \tilde{J}_{\rm min}^2}, \tag{9}$$

where $k_R = k_z = 0.25$ and following S22 we set $\tilde{J}_{\rm min} = 10\,{\rm kpc\,km\,s^{-1}}$. The two dispersion functions, $\tilde{\sigma}_R$ and $\tilde{\sigma}_z$, are given by

$$\begin{aligned}\tilde{\sigma}_R^2(R_c) &= \sigma_{R0}^2 \exp(-2(R_c - R_{\rm NSD})/R_{\sigma,R}) + \sigma_{\rm min}^2,\\ \tilde{\sigma}_z^2(R_c) &= 2H_{\rm NSD}^2 \nu^2(R_c) + \sigma_{\rm min}^2.\end{aligned} \tag{10}$$

where again following S22 we set $\sigma_{\rm min} = 10\,{\rm km\,s^{-1}}$.

A disadvantage of the quasi-isothermal DFs is they make explicit reference to the potential through the use of the epicyclic frequencies. This makes them slightly awkward when constructing self-consistent distribution functions. For this reason, alternative DFs have been proposed by Vasiliev (2019) and Binney & Vasiliev (2023). However, we opt to use the quasi-isothermal DFs as (i) they were used in S22, (ii) our aim is not to construct self-consistent DFs (the potential is fixed), and (iii) we have found the quasi-isothermal DF parametrization is more physically interpretable than the alternatives.

The introduced DF has four free parameters (for our purposes, the normalization Σ_0 can be absorbed into the mass ratio $w(P)$): the scale length $R_{\rm NSD}$, the scale height $H_{\rm NSD}$, the radial dispersion at the scale length of the disc σ_{R0} and the radial scale length of the dispersion fall-off, $R_{\sigma,R}$ (note this is slightly different to S22, who normalize the radial dispersion at $R = 0$). In our modelling, these are either fixed independent of period or modelled as flexible interpolated cubic splines in period.

S22 used the quasi-isothermal DF models to fit the line-of-sight velocity from Fritz et al. (2021) and proper motion distributions from VIRAC (Smith et al. 2018). They used a fully self-consistent procedure assuming the observed stars traced the underlying mass distribution (modified by a selection function) which together with a fixed potential for the NSC then sourced the potential the DF was computed in. As fiducial values, we consider the probability $f_{\rm S22}(\boldsymbol{J})$ of being part of the NSD as defined by the fitted S22 distribution function. Note the best-fitting quasi-isothermal DF parameters reported in S22 do not quite give the best quasi-isothermal DF fit to the NSD in our chosen potential due to the impact of the iterative self-consistency procedure in S22. The best-fitting parameters are $(R_{\rm NSD}, H_{\rm NSD}, \sigma_{R0}, R_{\sigma,R})_{\rm S22} = (51\,{\rm pc}, 19\,{\rm pc}, 112\,{\rm km\,s^{-1}}, 330\,{\rm pc})$, which will serve as useful data for priors in the subsequent modelling. S22 found the uncertainty on these parameters was 5 per cent for $R_{\rm NSD}$, 10 per cent for $H_{\rm NSD}$, 5 per cent for σ_{R0} and $R_{\sigma,R}$ is prior-dominated and only a lower limit is obtained. Note that these parameters are only loosely related to the genuine scale length, scale height, etc., of the model. For example, the scale length and scale height can be measured directly from a model realization as 74 pc and 26 pc respectively.

3.2 Bar–disc contamination model

In addition to the component of interest, the NSD, we must also model the contribution from the other foreground and background components of the Galaxy. S22 have shown that the 'contamination' for a spectroscopically targeted sample is significant across the entire NSD region so it is reasonable to assume our Mira variable sample suffers from similar contamination.

The model for this contaminant component, $p_{\rm bar}$, is taken from an N-body model from Portail et al. (2017) with the central nuclear component removed (the NSD model fulfils this role in our model). This model has been fitted using the made-to-measure method to bar(-bulge) star counts, spectroscopy and proper motion data and was demonstrated by S22 to accurately capture the 3D velocity distributions of giant stars in the NSD region when combined with an NSD model. We remove model star particles at distances beyond 8 kpc from the Galactic centre as (i) the outer disc ($R > 5.5$ kpc) of the Portail et al. (2017) model was imposed as a data-motivated constant-scale height exponential disc but not explicitly fitted to data, (ii) it is likely extinction means few very distant disc stars enter our sample and the extinction maps used in Appendix A only extend to ~ 10.5 kpc from the Sun, and most importantly (iii) the Wesenheit cuts employed by Sanders et al. (2022b) to clean the sample remove any stars more distant than ~ 16 kpc from the Sun (see fig. D1 from Sanders et al. 2022b). The Portail et al. (2017) model is rotated such that the major axis of the bar lies at an angle of 25 deg with respect to the Sun–Galactic Centre line. This component has a fixed functional form in the fitting. Only its relative contribution as a function of P is considered [via $w(P)$]. We compute kernel density estimates (KDEs) of $p_{\rm bar}(\tilde{\boldsymbol{\mu}}, s|\ell, b)$ and $p_{\rm bar}(s|\ell, b)$ weighted by the

mass of each particle for a regular grid of small circular regions in (ℓ, b) of solid angle $\Delta\Omega$. The KDEs are computed using a fast Fourier transform with the KDEpy package.[2] Furthermore, we store $p_{\rm bar}(\ell, b)$ at these locations by simply summing the mass in each region and dividing by the on-sky area of the region. When required, these quantities are linearly interpolated in (ℓ, b) for an arbitrary star location.

3.3 Coordinate systems

We follow S22 and define two Cartesian coordinate systems: the first is centred a distance 8.275 kpc from the Sun (Gravity Collaboration et al. 2021) towards $(\ell, b) = (0, 0)$ with the Sun-Galactic centre line in the (x, y) plane, and the second is aligned with the first but shifted by 8.0 pc in the direction opposite Galactic rotation and 6.6 pc towards the Galactic South Pole. In this way, the origin of the second system is centred on Sgr A$*$ at $(\ell, b) = (-0.056, -0.046)$ deg, the bottom of the potential well used in the modelling. We set the solar motion as $4.74 \times 6.41\,{\rm mas\,yr^{-1}} \times 8.275\,{\rm kpc} = 251.4\,{\rm km\,s^{-1}}$ in the ℓ direction (Reid & Brunthaler 2020), $11.1\,{\rm km\,s^{-1}}$ towards the centre of the Galaxy and $7.25\,{\rm km\,s^{-1}}$ towards the Galactic North Pole (Schönrich et al. 2010). The multicomponent potential and action coordinates are computed in the frame centred on Sgr A$*$.

3.4 Computational specifics

Our fitting procedure relies on marginal distributions so per-star integrals are required. We follow Zhang & Sanders (2023) in the efficient computation of these integrals. N_s proper motion samples are generated for each star from the uncertainty distributions: $\boldsymbol{\mu}' \sim \mathcal{N}(\tilde{\boldsymbol{\mu}}, \boldsymbol{\Sigma}_{\boldsymbol{\mu}})$.[3] These are complemented by line-of-sight velocity samples and distance samples drawn from two Gaussians: $v' \sim \mathcal{N}(v_0, \sigma_v)$ with $v_0 = 0$ and $\sigma_v = 200\,{\rm km\,s^{-1}}$ and $s' \sim \mathcal{N}(s_0, \sigma_s)$ with $s_0 = 8.275$ kpc and $\sigma_s = 0.2$ kpc, respectively. The actions and frequencies for each star's set of samples are pre-computed. Then, for each star, we compute $p_{\rm NSD}(\ell, b, s, \tilde{\boldsymbol{\mu}}|P)$ as

$$p_{\rm NSD}(\ell, b, \tilde{\boldsymbol{\mu}}|P) \approx \frac{1}{N_s}\sum_i^{N_s} \frac{S(\ell, b, s_i', P)\, s_i'^4 \cos b\, f(\boldsymbol{J}')}{\mathcal{N}(v'|v_0, \sigma_v)\mathcal{N}(s'|s_0, \sigma_s)}. \tag{11}$$

Similarly, the denominator terms are found by drawing samples for $\boldsymbol{v}' = (v_\ell', v_b', v_{\rm los}') \sim \mathcal{N}(\boldsymbol{v}_0, {\rm diag}(\boldsymbol{\sigma}_v))$ with $\boldsymbol{v}_0 = (-250, 0, 0)\,{\rm km\,s^{-1}}$ and $\boldsymbol{\sigma}_v = (100, 100, 100)\,{\rm km\,s^{-1}}$ pre-computing all of the actions and frequencies $(\boldsymbol{J}')$, and evaluating

$$p_{\rm NSD}(\ell, b|P) \approx \frac{1}{N_s}\sum_i^{N_s} \frac{S(\ell, b, s_i', P)\, s_i'^2 \cos b\, f(\boldsymbol{J}')}{\mathcal{N}(\boldsymbol{v}'|\boldsymbol{v}_0, \boldsymbol{\sigma}_v)\mathcal{N}(s'|s_0, \sigma_s)}. \tag{12}$$

Note that as we are generating samples in the physical velocity space, not the proper motion space the Jacobian is s^2 instead of s^4. We use $N_s = 500$ samples for the 'numerator' quantities and $N_s = 1000$ for the 'denominator'. We adopt a very similar procedure for the computation of the integrals over the background bar model as

$$p_{\rm bar}(\ell, b, \tilde{\boldsymbol{\mu}}|P) \approx \frac{1}{N_s'}\sum_i^{N_s'} \frac{S(\ell, b, s_i', P) p_{\rm bar}(s, \tilde{\boldsymbol{\mu}}|\ell, b) p_{\rm bar}(\ell, b)}{\mathcal{N}(s'|s_0, \sigma_s')}, \tag{13}$$

where $N_s' = 10N_s$, $\sigma_s' = 3$ kpc and the terms in the numerator are described in Section 3.2. The computation of $p_{\rm bar}(\ell, b|P)$ is very similar (using $p_{\rm bar}(s|\ell, b)$ instead of $p_{\rm bar}(s, \tilde{\boldsymbol{\mu}}|\ell, b)$). Where possible, these quantities are pre-computed. With all components of our modelling approach defined, we now turn to fitting the models to the data sample.

Table 1. Fitted DF parameters for models independent of period. The first columns give the results fixing the NSD DF parameters to their values from Sormani et al. (2022a, given in the fourth column) with and without a selection function. The third column gives the results when the DF parameters are also fitted incorporating the priors implemented in log-space using the means as the logarithms of the values from the fourth column and the prior widths in log-space from the fifth column as shown in equation (14).

Parameter	Fixed	No SF	Marg.	S22	Prior width
w	0.8 ± 0.1	0.9 ± 0.2	0.7 ± 0.1	1	3
$R_{\rm NSD}$ [pc]	–	–	56 ± 7	51	0.6
$H_{\rm NSD}$ [pc]	–	–	16 ± 1	19	0.6
$\sigma_{R,0}$ [km s^{-1}]	–	–	124 ± 17	112	0.3
$R_{\sigma,R}$ [pc]	–	–	257 ± 83	331	0.3

3.5 Priors and model implementation

We place priors on the logarithms of the various parameters to ensure positivity. For period-independent models, we use the priors listed in Table 1 (identical to those given in equation 14). For more flexible models we allow the relative weight $\ln w(P)$ and the NSD DF parameters $(\ln R_{\rm NSD}, \ln H_{\rm NSD}, \ln \sigma_{R,0}, \ln R_{\sigma,R})$ to be fitted interpolated cubic splines in $\log_{10}P$. By default, we use 12 equally spaced knots in $\log_{10}P$/d between 1.9 and 3.1 (although we also consider equally spaced in age later). The 12 knot values for each free function are the parameters of the model on which we place priors. Extreme regions of parameter space can be assigned high likelihood due to poor estimates of the integrals from the samples we use. For this reason, we introduce smooth lower limits to the parameters at $(R_{\rm NSD}, H_{\rm NSD}, \sigma_{R,0}, R_{\sigma,R}, 1/w(P)) = (5\,{\rm pc}, 5\,{\rm pc}, 50\,{\rm km\,s^{-1}}, 50\,{\rm pc}, 10^{-6})$.

Identifiability of the model components is important. Giving the NSD component too much freedom allows it to replicate the bar-bulge component and we are unable to distinguish NSD versus bar-bulge stars. From the work of S22, we have prior knowledge of the approximate form of the NSD DF. We, therefore, place priors on all the knots as

$$\begin{aligned}
\ln w(P) &\sim \mathcal{N}(0, 3),\\
\ln R_{\rm NSD} &\sim \mathcal{N}(\ln R_{\rm NSD,S22}, 0.6),\\
\ln H_{\rm NSD} &\sim \mathcal{N}(\ln H_{\rm NSD,S22}, 0.6),\\
\ln \sigma_{R,0} &\sim \mathcal{N}(\ln \sigma_{R,0,\rm S22}, 0.3),\\
\ln R_{\sigma,R} &\sim \mathcal{N}(\ln R_{\sigma,R,\rm S22}, 0.3),
\end{aligned} \tag{14}$$

and then we introduce a smoothing prior $P_{\rm smooth}$ e.g.

$$\ln P_{\rm smooth} \propto \sum_i -\frac{(\ln w_i(P) - \ln w_{i+1}(P))^2}{2\ln\sigma_w^2} - \ln\ln\sigma_w \tag{15}$$

[2] https://kdepy.readthedocs.io/en/latest/index.html

[3] We use the notation $x \sim \mathcal{N}(\mu, \sigma)$ to denote that x is a random variate drawn from a normal distribution with mean μ and standard deviation σ, and $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ for the vector version where $\boldsymbol{\Sigma}$ is a covariance matrix. $\mathcal{N}(x|\mu, \sigma)$ denotes evaluating the normal distribution with mean μ and standard deviation σ at x, and similar for the vector version.

for $w(P)$ where the smoothing scales are hyperparameters following half-normal priors $\sigma_i \sim \mathcal{N}(0, \tau_i)$ where

$$\ln(\tau_{R_{\rm NSD}}, \tau_{H_{\rm NSD}}, \tau_{\sigma_{R,0}}, \tau_{R_{\sigma,R}}, \tau_w) = (0.3, 0.3, 0.3, 0.3, 1).$$

Note that this combination of a Gaussian prior on each knot value and a smoothing prior makes the effective prior on each knot value tighter than just the Gaussian prior would suggest. 12 knots is a good compromise between speed and flexibility. As smoothing priors are adopted, the risk of 'overfitting' is minimal and more knots are preferable to capture all significant data features. To check the results, we also fit models binned by period adopting the priors given in equation (14). The models are implemented in JAX (Bradbury et al. 2018) and NUMPYRO (Bingham et al. 2019; Phan et al. 2019), and sampled using the NUTS sampler (Hoffman et al. 2014).

4 DYNAMICAL MODEL FITTING RESULTS

We use the sample of 1163 stars described in Section 2: 'reliable', high-amplitude ($\Delta K_s > 0.4$), low latitude ($|b| < 0.4$ deg) and with $>3\sigma$ outliers in proper motion removed.

4.1 Global models

We begin by performing 'global fits' which neglect any dependence of the NSD properties with period. The results of these investigations are shown in Table 1.

First, we consider a model where the relative weight w is independent of period. The uncertainty on the mass from the S22 model is around 10 per cent and the bar–disc total mass within the region is uncertain to around 6 per cent (Portail et al. 2017) such that from prior data w is constrained as $w = (1.00 \pm 0.12)$. We use the NSD weights from the S22 model, $f_{\rm S22}(\boldsymbol{J})$, finding $w = (0.85 \pm 0.14)$. Excluding the selection function leads to $w = (0.92 \pm 0.16)$ as the selection function biases towards the more distant, background stars which mimic the colder dispersion of the NSD meaning a lower weight of NSD is required. However, in both cases w is consistent with the unity value found by S22 demonstrating that the sample is representative of stars in the NSD region. If we simultaneously fit the DF parameters alongside the relative weight using the priors in equation (14), we find NSD parameters highly consistent with the results of S22 (see Table 1) although our constraint on $R_{\sigma,R}$ is entirely prior-dominated. In this model, the relative weight decreases to $w = (0.66 \pm 0.13)$ but is still consistent with the measurement from S22 within the respective uncertainties. These models further confirm that our sample contains NSD and bar-bulge stars in similar proportions to that found by S22 so is unbiased.

4.2 Period-dependent models

We now model the variation of the weighting term, $w(P)$, with period. As described in the previous section, we use an interpolated cubic spline. The DF parameters are fixed to their values from S22. We refer to this model as our '*Fiducial*' model. We show the resulting fit in Fig. 5. We notice that at short period the relative NSD to bar weight is significantly smaller than $w = 1$ (the average value over the full population) whilst at long period the NSD dominates with $w \approx 2$. The transition occurs around a period of 350 d (corresponding to an approximate age of 7 Gyr using the relations from Zhang & Sanders 2023). This transition appears suggestive of the formation of the NSD but we delay the detailed consequences of this model to a later section.

We further fit a model allowing all the NSD DF parameters to be free functions of P, dubbed '*Fully marginalized*' model, using the priors given in Section 3.5. As shown in Fig. 5, we find a much weaker transition from NSD kinematics to bar kinematics, but still the transition occurs around the same period. As discussed previously, identifiability becomes an issue and a very thick NSD can resemble the bar population. Indeed that appears to be happening as seen in the lower panels of Fig. 5 where panel (c) shows how the NSD scale height, $H_{\rm NSD}$ increases at short period to mimic the broader bar-bulge population echoing the conclusion of Fig. 3 where the vertical dispersion of the NSD component behaves similarly. The other parameters are relatively flat with period. We observe that periods longer than 300 day (so ages less than ~ 8 Gyr) there is a weak suggestion of inside-out formation ($R_{\rm NSD}$ decreasing with age) and dynamical heating ($H_{\rm NSD}$ increasing with age). However, largely the data appears consistent with no gradients in period. This is somewhat at odds with the results from Nogueras-Lara et al. (2022) for the Milky Way and Bittner et al. (2020) for external galaxies. However, dynamical mixing is likely significant in the NSD (where the orbital time is ~ 1 Myr) so any formation gradients may get rapidly washed out (e.g. Frankel et al. 2020). These conclusions are corroborated by the fits binned by period also shown in Fig. 5. In the bin around ~400 d, a very large NSD fraction is found, possibly suggesting a very significant burst at this epoch that is smoothed over by the spline model, but the uncertainties are large.

4.3 Model variants

We now explore a number of model variants that test different assumptions in our modelling. Although our selection function is well motivated, it is quite likely that there are some shortcomings in the approach. As an alternate extreme, we now consider a model *without* a selection function term (i.e. $S = 1$), dubbed '*No selection function*' model. We have already seen in the 'global model' that the selection function has minimal impact. Here we opt to fix the DF parameters with period to the S22 values. We display the result in the top panel of Fig. 5 which agrees well with the fiducial model incorporating the selection function but with a slightly shorter period transition from bar-dominated to NSD-dominated kinematics.

One concern with the modelling approach is that the 3D extinction maps are not sufficient to resolve extinction within the NSD region. It could well be that in some regions we fail to observe Mira variables on the far side of the NSD due to extreme extinction inside the NSD. This effect will severely bias the Galactic longitude proper motion distributions reducing the height of the more negative μ_ℓ peak. However, the Galactic latitude proper motion distributions will be less biased as at the distance of the NSD the near- and far-sides of the NSD have effectively the same distributions. Similarly, any biasing from missing bright nearside NSD stars due to saturation effects will have a lesser impact on the NSD μ_b distributions. We, therefore, run a model, dubbed the '*Only using* $\mu_{\rm b}$' model, ignoring the μ_ℓ measurements and instead marginalize over them in the way described in Section 3.4. This still ignores any extreme extinction variation in the NSD but makes the modelling less sensitive to such effects. Finally, as highlighted in Section 2, the proper motion uncertainties may be poorly estimated both from saturation effects (see Fig. 1) and also more broadly due to other calibration issues (Luna et al. 2023). We, therefore, consider the proper motion uncertainties inflated by the unit weight error shown in Fig. 1 and broadened by a further factor of 1.1 (Luna et al. 2023). This model is dubbed the '*Scaled proper motion errors*' model. Both the '*Only using* $\mu_{\rm b}$' model and the '*Scaled proper motion errors*' model

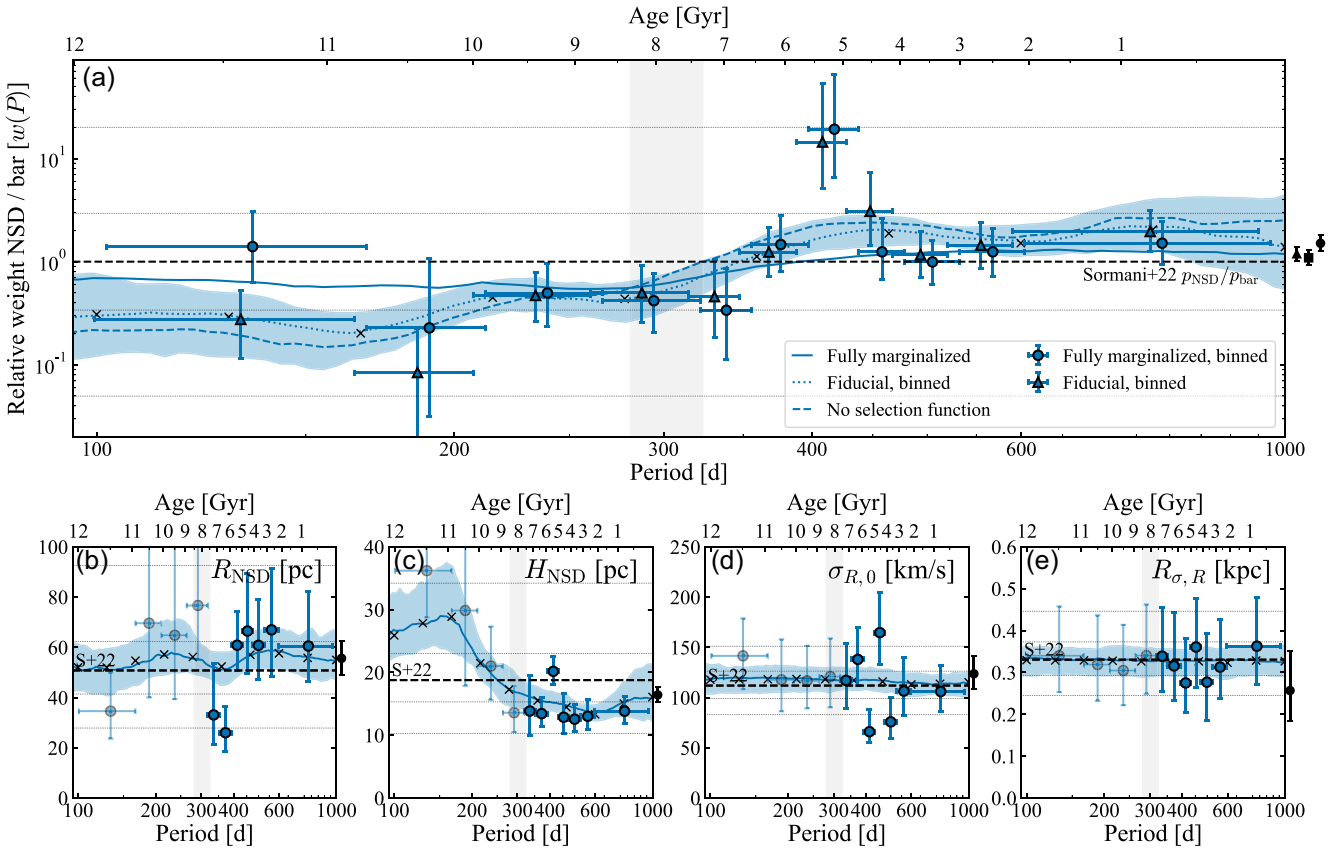

Figure 5. Dynamical model fitting results: panel (a) shows the relative weight of the NSD to the bar [$w(P)$ in the modelling]. The points are estimates using Mira variables split into bins [circles are the fully marginalized model and triangles the model with a fixed DF from Sormani et al. (2022a)]. The three lines are from the spline models with fixed weights (solid), no selection function (dashed), and fully marginalized (dotted). Panels (b–e) show the NSD DF fit results for the fully marginalized model as a function of period (age) [the scale length $R_{\rm NSD}$ in panel (b), scale height $H_{\rm NSD}$ in panel (c), central radial dispersion $\sigma_{R,0}$ in panel (d), and the radial scale length of the radial dispersion fall-off $R_{\sigma,R}$ in panel (e)]. The binned fits are the points and the line is the spline fit result. Faint horizontal lines show the $\pm1\sigma$ for the prior width for the spline (inner) and binned (outer) models. The black error bars outside each panel show the results for the period-independent models [triangle: fiducial, square: no selection function, circle: fully marginalized; note as the fiducial and no selection function models use fixed NSD DF parameters, only the fully marginalized results are shown for panels (b–e)]. In all panels, the crosses show the locations of spline knots.

produce very similar results to the fiducial model as we will see in Section 4.5.

4.4 Posterior predictive checks from mock sample generation

To perform posterior predictive checks, we generate mock samples from the fitted models. We use EMCEE (Foreman-Mackey et al. 2017) to generate samples in the observable space $(\ell, b, s, \boldsymbol{\mu}, v_{||}, P)$ from $s^4 \cos b\, f(\boldsymbol{J}|P)\mathcal{U}(\log_{10} P)$ where $\mathcal{U}$ is a uniform distribution and assign a total mass of $M_{\rm NSD}$ to the samples. We then select a subset of particles with associated masses from the Portail et al. (2017) models within the on-sky NSD region and assign them periods log-uniformly sampled between 100 and 1000 d. We multiply the NSD particle masses by $w(P)$ and then all particle masses by the selection fraction $S(\ell, b, s, P)$. From the combined set of NSD and bar samples, we randomly draw a subset of particles with probability proportional to their particle mass. This sample has the correct balance of bar and NSD at each location but does not capture the on-sky and period distributions of the data. To do this, we, for each datum, find the 100 nearest mock stars in (ℓ, b) and P using a Euclidean distance with scales of $\Delta\log_{10}P = 0.05$ and $(\Delta\ell, \Delta b) = (0.05, 0.05)$ deg. Finally, we convolve the mock proper motions with the errors of the matched datum. When a sample unaffected by selection function effects is required, we do not multiply the masses by the selection function and instead compute the period distribution of the data using the procedure described in the next subsection, and then reweight the samples by this distribution.

The results of this mock generation procedure for the fiducial model are shown in Figs 6 and B1 in Appendix B where the proper motion and line-of-sight velocity histograms for the data are compared to the model samples. From Fig. 6, we observe, for all period bins, a good model fit to the data. We note a weakly significant but systematic model underestimate of the $\mu_b \approx 0$ bins. Inspection of Fig. B1 shows this arises from model deficiencies at higher latitude that we attribute to limitations of the background disc model. The line-of-sight velocities are not used in the fits so they give good corroboration of the results and demonstrate the power of using a dynamical model. As the period increases, there is an increasing dominance of the NSD relative to the bar.

Further evidence of the presence of NSD member stars in our sample at long periods is given in Fig. 7. By restricting to stars in a 0.4 deg × 0.4 deg region centred on $(\ell, b) = (0, 0)$, we see that short-period ($P < 250$ d, old) stars have a unimodal μ_ℓ distribution centred on the motion of Sgr A∗ with a skew towards positive μ_ℓ

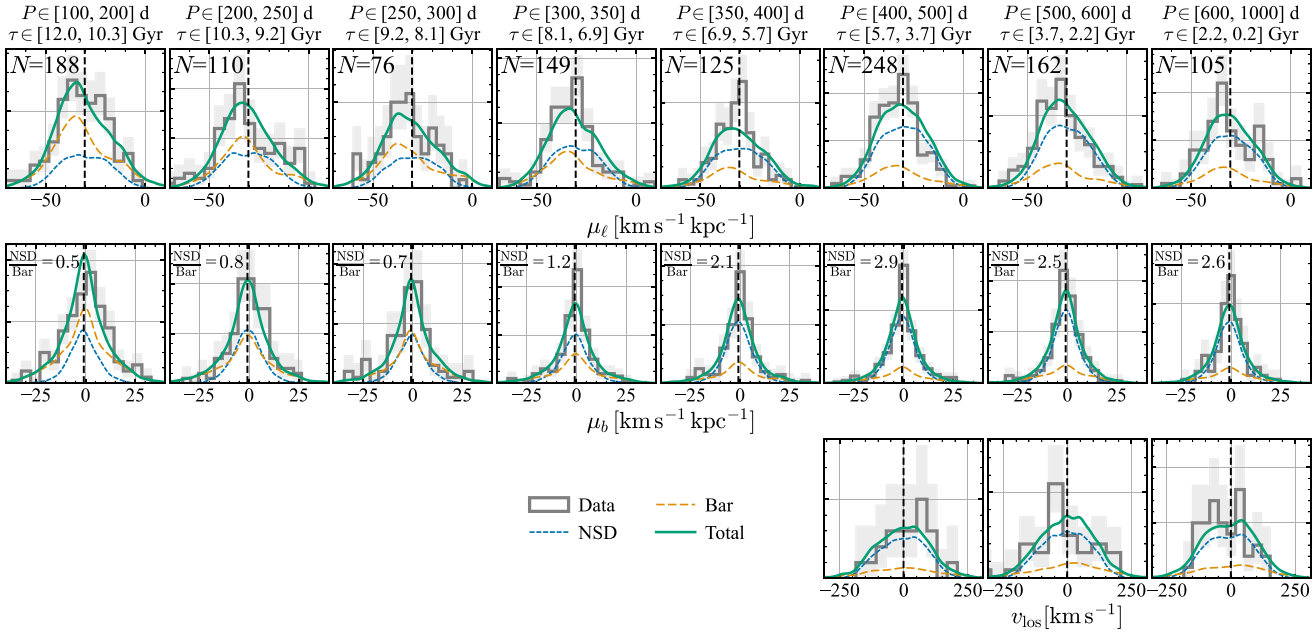

Figure 6. Velocity distributions for the data (grey with Poisson uncertainty shaded interval) compared to the best-fitting model split by period (age). The green line is the total model, blue short-dashed line is the NSD model, and orange long-dashed line is the bar/disc contamination model. The top row shows the Galactic longitude proper motions, μ_ℓ, the middle row shows the Galactic latitude proper motions μ_b, and the bottom row shows line-of-sight velocities, $v_{\rm los}$, not used in the analysis. Each column corresponds to a different period bin as indicated above the top row. In the top row, the number of stars in the proper motion panels is given and in the middle panel, the model fraction of the NSD to the bar is shown.

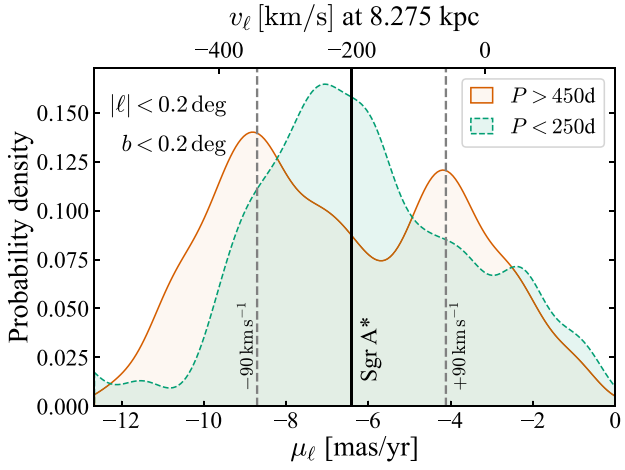

Figure 7. Galactic longitude proper motion distributions for two samples located $|\ell| < 0.2$ deg and $b < 0.2$ deg and with proper motion uncertainties < 1.5 mas yr^{-1}. The blue solid line shows a 'young' sample with periods >450 d where the rotation of the NSD is visible whilst the orange dashed line shows an 'old' sample with periods <250 d showing no net rotation with respect to Sgr A∗.

characteristic of the background bar–disc population, whilst the long-period ($P > 450$ d, young) stars have a bimodal distribution characteristic of the rotating NSD population (e.g. Shahzamanian et al. 2022) with a rotation amplitude of ~ 90 km s^{-1}. This demonstrates that our sample appears to probe both sides of the NSD and we are not limited by dust within the NSD region (at least in some parts).

4.5 NSD star formation history and the bar formation time

We now proceed to estimate the star formation history of the NSD and in turn the epoch of formation of the Galactic bar. Under the assumptions of our model, the relative contribution of NSD stars and contaminant bar/disc stars at each location is given by $F(\ell, b, P)$ from equation (4). Therefore, we can convert the observed period distribution at a given location $p(P|\ell, b)$ into the period distribution of the NSD, $p_{\rm NSD}(P)$, as

$$p_{\rm NSD}(P) = \frac{p_{{\rm NSD},S=1}(\ell, b|P)}{p_{\rm NSD}(\ell, b|P)} \frac{p(P|\ell, b)}{1 + F(\ell, b, P)}, \tag{16}$$

and similarly using the weight factor $F/(1 + F)$ for $p_{\rm bar}(P)$. The first ratio is the impact of the selection function on the period distributions: the ratio of the on-sky density given period under the selection function divided by the same but without considering the selection function. The second ratio gives the observed period distribution of the NSD which is affected by the selection function. Multiplication by the first term undoes this effect. The left-hand side is independent of the on-sky position so can be estimated using any subset of stars. We, therefore, construct histograms of the period distribution of the sample weighted by $(1 + F)^{-1}$ and have confirmed that similar histograms are obtained when limiting only to 'high'-latitude stars. To convert these distributions into star formation histories we have to first adopt a period–age relation. We consider the 'Both' and 'With GC'[4] relations from table 3 of Zhang & Sanders (2023). These share the parametric form

$$\tau(P) = \frac{\tau_0}{2}\left(1 + \tanh\left[\frac{330\,{\rm d} - P}{P_0}\right]\right), \tag{17}$$

[4]GC for globular cluster not to be confused with GC for Galactic Centre.

with $(\tau_0, P_0) = (13.7\,\mathrm{Gyr}, 401\,\mathrm{d})$ for the 'Both' fit and $(\tau_0, P_0) = (14.7\,\mathrm{Gyr}, 308\,\mathrm{d})$ for the 'With GC' fit. The first of these fits is a joint fit to the period–velocity dispersion in both the radial and vertical directions, whilst the second also includes information from cluster members. Both relations are quite similar but the 'With GC' relation is slightly steeper. They are also consistent with other relations used in the literature (e.g. Wyatt & Cahn 1983; Feast & Whitelock 1987, 2014; Catchpole et al. 2016; López-Corredoira 2017; Grady et al. 2020; Nikzat et al. 2022). There is likely significant scatter in the Mira period–age relation as demonstrated in the theoretical models from Trabucchi et al. (2019). Based on the differences between dynamical ages and cluster member ages, Zhang & Sanders (2023) quote a relative age scatter at fixed period in their 'With GC' fit of 11 per cent.

The next step is to convert the number density in age into a star formation rate. The number of Mira variables per unit age, τ, given a star formation rate law $\Gamma(\tau) \equiv \mathrm{d}M_\star/\mathrm{d}\tau$ and an IMF $\xi(M) \equiv \mathrm{d}N/\mathrm{d}M \propto (M/M_\odot)^{-2.3}$ (the number of stars formed per unit stellar mass normalized to give a total mass of star formation of $M_\star = M_\odot$) is

$$\frac{\mathrm{d}N}{\mathrm{d}\tau} = \int_{M_\mathrm{l}(\tau)}^{M_\mathrm{u}(\tau)} \mathrm{d}M\, \Gamma(\tau)\xi(M), \tag{18}$$

where $M_\mathrm{u}(\tau)$ and $M_\mathrm{l}(\tau)$ are the upper and lower initial masses of stars of age τ in the Mira phase. If the Mira phase lasts $\tau_\mathrm{Mira}(M)$ and a star becomes a Mira variable after approximately the main sequence lifetime of $\tau = (10\,\mathrm{Gyr})(M/M_\odot)^{-2.5}$ then we can write

$$\begin{aligned}\frac{\mathrm{d}N}{\mathrm{d}\tau} &\approx \tau_\mathrm{Mira}(\tau)\left(-\frac{\mathrm{d}M}{\mathrm{d}\tau}\right)\Gamma(\tau)\xi(M),\\ &\propto \tau_\mathrm{Mira}(\tau)\left(\frac{10\,\mathrm{Gyr}}{\tau}\right)^{0.48}\Gamma(\tau).\end{aligned} \tag{19}$$

The $\tau^{0.48}$ factor gives the combination of the number of stars formed of age τ that have masses consistent with being giant stars and the rate at which these giant stars are forming. $\tau_\mathrm{Mira}(\tau)$ gives the time stars born an age τ ago will be Mira variables for. This quantity is quite uncertain. Trabucchi et al. (2019) presented models showing the time spent in the Mira fundamental pulsation phase for stars of metallicity $Z = 0.008$ and masses $M = (1.5, 2.6, 4.8)M_\odot$ is $\sim (0.2, 0.15, 0.3)$ Myr. This suggests we can assume τ_Mira is approximately constant with age. In conclusion, we can map the period distribution to a star formation rate using

$$\frac{\mathrm{d}M_\star}{\mathrm{d}\tau} \propto \tau^{0.48}\frac{\mathrm{d}P}{\mathrm{d}\tau}\frac{\mathrm{d}N}{\mathrm{d}P}. \tag{20}$$

We apply this procedure using the 'fiducial' model as displayed in Fig. 8. The top panel of this figure shows the 'raw' age distribution using both period–age relations from Zhang & Sanders (2023). In the lower panel, we show the inferred star formation history of the NSD and the bar contaminant model. We define the 'bar formation epoch' as the time at which the NSD star formation history is maximally increasing. Baba & Kawata (2020) show from simulations that the NSD star formation history has a peak after bar formation validating our choice. From Fig. 8 we see this corresponds to ~ 8 Gyr for the fiducial model using the 'with GC' period–age relation from Zhang & Sanders (2023). The full set of results is given for each of the model variants in Table 2 and the inferred star formation histories are displayed in Fig. 9. We see that all models are consistent with the ~ 8 Gyr bar age estimate. Fig. 9 shows that in the *fully marginalized* model there is less separation between the bar-bulge and the NSD at old ages due to the identifiability issue discussed in Section 4.2. Using no selection function weakly reduces the NSD contribution at older ages as the selection function biases weakly against observing

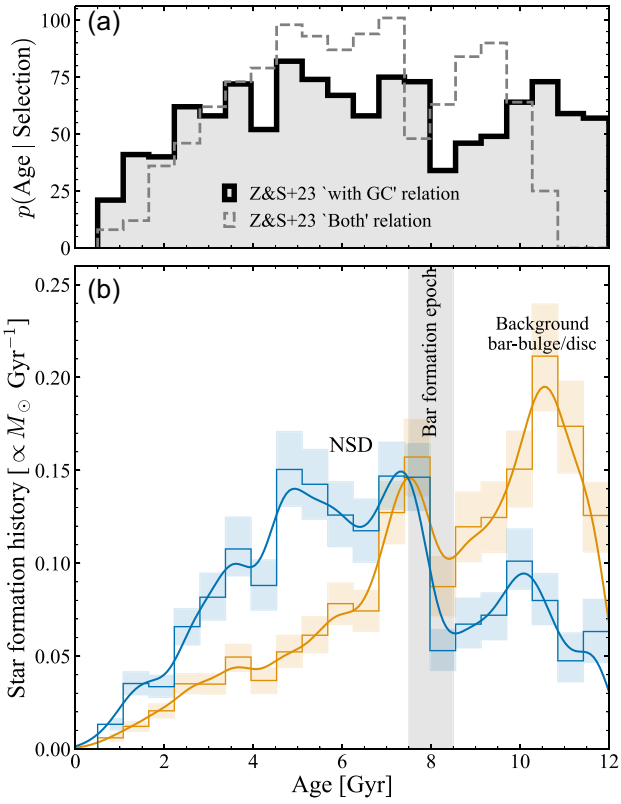

Figure 8. Star formation history of the NSD: panel (a) shows the 'raw' age distribution of the sample using two period–age relations as described in the text. Panel (b) shows the resulting inferred *normalized* star formation histories for the NSD (blue) and the background bar-bulge/disc (orange) using the 'with GC' distribution from panel (a).

Table 2. Bar formation epoch estimates: best estimates of the bar age using different modelling configurations. The scatter and possible systematic error in the period–age relation is around 1 Gyr as is the typical time difference between bar formation and the peak star formation in the NSD. We quote three significant figures just to illustrate the differences between the models.

Setup	Estimate [Gyr]
Fiducial	7.90
Alternate period–age relation	7.29
Marginalized	7.91
No selection function	7.87
Only μ_b	7.88
Scaled proper motion errors	7.90
Age-separated knots	7.91

NSD stars. Using only μ_b produces a stronger age separation of NSD and bar-bulge evidencing the robustness of the conclusions to the uncertain interstellar dust distribution that could bias μ_ℓ. Both the *scaled proper motion uncertainties* model and the *linear-spaced age knots* model are very similar to the fiducial case.

Using the period–age relation from Zhang & Sanders (2023) that doesn't consider the cluster results ('Both') produces a 0.6 Gyr lower bar age estimate so we can consider an approximate 1 Gyr systematic uncertainty in the bar age estimate arising from the choice of period–age relation. Zhang & Sanders (2023) also investigated biases arising from velocity uncertainties in the age-velocity-dispersion relation

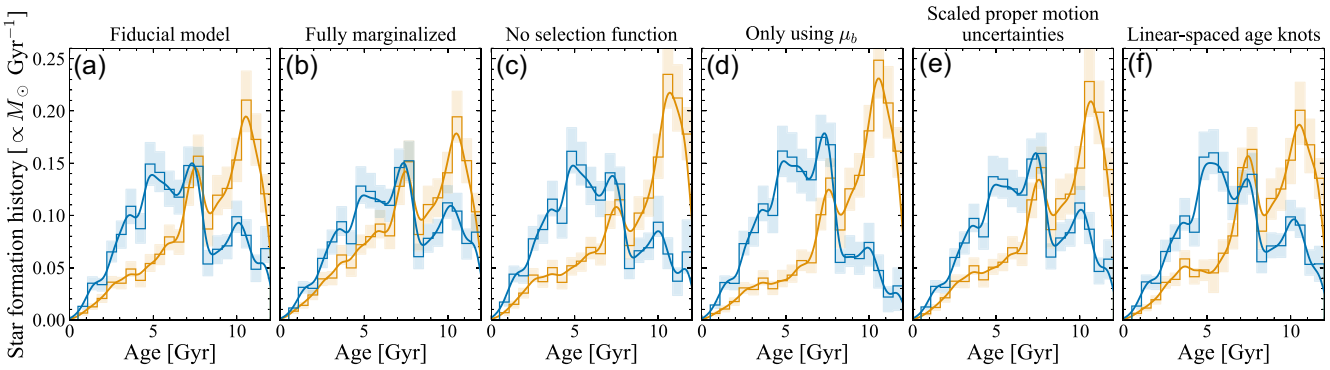

Figure 9. Inferred normalized star formation histories for the NSD (blue) and the bar/disc background (orange) with different modelling assumptions (as denoted above each panel; see Sections 4.2 and 4.3 for more details).

calibration. By assuming a uniform stellar age distribution they found that the period–age relation is essentially insensitive to velocity uncertainties around 8 Gyr meaning the systematic uncertainty on our result from this source of calibration error is subdominant. Fig. 8 also supports the metric proposed by Baba, Kawata & Schönrich (2022) where in a simulation without bar buckling the bar formation time is estimated from the transition age between the bar-bulge star formation history and that of the NSD. In a similar vein, in Appendix C we consider how the Milky Way bar age would be inferred by an external observer of the galaxy using the integrated field unit method of De Sá-Freitas et al. (2023a, b) finding consistent estimates. This gives good corroboration that methods for external galaxies match expectations when applied to observational data.

5 DISCUSSION AND CONCLUSIONS

In summary, our results from analysing the proper motions of Mira variables in the NSD region indicate:

(i) Around the 350-d period there is a transition in the proper motion kinematics with the shorter period (older) Mira variables appearing to have more bar-like kinematics whilst the longer period (younger) have more NSD-like kinematics (both colder kinematics and a multimodal rotation signature). This is evidenced by detailed distribution function modelling accounting for selection effects (Fig. 6), more basic Gaussian mixture models (Fig. 3) and simply plotting the data (Fig. 7).

(ii) Using a period–age relation, this implies that there is a sharp increase in star formation in the NSD around (8 ± 1) Gyr which we have identified with the time of bar formation. This feature is present irrespective of whether the selection function is included, whether the longitudinal proper motions are ignored and whether the proper motion errors are weakly modified. There is evidence for some NSD members older than 8 Gyr although as the bar kinematics are dominant in this regime this is somewhat uncertain.

(iii) More flexible models of the NSD with period demonstrate weak gradients in the NSD scale length, scale height and radial velocity dispersion with period consistent with the NSD forming inside-out and having undergone dynamical heating over time but the formation gradients have likely been washed out by significant dynamical mixing.

To close, we will now discuss the significance of these results and highlight some limitations of the work and possible future extensions.

5.1 Modelling limitations

There are several assumptions in our modelling that could be relaxed in future work or with a better understanding of the NSD and its stellar tracers. For example, we have assumed that the age distribution of the background bar–disc 'contamination' model is independent of distance and that the velocity distributions at a given location are independent of age. It is well known that velocity dispersion-age-position correlations exist in the disc and the bar-bulge (e.g. Hasselquist et al. 2020; Sharma et al. 2021). Too much freedom to the background model will remove our constraining power as the background can become confused with the NSD component, so more prior information from the literature and other data sets would be required, or a more detailed understanding of the distance distribution of our sample employed.

Our discussion of the selection function highlighted that different Mira variable populations could trace different K_s period–luminosity relations (Sanders 2023). The cause for this is unclear but the first-order effects might be related to metallicity (Trabucchi et al. 2019). Without detailed spectroscopy, even approximate metallicities for our sample are not possible, but it seems likely there is some metallicity variation in our sample and in particular, a variation in mean metallicity between the NSD and bar populations (Schultheis et al. 2021). This adds a layer of complexity to the selection function where we may be losing the intrinsically brighter more metal-poor objects that make up the bar population in favour of the fainter metal-rich objects that are preferentially part of the NSD population. We have demonstrated that our conclusions are independent of selection effects so this effect is not a significant concern but it highlights that there is potential in future work to combine the models with spectroscopic data sets (Fritz et al. 2021; Schultheis et al. 2021, S22) to yield more powerful constraints.

5.2 Comparison with other NSD studies

Our derived formation epoch for the NSD is consistent with the idea of a 'middle-aged' to old Galactic bar as found in other studies (Bovy et al. 2019; Nogueras-Lara et al. 2020; Wylie et al. 2022; Schödel et al. 2023; Nogueras-Lara et al. 2023b) and seems to strongly rule out any suggestion that the Milky Way bar is only a few Gyr old. We here discuss more critically the comparison of our result with other related work. Our derived star formation histories for the NSD and bar are compared to other studies in Fig. 10. We have opted to plot the results against fractional time to divide out small age systematics we will discuss below.

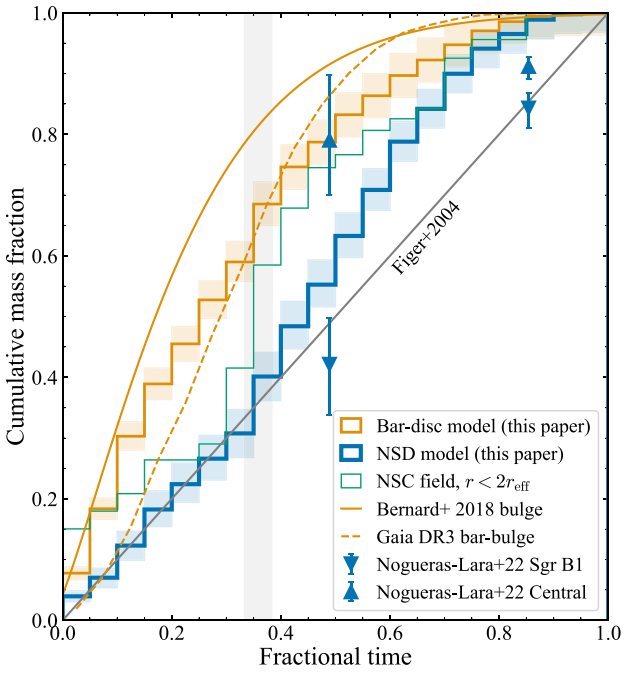

Figure 10. Comparison with other studies: the histograms show cumulative mass growth of the NSD (thick blue), bar-bulge (thinner orange), and the NSC region (thinnest green). The orange line shows the cumulative star formation history of the bar-bulge inferred by Bernard et al. (2018), the dashed orange line shows the cumulative bar-bulge star formation history inferred from *Gaia* DR3 Mira variables within $|\ell| < 1.5$ deg and $|b| < 1.5$ deg, and the blue error bars are cumulative mass measurements for two regions of the NSD from Nogueras-Lara et al. (2022). The grey line is a uniform star formation history as advocated by Figer et al. (2004). The x-axis is fractional time (time/maximum time) to accommodate slightly different age scales (see the text for details).

Our NSD star formation history is approximately consistent with the idea of an old burst as advocated by Nogueras-Lara et al. (2020, 2023b) and Schödel et al. (2023). Somewhat different to those studies, we find weak evidence for very ancient ($\gtrsim 10$ Gyr) populations in the NSD. Our derived star formation history is relatively continuous with age (Figer et al. 2004). Simulations of the nuclear stellar discs, such as those of Seo et al. (2019), predict a highly bursty star formation history highly dependent on the large-scale supply of gas. However, when smoothed over the typical scatter in the Mira variable period–age relation or ~ 1 Gyr (Trabucchi & Mowlavi 2022; Zhang & Sanders 2023), these simulations appear relatively consistent with the results presented here. One feature of the Nogueras-Lara et al. (2020) picture we are missing is a recent 1 Gyr ago star formation burst in the NSD (Schödel et al. 2023, finds this burst could constitute ∼15 per cent of the mass in the NSD). It is clear from the presence of very young stellar tracers that there is ongoing star formation in the NSD region (Morris & Serabyn 1996; Matsunaga et al. 2015; Henshaw et al. 2023) but the identification of these less recent significant star formation episodes comes primarily from the presence of secondary red clump stars (Nogueras-Lara et al. 2020; Schödel et al. 2023). Another indicator of recent star formation is C-rich Mira variables (Boyer et al. 2013; Matsunaga et al. 2017; Sanders & Matsunaga 2023). The single star formation channel for these objects is through dredge-up making their production more effective for massive and/or metal-poor stars. However, binary formation channels are also likely (Matsunaga et al. 2017; Sanders & Matsunaga 2023). Sanders et al. (2022b) used colour–colour diagrams to demonstrate $\lesssim 10$ per cent of the present sample is consistent with being C-rich but this number is somewhat uncertain due to extinction effects. It is uncertain whether this number of objects is consistent solely with production through binary channels or whether it indicates some recent star formation in the population. Another possibility for missing a significant recent star formation episode is that our Mira variable selection is lacking long-period objects. Long-period Mira variables are easily identified as they are high amplitude, but they could be missed due to significant dust embedding or very irregular light curves due to circumstellar dust. A final reason for the discrepancy at recent times could be from our simple conversion from Mira variable number density to star formation history and/or the employed Mira period–age relation is inappropriate for the youngest stars (our employed period–age relations are an extrapolation below ~ 3 Gyr).

It is interesting to briefly consider the relationship between the NSD and NSC star formation histories. The NSC appears to predominantly contain stars that are older than 5 Gyr but on average younger than the bar-bulge (see the summary and references given by Neumayer et al. 2020). The recent study of Chen et al. (2023) argued the NSC has an age of 5^{+3}_{-2} Gyr, younger than previous estimates due to their more flexible modelling of the metallicity distribution. Interestingly, this is very similar to the mean age of the NSD from our analysis. In Fig. 10, we show the cumulative mass fraction from the 51 Mira variables we have within a projected distance of $2r_{\rm eff} = 8.4$ pc from the NSC centre where we notice an uptick at exactly the bar formation time proposed in this work. These two pieces of evidence suggest a close evolutionary link between the NSD and the NSC (Nogueras-Lara et al. 2023a).

Although not the focus of the work, our inferred bar star formation history at $|b| \lesssim 0.4$ deg is approximately consistent with that presented by Bernard et al. (2018) using *HST* data in several bar-bulge fields at 2 deg $\lesssim |b| \lesssim 5$ deg. Our age scale saturates around 12 Gyr whilst the age scale of Bernard et al. (2018) extends to 14 Gyr. This is suggestive of a $12/14 \times 8$ Gyr ≈ 1 Gyr systematic in our bar age estimate. For this reason, we have opted to plot the fractional time in Fig. 10. Even accounting for this, our bar star formation history is biased slightly young with respect to the results of Bernard et al. (2018). This result could be interpreted as the product of fractionation when the bar formed (Debattista et al. 2017) whereby the older hotter population forms a more vertically extended boxy-peanut bulge than the younger cooler population, thus producing a vertical age gradient through the bulge. However, the models using only the Galactic latitude proper motion (see Fig. 9) are biased older and more consistent with the results of Bernard et al. (2018) so this discrepancy appears within the systematics of our methodology. Finally, the central bar-bulge Mira variables in *Gaia* DR3 (Lebzelter et al. 2023), as identified with a G-band semi-amplitude cut of 0.45, periods between 100 and 1000 d and in the region $|\ell| < 1.5$ deg and $|b| < 1.5$ deg, produce a star formation history consistent with our bar-bulge star formation history with a slightly younger bias possibly due to foreground disc contamination.

5.3 Alternative NSD formation scenarios

We have weak evidence that the NSD is older than the ~ 8 Gyr star formation burst signal as some fraction of the oldest stars are attributed to the NSD. This could reflect contamination in the sample or shortcomings of the bar–disc model as discussed above, or could genuinely reflect a much more ancient NSD that had a 'slow start'. A slow-start picture is somewhat opposed to the expectation from

simulations (Baba & Kawata 2020) where a significant burst in star formation is expected at the time of formation. However, potentially the Galaxy was relatively gas-poor at these early times (although this seems unlikely, Daddi et al. 2010) and only at some later epoch (possibly coinciding with a merger event) was significantly more gas accreted leading to a star formation burst. This discussion highlights that our bar age estimate is only ever going to be a lower bound as gas is required in the Galaxy at the time of bar formation and then some time is required for its transport to the centre. Another interpretation is that an ancient NSD formed after an early bar formation, the bar was destroyed (again possibly related to merger events) and then later a second bar formed. This idea is somewhat supported by the differing NSD kinematics for stars younger and older than 8 Gyr. However, there is limited evidence for the destruction of bars after formation both theoretically and observationally [see discussion in section V.B.6 of Sellwood 2014, although see recent simulation work from Bi et al. (2022) and Cavanagh et al. (2022)]. Therefore, we are well-justified in discussing a single bar formation epoch. Finally, it is possible that the old NSD material formed through a different channel. For instance, it could be produced by accretion through mergers that sink and remain intact deep into the Milky Way potential. From our calculations, the old NSD material makes up a significant fraction of the mass (30 per cent so around $3 \times 10^8 M_\odot$) which would require the order of hundreds of accreted clusters so seems unlikely. Another possibility is we are seeing the remnant of the very earliest 'disc' that formed in the primordial Milky Way, an analogue of those seen at high redshift (Kikuchihara et al. 2020; Ono et al. 2023).

5.4 Future observational prospects

One limitation of using VVV data to probe the Mira variables in the NSD region is that, due to saturation, for $K_s \lesssim 10.5$ mag we only detect those NSD Mira variables with significant circumstellar and/or interstellar extinction. Additionally, this means our sample is strongly confined to high-extinction regions in the Galactic plane. In the presented modelling, we have included this effect and demonstrated it has a relatively weak impact on the conclusions. Ideally, we would use *all* Mira variables in the NSD. However, this requires both their identification from multiepoch data and measured proper motions. One option is *Gaia*: unfortunately due to the effects of extinction the *Gaia* DR3 Mira variable sample (Lebzelter et al. 2023) is confined to $|b| \gtrsim 0.5$ deg where the NSD stars are expected to only contribute $\lesssim 10$ per cent relative to the background bar-bulge (see Fig. 2). Therefore, there is a 'gap' in both the Mira variable on-sky and magnitude ($9 \lesssim K_s \lesssim 10.5$) distributions between the *Gaia* stars and the VIRAC sample presented here. This is somewhat filled in by the Matsunaga et al. (2009) sample, but these lack proper motion information. One promising future candidate for filling this gap is the *Japan Astrometry Satellite Mission for INfrared Exploration* (*JASMINE*, Kawata et al. 2023), a planned pointed astrometric satellite mission of the central Galactic bulge. *JASMINE*'s Galactic Centre Survey is designed to cover exactly this on-sky and magnitude range ($9.5 \lesssim H_w \lesssim 14.5$) so if there are Mira variables there, *JASMINE* should find them. Finally, it is possible that the Roman telescope could reach down to the main sequence turn-off in the NSD (Terry et al. 2023) although confusion, crowding and foreground/background subtraction may still be challenging. This would provide an independent way to date the different components of the NSD and also give a way to measure the age-metallicity variations in the NSD, an option not possible with Mira variables and currently only possible through combination with spectroscopic data sets (Fritz et al. 2021; Schultheis et al. 2021; Sormani et al. 2022a). Our results demonstrate the great promise of further pursuing more sophisticated modelling of the limited NSD data sets as well as future surveys and projects to better understand this important but hard-to-study region of the Galaxy.

5.5 The timeline of the Galaxy and the wider context

We close by returning to our initial goal: placing the bar formation epoch in the timeline of the Milky Way. We have presented evidence that the Milky Way bar formed ~ 8 Gyr ago. Within the expected systematic uncertainties of both estimates (~ 1 Gyr), this could place the bar formation time close to the infall of the Gaia-Sausage-Enceladus merger around 9–10 Gyr ago (Belokurov et al. 2018, 2020; Bonaca et al. 2020) such that potentially the Milky Way has a tidally induced bar (Łokas et al. 2014; Merrow et al. 2023). However, it is also likely that the early disc growth (Belokurov & Kravtsov 2022; Dillamore et al. 2023; Khoperskov et al. 2023; Semenov et al. 2023) and relative dynamical quietness after the early mergers in the Milky Way gave rise to conditions in the disc that were more conducive to bar formation. Either hunting for observational signatures that the Milky Way bar is tidally induced (Miwa & Noguchi 1998) or more accurate timing of both the bar formation and the merger time of GS/E will be needed to answer this question more concretely. It is also intriguing that around this $\sim$ 8–10 Gyr age the Milky Way transitions from thick disc dominance to thin disc dominance. There is then the suggestion that the bar formation plays a role in this population separation (Khoperskov et al. 2018) or possibly as with the merger discussion both are driven by the same root cause (Grand et al. 2020). Finally, measurement of the bar age can be combined with the observations of a presently slowing bar (Chiba & Schönrich 2021; Chiba et al. 2021) to approximately extrapolate to find the pattern speed of the bar at formation. This is ambitious and would require detailed modelling through simulations but may reveal insights into the relative importance of dark matter and gas for angular momentum transfer over the lifetime of the Galaxy. In conclusion, the synergy between high-redshift studies, large cosmological simulation suites and resolved stellar studies in the Milky Way of the type presented here are beginning to present a coherent picture of bar formation across the Universe.

ACKNOWLEDGEMENTS

We thank the referee for their close reading of the manuscript and useful comments. JLS acknowledges the support of the Royal Society (URF\R1\191555). DK acknowledges the support of the UK's Science & Technology Facilities Council (STFC grants ST/S000216/1 and ST/W001136/1) and MWGaiaDN, a Horizon Europe Marie Skłodowska-Curie Actions Doctoral Network funded under grant agreement no. 101072454 and also funded by UK Research and Innovation (EP/X031756/1). MCS acknowledges the financial support of the Royal Society (URF\R1\221118). DM gratefully acknowledges support from the ANID BASAL projects ACE210002 and FB210003, from FONDECYT project no. 1220724, and from CNPq Brasil project 350104/2022-0. We thank the organizers and contributors of the Galactic Bars 2023 conference for their excellent updates on the interesting field of bar formation. 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. Based on data products from observations made with ESO Telescopes at the La Silla or Paranal Observatories under ESO programme ID 179.B-2002. This paper made use of NUMPY (van der Walt, Colbert & Varoquaux 2011), JAX (Bradbury et al. 2018), NUMPYRO (Bingham et al. 2019; Phan et al. 2019), SCIPY (Virtanen et al. 2020), MATPLOTLIB (Hunter 2007), SEABORN (Waskom 2021), PANDAS (McKinney 2010), ASTROPY (Astropy Collaboration 2013; Price-Whelan et al. 2018), and AGAMA (Vasiliev 2019).

DATA AVAILABILITY

The Mira variable data set from the work of Sanders et al. (2022b) will be made available along with the proper motions via Vizier. Code related to this project is available at https://github.com/jls713/mira_nsd.

REFERENCES

Alard C., 2001, A&A, 379, L44
Anderson L. D. et al., 2020, ApJ, 901, 51
Antoja T. et al., 2014, A&A, 563, A60
Antoja T. et al., 2018, Nature, 561, 360
Arentsen A. et al., 2020, MNRAS, 491, L11
Astropy Collaboration, 2013, A&A, 558, A33
Athanassoula E., 1992, MNRAS, 259, 345
Athanassoula E., 2003, MNRAS, 341, 1179
Athanassoula E., 2008, MNRAS, 390, L69
Athanassoula E., Misiriotis A., 2002, MNRAS, 330, 35
Baba J., Kawata D., 2020, MNRAS, 492, 4500
Baba J., Kawata D., Schönrich R., 2022, MNRAS, 513, 2850
Beane A. et al., 2022, ApJ, 953, 15
Belokurov V., Kravtsov A., 2022, MNRAS, 514, 689
Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, MNRAS, 478, 611
Belokurov V., Sanders J. L., Fattahi A., Smith M. C., Deason A. J., Evans N. W., Grand R. J. J., 2020, MNRAS, 494, 3880
Bensby T. et al., 2013, A&A, 549, A147
Bensby T. et al., 2017, A&A, 605, A89
Bernard E. J., Schultheis M., Di Matteo P., Hill V., Haywood M., Calamida A., 2018, MNRAS, 477, 3507
Bhardwaj A. et al., 2019, ApJ, 884, 20
Bi D., Shlosman I., Romano-Díaz E., 2022, ApJ, 934, 52
Bingham E. et al., 2019, J. Mach. Learn. Res., 20, 28:1
Binney J., 2010, MNRAS, 401, 2318
Binney J., 2012, MNRAS, 426, 1324
Binney J., McMillan P., 2011, MNRAS, 413, 1889
Binney J., Vasiliev E., 2023, MNRAS, 520, 1832
Binney J., Gerhard O. E., Stark A. A., Bally J., Uchida K. I., 1991, MNRAS, 252, 210
Bittner A. et al., 2020, A&A, 643, A65
Bland-Hawthorn J., Tepper-Garcia T., Agertz O., Freeman K., 2023, ApJ, 947, 80
Blitz L., Spergel D. N., 1991, ApJ, 379, 631
Blommaert J. A. D. L., van der Veen W. E. C. J., van Langevelde H. J., Habing H. J., Sjouwerman L. O., 1998, A&A, 329, 991
Bonaca A. et al., 2020, ApJ, 897, L18
Bournaud F., Combes F., 2002, A&A, 392, 83
Bovy J., Leung H. W., Hunt J. A. S., Mackereth J. T., García-Hernández D. A., Roman-Lopes A., 2019, MNRAS, 490, 4740
Boyer M. L. et al., 2013, ApJ, 774, 83
Bradbury J. et al., 2018, http://github.com/google/jax
Buck T., Ness M. K., Macciò A. V., Obreja A., Dutton A. A., 2018, ApJ, 861, 88
Burton W. B., Liszt H. S., 1978, ApJ, 225, 815
Campagne J.-E. et al., 2023, Open J. Astrophys, 6, 15
Catchpole R. M., Whitelock P. A., Glass I. S., 1990, MNRAS, 247, 479
Catchpole R. M., Whitelock P. A., Feast M. W., Hughes S. M. G., Irwin M., Alard C., 2016, MNRAS, 455, 2216
Cavanagh M. K., Bekki K., Groves B. A., Pfeffer J., 2022, MNRAS, 510, 5164
Chatzopoulos S., Fritz T. K., Gerhard O., Gillessen S., Wegg C., Genzel R., Pfuhl O., 2015, MNRAS, 447, 948
Chen C.-C. et al., 2022, ApJ, 939, L7
Chen Z. et al., 2023, ApJ, 944, 79
Chiba R., Schönrich R., 2021, MNRAS, 505, 2412
Chiba R., Friske J. K. S., Schönrich R., 2021, MNRAS, 500, 4710
Churchwell E. et al., 2009, PASP, 121, 213
Clarke J. P., Wegg C., Gerhard O., Smith L. C., Lucas P. W., Wylie S. M., 2019, MNRAS, 489, 3519
Cole A. A., Weinberg M. D., 2002, ApJ, 574, L43
Cole D. R., Debattista V. P., Erwin P., Earp S. W. F., Roškar R., 2014, MNRAS, 445, 3352
Combes F., Sanders R. H., 1981, A&A, 96, 164
Combes F., Debbasch F., Friedli D., Pfenniger D., 1990, A&A, 233, 82
Daddi E. et al., 2010, ApJ, 713, 686
Debattista V. P., Ness M., Gonzalez O. A., Freeman K., Zoccali M., Minniti D., 2017, MNRAS, 469, 1587
Deguchi S. et al., 2004, PASJ, 56, 765
Dehnen W., 2000, AJ, 119, 800
Dillamore A. M., Belokurov V., Kravtsov A., Font A. S., 2023, MNRAS, 527, 7070
Drimmel R. et al., 2023, A&A, 674, A37
Eggen O. J., 1998, AJ, 115, 2435
Emsellem E., Renaud F., Bournaud F., Elmegreen B., Combes F., Gabor J. M., 2015, MNRAS, 446, 2468
Engels D., Bunzel F., 2015, A&A, 582, A68
Erwin P., 2018, MNRAS, 474, 5372
Erwin P., 2024, MNRAS, 528, 3613
Erwin P., Sparke L. S., 2002, AJ, 124, 65
Eskridge P. B. et al., 2000, AJ, 119, 536
Feast M. W., 1963, MNRAS, 125, 367
Feast M. W., Whitelock P. A., 1987, in Kwok S., Pottasch S. R.eds, Late Stages of Stellar Evolution. Springer, p. 33
Feast M., Whitelock P., 2000, in Matteucci F., Giovannelli F.eds, Astrophysics and Space Science LibraryVol. 255, p. 229
Feast M., Whitelock P. A., 2014, in Feltzing S., Zhao G., Walton N. A., Whitelock P.eds, IAU SymposiumVol. 298, Setting the scene for Gaia and LAMOST. p. 40
Feast M. W., Glass I. S., Whitelock P. A., Catchpole R. M., 1989, MNRAS, 241, 375
Feast M. W., Whitelock P. A., Menzies J. W., 2006, MNRAS, 369, 791
Ferreira L. et al., 2022, ApJ, 955, 15
Figer D. F., Rich R. M., Kim S. S., Morris M., Serabyn E., 2004, ApJ, 601, 319
Foreman-Mackey D., Agol E., Ambikasaran S., Angus R., 2017, AJ, 154, 220
Fragkoudi F. et al., 2020, MNRAS, 494, 5936
Fragkoudi F., Grand R. J. J., Pakmor R., Springel V., White S. D. M., Marinacci F., Gomez F. A., Navarro J. F., 2021, A&A, 650, L16
Frankel N., Sanders J., Ting Y.-S., Rix H.-W., 2020, ApJ, 896, 15
Frankel N. et al., 2022, ApJ, 940, 61
Fraser O. J., Hawley S. L., Cook K. H., 2008, AJ, 136, 1242
Freytag B., Liljegren S., Höfner S., 2017, A&A, 600, A137
Fritz T. K. et al., 2011, ApJ, 737, 73
Fritz T. K. et al., 2021, A&A, 649, A83
Fujii T., Deguchi S., Ita Y., Izumiura H., Kameya O., Miyazaki A., Nakada Y., 2006, PASJ, 58, 529
Fux R., 1999, A&A, 345, 787
Gadotti D. A., Seidel M. K., Sánchez-Blázquez P., Falcón-Barroso J., Husemann B., Coelho P., Pérez I., 2015, A&A, 584, A90
Gadotti D. A. et al., 2018, MNRAS, 482, 506
Gadotti D. A. et al., 2020, A&A, 643, A14

Gallego-Cano E., Schödel R., Nogueras-Lara F., Dong H., Shahzamanian B., Fritz T. K., Gallego-Calvente A. T., Neumayer N., 2020, A&A, 634, A71
Genzel R., Thatte N., Krabbe A., Kroker H., Tacconi-Garman L. E., 1996, ApJ, 472, 153
Gerhard O., Martinez-Valpuesta I., 2012, ApJ, 744, L8
Gerin M., Combes F., Athanassoula E., 1990, A&A, 230, 37
Ghosh S., Saha K., Di Matteo P., Combes F., 2021, MNRAS, 502, 3085
Glass I. S., Evans T. L., 1981, Nature, 291, 303
Glass I. S., Matsumoto S., Carter B. S., Sekiguchi K., 2001, MNRAS, 321, 77
Gonzalez O. A., Rejkuba M., Zoccali M., Valenti E., Minniti D., Schultheis M., Tobar R., Chen B., 2012, A&A, 543, A13
Grady J., Belokurov V., Evans N. W., 2019, MNRAS, 483, 3022
Grady J., Belokurov V., Evans N. W., 2020, MNRAS, 492, 3128
Grand R. J. J. et al., 2020, MNRAS, 497, 1603
Gravity Collaboration, 2021, A&A, 647, A59
Groenewegen M. A. T., 2004, A&A, 425, 595
Guo Y. et al., 2023, ApJ, 945, L10
Hasselquist S. et al., 2020, ApJ, 901, 109
Hatchfield H. P., Sormani M. C., Tress R. G., Battersby C., Smith R. J., Glover S. C. O., Klessen R. S., 2021, ApJ, 922, 79
Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, Nature, 563, 85
Henshaw J. D., Barnes A. T., Battersby C., Ginsburg A., Sormani M. C., Walker D. L., 2023, in Inutsuka S., Aikawa Y., Muto T., Tomida K., Tamura M.eds, Astronomical Society of the Pacific Conference SeriesVol. 534, p. 83
Hoffman M. D., Gelman A. et al., 2014, J. Mach. Learn. Res., 15, 1593
Hohl F., 1971, ApJ, 168, 343
Hohl F., 1976, AJ, 81, 30
Hunt J. A. S., Bovy J., 2018, MNRAS, 477, 3945
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Ita Y., Matsunaga N., 2011, MNRAS, 412, 2345
Ita Y. et al., 2004, MNRAS, 347, 720
Iwanek P., Soszyński I., Kozłowski S., 2021, ApJ, 919, 99
Jacobs C. et al., 2023, ApJ, 948, L13
Kawata D. et al., 2023, PASJ, preprint (arXiv:2307.05666)
Khoperskov S., Haywood M., Di Matteo P., Lehnert M. D., Combes F., 2018, A&A, 609, A60
Khoperskov S., Di Matteo P., Gerhard O., Katz D., Haywood M., Combes F., Berczik P., Gomez A., 2019, A&A, 622, L6
Khoperskov S., Minchev I., Steinmetz M., Ratcliffe B., Walcher J. C., Libeskind N., 2023, preprint (arXiv:2309.07321)
Kikuchihara S. et al., 2020, ApJ, 893, 60
Koposov S. E. et al., 2023, MNRAS, 521, 4936
Kraljic K., Bournaud F., Martig M., 2012, ApJ, 757, 60
Launhardt R., Zylka R., Mezger P. G., 2002, A&A, 384, 112
Le Conte Z. A. et al., 2023, preprint (arXiv:2309.10038)
Lebzelter T. et al., 2023, A&A, 674, A15
Lewis M. O., Bhattacharya R., Sjouwerman L. O., Pihlström Y. M., Pietrzyński G., Sahai R., Karczmarek P., Górski M., 2023, A&A, 677, 14
Li Z., Shen J., Gerhard O., Clarke J. P., 2022, ApJ, 925, 71
Lindqvist M., Habing H. J., Winnberg A., 1992, A&A, 259, 118
Łokas E. L., 2021, A&A, 647, A143
Łokas E. L., Athanassoula E., Debattista V. P., Valluri M., Pino A. d., Semczuk M., Gajda G., Kowalczyk K., 2014, MNRAS, 445, 1339
López-Corredoira M., 2017, ApJ, 836, 218
Luna A., Marchetti T., Rejkuba M., Minniti D., 2023, A&A, 677, 16
Macri L. M., Ngeow C.-C., Kanbur S. M., Mahzooni S., Smitka M. T., 2015, AJ, 149, 117
Marigo P. et al., 2022, ApJS, 258, 43
Matsunaga N., Fukushi H., Nakada Y., 2005, MNRAS, 364, 117
Matsunaga N., Kawadu T., Nishiyama S., Nagayama T., Hatano H., Tamura M., Glass I. S., Nagata T., 2009, MNRAS, 399, 1709
Matsunaga N. et al., 2015, ApJ, 799, 46
Matsunaga N., Menzies J. W., Feast M. W., Whitelock P. A., Onozato H., Barway S., Aydi E., 2017, MNRAS, 469, 4949
McGinn M. T., Sellgren K., Becklin E. E., Hall D. N. B., 1989, ApJ, 338, 824
McKinney W., 2010, in van der Walt Stéfan, Millman Jarrod eds, Proceedings of the 9th Python in Science Conference. p. 56
Melvin T. et al., 2014, MNRAS, 438, 2882
Merrill P. W., 1923, ApJ, 58, 215
Merrow A., Grand R. J. J., Fragkoudi F., Martig M., 2023, preprint (arXiv:2312.02318)ar
Messineo M., Habing H. J., Sjouwerman L. O., Omont A., Menten K. M., 2002, A&A, 393, 115
Messineo M., Habing H. J., Menten K. M., Omont A., Sjouwerman L. O., 2004, A&A, 418, 103
Minchev I., Famaey B., 2010, ApJ, 722, 112
Minniti D. et al., 2010, New A, 15, 433
Minniti D., Contreras Ramos R., Zoccali M., Rejkuba M., Gonzalez O. A., Valenti E., Gran F., 2016, ApJ, 830, L14
Miwa T., Noguchi M., 1998, ApJ, 499, 149
Monari G., Kawata D., Hunt J. A. S., Famaey B., 2017, MNRAS, 466, L113
Monari G., Famaey B., Siebert A., Wegg C., Gerhard O., 2019, A&A, 626, A41
Morris M., Serabyn E., 1996, ARA&A, 34, 645
Nepal S. et al., 2024, A&A, 681, L8
Neumayer N., Seth A., Böker T., 2020, A&A Rev., 28, 4
Nikzat F. et al., 2022, A&A, 660, A35
Nishiyama S. et al., 2013, ApJ, 769, L28
Noguchi M., 1987, MNRAS, 228, 635
Nogueras-Lara F. et al., 2019, A&A, 631, A20
Nogueras-Lara F. et al., 2020, Nat. Astron., 4, 377
Nogueras-Lara F., Schödel R., Neumayer N., 2022, Nat. Astron., 6, 1178
Nogueras-Lara F. et al., 2023a, A&A, 680, 13
Nogueras-Lara F., Schultheis M., Najarro F., Sormani M. C., Gadotti D. A., Rich R. M., 2023b, A&A, 671, L10
Ono Y. et al., 2023, ApJ, 951, 72
Ostriker J. P., Peebles P. J. E., 1973, ApJ, 186, 467
Pérez-Villegas A., Portail M., Wegg C., Gerhard O., 2017, ApJ, 840, L2
Petersen M. S., Peñarrubia J., 2021, Nat. Astron., 5, 251
Phan D., Pradhan N., Jankowiak M., 2019, NeurIPS 2019 Program Transformations for Machine Learning Workshop, preprint (arXiv:1912.11554)
Pizzella A., Corsini E. M., Morelli L., Sarzi M., Scarlata C., Stiavelli M., Bertola F., 2002, ApJ, 573, 131
Planck Collaboration, 2020, A&A, 641, A6
Portail M., Gerhard O., Wegg C., Ness M., 2017, MNRAS, 465, 1621
Price-Whelan A. M. et al., 2018, AJ, 156, 123
Quillen A. C., 2003, AJ, 125, 785
Quillen A. C., Minchev I., Sharma S., Qin Y.-J., Di Matteo P., 2014, MNRAS, 437, 1284
Raha N., Sellwood J. A., James R. A., Kahn F. D., 1991, Nature, 352, 411
Ramírez S. V., Arendt R. G., Sellgren K., Stolovy S. R., Cotera A., Smith H. A., Yusef-Zadeh F., 2008, ApJS, 175, 147
Reid M. J., Brunthaler A., 2020, ApJ, 892, 39
Riebel D., Meixner M., Fraser O., Srinivasan S., Cook K., Vijh U., 2010, ApJ, 723, 1195
Rix H.-W. et al., 2022, ApJ, 941, 45
Romeo A. B., Agertz O., Renaud F., 2023, MNRAS, 518, 1002
Rosas-Guevara Y. et al., 2022, MNRAS, 512, 5339
Roshan M., Ghafourian N., Kashfi T., Banik I., Haslbauer M., Cuomo V., Famaey B., Kroupa P., 2021, MNRAS, 508, 926
De Sá-Freitas C. et al., 2023a, A&A, 678, 12
De Sá-Freitas C. et al., 2023b, A&A, 671, A8
Saito R. K. et al., 2012, A&A, 537, A107
Sanders J. L., 2023, MNRAS, 523, 2369
Sanders J. L., Binney J., 2016, MNRAS, 457, 2107
Sanders J. L., Matsunaga N., 2023, MNRAS, 521, 2745
Sanders J. L., Smith L., Evans N. W., Lucas P., 2019, MNRAS, 487, 5188
Sanders J. L., Smith L., González-Fernández C., Lucas P., Minniti D., 2022a, MNRAS, 514, 2407
Sanders J. L., Matsunaga N., Kawata D., Smith L. C., Minniti D., Lucas P. W., 2022b, MNRAS, 517, 257

Savino A., Koch A., Prudil Z., Kunder A., Smolec R., 2020, A&A, 641, A96
Schödel R. et al., 2023, A&A, 672, L8
Schönrich R., Binney J., Dehnen W., 2010, MNRAS, 403, 1829
Schönrich R., Aumer M., Sale S. E., 2015, ApJ, 812, L21
Schultheis M. et al., 2014, A&A, 566, A120
Schultheis M. et al., 2021, A&A, 650, A191
Sellwood J. A., 2014, Rev. Mod. Phys., 86, 1
Sellwood J. A., Gerhard O., 2020, MNRAS, 495, 3175
Semczuk M., Dehnen W., Schönrich R., Athanassoula E., 2022, MNRAS, 517, 6060
Semenov V. A., Conroy C., Chandra V., Hernquist L., Nelson D., 2023, ApJ, 962, 18
Seo W.-Y., Kim W.-T., Kwak S., Hsieh P.-Y., Han C., Hopkins P. F., 2019, ApJ, 872, 5
Shahzamanian B., Schödel R., Nogueras-Lara F., Martínez-Arranz A., Sormani M. C., Gallego-Calvente A. T., Gallego-Cano E., Alburai A., 2022, A&A, 662, A11
Sharma S. et al., 2021, MNRAS, 506, 1761
Sheth K. et al., 2008, ApJ, 675, 1141
Sheth K., Melbourne J., Elmegreen D. M., Elmegreen B. G., Athanassoula E., Abraham R. G., Weiner B. J., 2012, ApJ, 758, 136
Shlosman I., Frank J., Begelman M. C., 1989, Nature, 338, 45
Simmons B. D. et al., 2014, MNRAS, 445, 3466
Smith L. C. et al., 2018, MNRAS, 474, 1826
Sormani M. C., Barnes A. T., 2019, MNRAS, 484, 1213
Sormani M. C., Binney J., Magorrian J., 2015a, MNRAS, 449, 2421
Sormani M. C., Binney J., Magorrian J., 2015b, MNRAS, 454, 1818
Sormani M. C. et al., 2022a, MNRAS, 512, 1857
Sormani M. C., Gerhard O., Portail M., Vasiliev E., Clarke J., 2022b, MNRAS, 514, L1
Sormani M. C., Sobacchi E., Sanders J. L., 2023, MNRAS, 528, 5742
Sparke L. S., Sellwood J. A., 1987, MNRAS, 225, 653
Surot F. et al., 2019, A&A, 623, A168
Tepper-Garcia T. et al., 2021, preprint (arXiv:2111.05466)
Terry S. K. et al., 2023, preprint (arXiv:2306.12485)
Trabucchi M., Mowlavi N., 2022, A&A, 658, L1
Trabucchi M., Wood P. R., Montalbán J., Marigo P., Pastorelli G., Girardi L., 2019, MNRAS, 482, 929
Tremaine S., Weinberg M. D., 1984, MNRAS, 209, 729
van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
Vasiliev E., 2019, MNRAS, 482, 1525
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Waskom M. L., 2021, J. Open Source Softw., 6, 3021
Wegg C., Gerhard O., 2013, MNRAS, 435, 1874
Whitelock P. A., Feast M. W., Van Leeuwen F., 2008, MNRAS, 386, 313
Wood P. R., 2015, MNRAS, 448, 3829
Wood P. R., Habing H. J., McGregor P. J., 1998, A&A, 336, 925
Wyatt S. P., Cahn J. H., 1983, ApJ, 275, 225
Wylie S. M., Clarke J. P., Gerhard O. E., 2022, A&A, 659, A80
Xiong D. R., Deng L., Zhang C., 2018, MNRAS, 480, 2698
Yu J., Liu C., 2018, MNRAS, 475, 1093
Yuan W., He S., Macri L. M., Long J., Huang J. Z., 2017a, AJ, 153, 170
Yuan W., Macri L. M., He S., Huang J. Z., Kanbur S. M., Ngeow C.-C., 2017b, AJ, 154, 149
Zhang H., Sanders J. L., 2023, MNRAS, 521, 1462
Zhao D., Du M., Ho L. C., Debattista V. P., Shi J., 2020, ApJ, 904, 170

APPENDIX A: SELECTION FUNCTION OF THE MIRA VARIABLE SAMPLE

In equation (1), $S(\ell, b, s, P)$ is the selection function, which is the fraction of stars at each (ℓ, b, s, P) entering our sample. We compute this quantity as

$$\begin{aligned} S(\ell, b, s, P) &= S(K_s(\ell, b, s, P)) \\ &= S\Big(M_{K_s,\mathrm{NSD}}(P) + 5\log_{10}(100[s/\mathrm{kpc}]) \\ &\quad + A_{K_s}(\ell, b, s)\Big). \end{aligned} \tag{A1}$$

This expression consists of several components. $S(K_s)$ is the completeness curve given by Sanders et al. (2022b), which is largely governed by the failure to detect bright objects due to saturation in VVV. Sanders et al. (2022b) assessed the impact of this effect using both artificial star tests and a comparison with the sample of Matsunaga et al. (2009). The two methods produce similar results that are well approximated by the simple functional form

$$S(K_s) = 0.45(1 + \tanh[2.5(K_s - 10.3)]), \tag{A2}$$

i.e. the sample is maximally around 90 per cent complete and stars with $K_s \lesssim 10.3$ have a significant probability to be missed. There are an insufficient number of Mira variables from Matsunaga et al. (2009) to fully assess the completeness at the faint end. The search of Sanders et al. (2022b) did not impose any explicit K_s cuts at faint magnitudes but period–Wesenheit magnitude cuts were employed by Sanders et al. (2022b) to clean the sample (see their fig. D1). These cuts are approximated by removing model star particles beyond ~ 16 kpc (see Section 3.2). The inclusion of a faint limit taper in $S(K_s)$ at $K_s \approx 13$ produces minimal changes to the selection-function-weighted bar model.

$A_{K_s}(\ell, b, s)$ is the 3D extinction map from Schultheis et al. (2014). We use a coefficient $A_{K_s}/E(H - K_s) = 1.306$ (Sanders et al. 2022a) to convert the $E(H - K_s)$ values reported by Schultheis et al. (2014). The resolution of the Schultheis et al. (2014) maps is 0.1 deg $\times$ 0.1 deg, so highly variable extinction on small on-sky scales will not be well captured. However, we later will describe how we also consider the reported extinction scatter in our procedure, which mitigates this effect. Furthermore, we have found the results in this paper change very little when we instead use the 2d extinction maps of Sanders et al. (2022a) combined with the distance dependence of the Schultheis et al. (2014) map assuming that the 2d map represents the extinction at the Galactic Centre distance.

A1 NSD Mira variable period–luminosity relation

$M_{K_s,\mathrm{NSD}}(P)$ is the period–luminosity relation of the Mira variables in our sample. As discussed extensively by Sanders et al. (2022b), after correcting the K_s magnitudes of the Mira variable sample for interstellar extinction (in that case using the extinction maps from Sanders et al. 2022a), the Mira variables in the NSD region are significantly fainter than those in the solar neighbourhood and LMC, as assessed by the period–luminosity relation $M_{K_s,\mathrm{LMC}}(P)$ (Sanders 2023, see also Glass et al. 2001 and Lewis et al. 2023) assuming the bulk of the observed stars are around the Galactic Centre distance. This can be seen in panel (b) of Fig. A1 and is slightly at odds with observations that suggest a near universality of the O-rich Mira variable K_s period–luminosity relation based on the comparison of observations of the solar neighbourhood and the LMC (Whitelock, Feast & Van Leeuwen 2008; Sanders 2023) although it is possible that metallicity effects for the supersolar NSD stars also play a role here (Schultheis et al. 2021). There are several other possible reasons for this discrepancy: (i) there is very significant background disc contamination – this would require some Mira variables being at distances far beyond the expected edge of the far side of the Galactic disc (e.g. see Nikzat et al. 2022) and also the colours of

the dereddened Mira variables do not agree with the expectation from the LMC (see panel a of Fig. A1). Furthermore, Sanders et al. (2022b) showed that several of the period–Wesenheit distributions for the NSD Mira variables (particularly those using mid-infrared bands) agree with LMC relations shifted to the Galactic Centre distance. These observations make this first hypothesis unlikely; (ii) the interstellar extinction is poorly estimated – this might be possible due to the highly fractal nature and extinction maps being averaged over large regions biased towards low extinction objects whilst we are biased towards high extinction objects). However, panel (a) shows that the colour excess discrepancy with respect to the LMC relation is a function of period which suggests that it is not a result of poorly estimated interstellar extinction unless there is a significant bias of finding longer period Mira variables in more heavily extincted regions which is a possibility; (iii) there is significantly more circumstellar extinction around these Mira variables than observed for the solar neighbourhood and LMC samples. This is likely to be a significant effect for the longer-period Mira variables but for $P \lesssim 300$ d circumstellar extinction is anticipated to be low (Ita & Matsunaga 2011). It appears likely that the cause of the discrepancy is a combination of options (ii) and (iii): interstellar and circumstellar extinction effects. Fortunately, the approach we take is agnostic to how these two effects are separated. Furthermore, the main results in the paper are insensitive to the inclusion or exclusion of the selection function.

We can relate the period–luminosity relation of the Mira variables in our sample to those in the LMC as

$$M_{K_s,\mathrm{NSD}}(P) = M_{K_s,\mathrm{LMC}}(P) + A_{K_s,R}(P), \tag{A3}$$

where $A_{K_s,R}(P)$ is a residual level of interstellar plus circumstellar extinction after the interstellar extinction correction from Schultheis et al. (2014). The residual extinction can be estimated by comparison of the colours of our sample to the period–colour relations for LMC Mira variables

$$\begin{aligned}(K_s - c) &\equiv (K_s - c)_0 + E_I(K_s - c),\\ &= (K_s - c)_{\mathrm{LMC}}(P) + E_I(K_s - c) + E_R(K_s - c),\end{aligned} \tag{A4}$$

where c is some arbitrary magnitude, $(K_s - c)_{\mathrm{LMC}}(P)$ is the period–colour relation for the LMC, $E_R(K_s - c)$ is the residual reddening and $(K_s - c)_0$ is the dereddened colour accounting solely for the interstellar reddening, $E_I(K_s - c)$, using the Schultheis et al. (2014) maps evaluated at $s = 8.275$ kpc as described above and the extinction law from Sanders et al. (2022a). We then estimate the residual extinction term as a function of period by computing a running median of the relation

$$E_R(K_s - c)(P) = (K_s - c)_0 - (K_s - c)_{\mathrm{LMC}}(P). \tag{A5}$$

We convert the residual reddening, $E_R(K_s - c)(P)$, into a residual extinction $A_{K_s,R}(P)$ using a choice of extinction coefficient, $C_c \equiv A_{K_s,R}/E_R(K_s - c)$, which we will discuss shortly. At each period P, we also naturally expect some variation in the circumstellar extinction properties (possibly through some metallicity variation). In a similar way to the computation of the period–luminosity relation, we compute its spread using the spread in the residual extinction via running medians in period as

$$\begin{aligned}\sigma_{K_s,\mathrm{NSD}}(P)^2 = C_R^2 \Big[&\mathrm{Var}(K_s - c) - \sigma_{E_I}^2(K_s - c)\\ &- \Delta E_I^2(K_s - c) - \Delta^2(K_s - c)\Big],\end{aligned} \tag{A6}$$

where $\mathrm{Var}(K_s - c)$ is the variance of data and $\sigma_{E_I}(K_s - c)$ is the extinction uncertainty reported by Schultheis et al. (2014) evaluated at a distance of 8.275 kpc and $\Delta E_I(K_s - c)$ is half the difference between the extinction at $s = 6.275$ kpc and $s = 10.275$ kpc (approximately the expected distance range of the sample). The extinction spread term is subdominant so this assumption is not very important for the overall selection function calculation. $\Delta(K_s - c)$ is the spread of the colour arising from the variability of the stars. We use a mean K_s measurement so this is only significant for the magnitude c. The uncertainties in the measurements of K_s and c are small so we do not consider them as contributing to the scatter. We put a 0.13 mag floor on $\sigma_{K_s,\mathrm{NSD}}$ as this is the scatter in the period–luminosity relation for O-rich LMC stars using the LMCNISS data (Macri et al. 2015; Yuan et al. 2017b; Sanders 2023) although this choice of floor is not too important as it is much narrower than the width of $S(K_s)$ in equation (A2).

A2 Application of the selection function procedure

This procedure is illustrated in panels (a) and (b) of Fig. A1 using the selection of stars described at the start of Section 2 (not restricted to low latitude only). We choose $c = [4.5]$ as our reference magnitude. because it is available from *Spitzer*/GLIMPSE catalogues (Ramírez et al. 2008; Churchwell et al. 2009) for the majority of our sample, and because, as [4.5] is a redder band, the variability is small. Additionally, Sanders et al. (2022b) demonstrated the period–Wesenheit magnitude relations using K_s and [4.5] for the sample were similar to the LMC and solar neighbourhood samples (particularly for short-period stars). This may be because circumstellar extinction effects become minimal around 3 μm (Ita & Matsunaga 2011; Lewis et al. 2023). The GLIMPSE data set is built by averaging over two epochs. Comparison with the model-average *W*2 WISE measurements from Sanders et al. (2022b) suggests the scatter arising from variability, $\Delta(K_s - [4.5])$, is around 0.3 mag with a weakly increasing trend towards longer periods.

In our procedure, we must adopt an extinction law for the Mira variables to connect the residual extinction $A_{K_s,R}$ to the residual reddening $E_R(K_s - [4.5])$, $C_{[4.5]} \equiv A_{K_s,R}/E_R(K_s - [4.5])$. As discussed in the previous section, this residual extinction is likely a combination of interstellar and circumstellar extinction so this is not a simple exercise. There have been several investigations into the properties of circumstellar dust around Mira variables in both the Galactic bar-bulge and the LMC in the J, H and K_s bands (Matsunaga, Fukushi & Nakada 2005; Ita & Matsunaga 2011; Yuan et al. 2017a; Lewis et al. 2023) with most studies concluding the circumstellar extinction law in these bands differ from the interstellar extinction law (note that Matsunaga et al. 2005 argue the interstellar and circumstellar extinction laws are similar although only when considering shallower interstellar extinction laws now disfavoured for the Galactic bar-bulge region).

In panel (c) of Fig. A1, we show the ratio of the discrepancy of $[4.5]_0$ dereddened using the Schultheis et al. (2014) maps with respect to the LMC relation shifted to the Galactic Centre distance modulus relative to the same for K_s. This ratio should give the residual $A_{[4.5]}/A_{K_s}$ reddening. We find a relatively flat trend with a median around the expectation from interstellar extinction laws (Fritz et al. 2011; Sanders et al. 2022a). There is a slight trend for higher coefficients for shorter-period stars, but this could be due to the relative importance of circumstellar and residual interstellar extinction with period or due to the distance distribution of the sample. Assuming different mean distances to the sample produces strong trends in this plane giving confidence that both the residual extinction coefficient is near the interstellar expectation and the sample is mostly located around the Galactic Centre distance. This

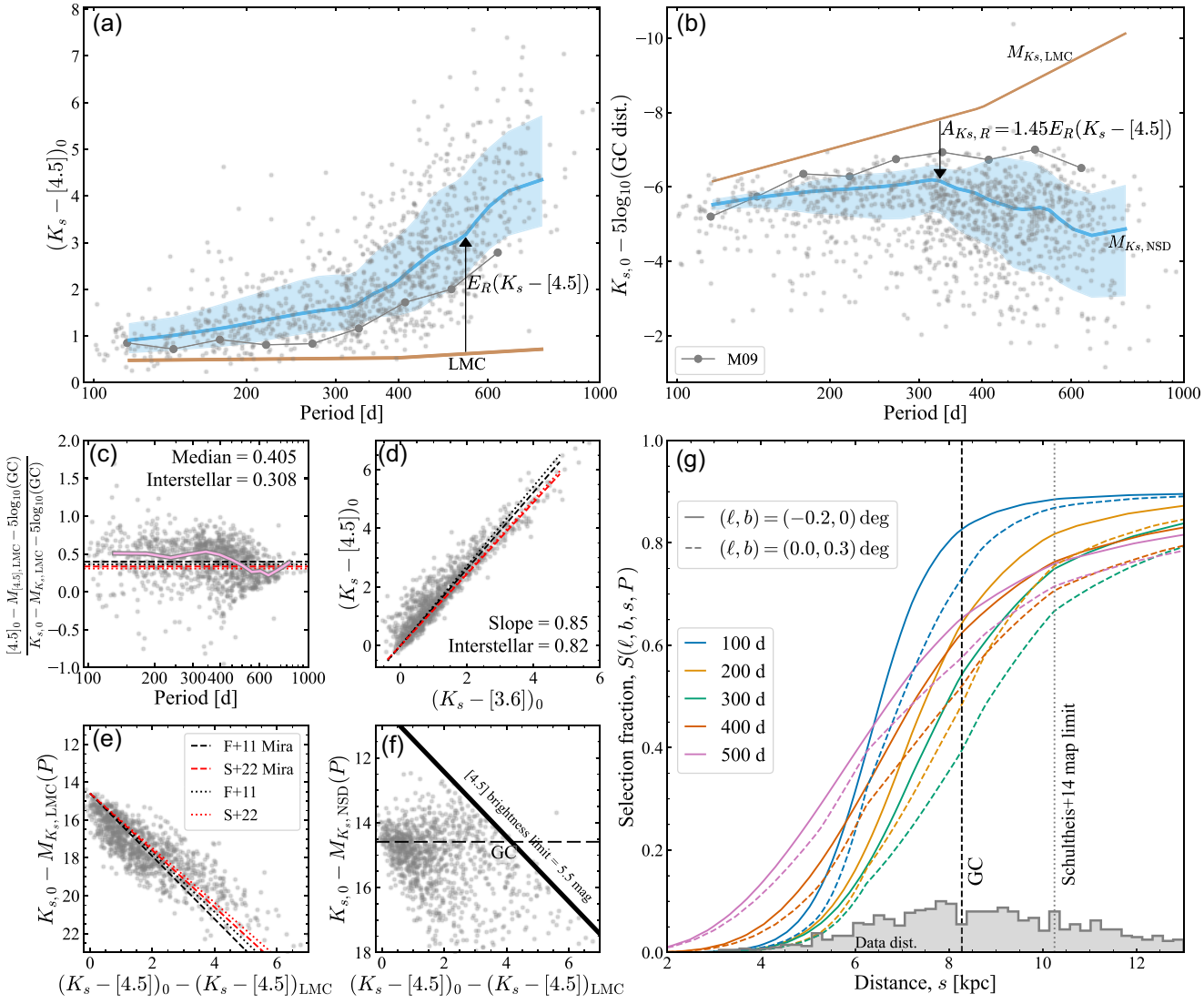

Figure A1. Illustration of the selection function calculation procedure. The grey points in all panels are the NSD Mira variable sample. Panel (a) shows the (K_s − [4.5]) colour corrected for interstellar extinction from the maps of Schultheis et al. (2014) versus period of the NSD Mira variable sample. The orange line is the measured relation for the LMC, and the cyan line and bracket show the median and $\pm 1\sigma$ for the NSD Mira variable sample. The grey line with larger grey points shows the median for (Matsunaga et al. 2009, M09) Mira variables with $K_s < 10.5$. Panel (b) shows similar but for dereddened absolute K_s (computed assuming all stars are at the Galactic Centre distance). The cyan line and bracket have been computed by adjusting the LMC period–luminosity relation (orange line) by the residual dust extinction as measured by the colour excess in panel (a) and removing the extinction scatter and multiepoch scatter for the bracket. Panel (c) shows the ratio of the residual reddening in [4.5] and K_s assuming all stars are at the Galactic Centre distance. The pink line is the median and the four horizontal lines are different interstellar extinction calculations as labelled in panel (e). Panel (d) shows (K_s − [4.5]) versus (K_s − [3.6]) dereddened by the Schultheis et al. (2014) maps with the interstellar extinction laws overplotted. Panel (e) shows K_s dereddened by the Schultheis et al. (2014) maps minus the absolute magnitude relation for the LMC, $M_{K_s,\mathrm{LMC}}$ [as displayed in panel (b)], against the colour excess from panel (a). Panel (f) shows the distance modulus computed from K_s dereddened by the Schultheis et al. (2014) maps and the NSD period–luminosity relation, $M_{K_s,\mathrm{NSD}}$, from panel (b) against the colour excess from panel (a). Finally, panel (g) shows the selection fraction, $S(\ell, b, s, P)$, with distance for two representative on-sky positions (distinguished by the line styles) and five periods (distinguished by the colours). Shorter-period stars are fainter so less affected by saturation. The Galactic Centre distance is marked along with the maximum distance resolved by the extinction maps of Schultheis et al. (2014). The distance distribution of the NSD Mira variable sample is shown assuming $M_{K_s,\mathrm{NSD}}$ from panel (b).

is further corroborated by panel (d) showing $(K_s - [4.5])_0$ versus $(K_s - [3.6])_0$ dereddened by the Schultheis et al. (2014) maps. This distribution has a slope very similar to the interstellar extinction law (within a few per cent) and panel (e) which shows the dereddened residual with respect to $M_{K_s,\mathrm{LMC}}$ (not correcting for an assumed distance modulus) versus the colour residual with respect to the LMC relation. The sample is consistent with being around a fixed distance and reddened according to an extinction law approximately similar to the interstellar extinction law. If we assume all of the residual extinction arises from circumstellar extinction, we confirm the result of Lewis et al. (2023) that the circumstellar extinction law has a smaller A_H/A_{K_s} than the interstellar extinction law for this sample. However, it appears that around the K_s, [3.6] and [4.5] wavelengths the two extinction laws are quite similar. Using the

interstellar extinction coefficient to relate $A_{K_s,R}$ to $E_R(K_s - [4.5])$ gives the blue line in panel (b). We see it neatly lies over the data distribution (which we assumed for the purpose of plotting is all located at the Galactic Centre distance). This is further evidenced in panel (f) showing the distance modulus assuming $M_{K_s,\mathrm{NSD}}$ versus the colour residual. The distribution is centred around the Galactic Centre distance with a bias arising from the lack of [4.5] measurements for stars brighter than 5.5 mag due to saturation. This bias does not affect our selection function as we are considering the selection in K_s.

It is advantageous that the residual extinction law is similar to the interstellar extinction law as it is then irrelevant how we portion out colour excess to each effect provided we do so consistently. The argument of the selection function is

$$\begin{aligned} K_s &= M_{K_s,\mathrm{LMC}}(P) + 5\log_{10}(100[s/\mathrm{kpc}]) \\ &+ C_{[4.5]}\Big[(K_s - [4.5]) - (K_s - [4.5])_{\mathrm{LMC}}(P)\Big] \\ &+ A_{K_s} - C_{[4.5]}E_I(K_s - [4.5]), \end{aligned} \tag{A7}$$

where if $C_{[4.5]} = A_{K_s}/E_I(K_s - [4.5])$ the interstellar part cancels. This is also true in the handling of the scatter term although one could consider making different assumptions about the scatter in colour excess arising from interstellar extinction variation along a given line-of-sight which might weakly affect the final scatter calculation. Multiplying the residual extinction coefficient $C_{[4.5]}$ by 0.9 to better match the distribution in panel (e) of Fig. A1 produces negligible changes to the selection function. A disadvantage of our procedure is that residual issues with interstellar extinction are not considered as a function of on-sky location, only period.

For our selection function approach to be valid, $M_{K_s,\mathrm{NSD}}(P)$ must reflect the period–magnitude relation of *all* Mira variables in the NSD region, not just those we have observed. This is naturally quite difficult to assess. In panels (a) and (b) we show $(K_s - [4.5])_0$ and $K_{s,0} - \mu_{\mathrm{GC}}$ both dereddened by the Schultheis et al. (2014) maps for the Matsunaga et al. (2009) sample, which contains many stars brighter than our sample and itself is subject to selection effects. As shown by Sanders et al. (2022b), the Matsunaga et al. (2009) sample is bluer and brighter than the sample inspected here suggesting our sample is not completely representative and we should be cautious extrapolating the properties of our Mira variable sample and $M_{K_s,\mathrm{NSD}}(P)$ to lower extinction regions.

With the period–luminosity relation, $M_{K_s,\mathrm{NSD}}(P)$, and its scatter, $\sigma_{K_s,\mathrm{NSD}}(P)$, defined, we can return to equation (A1) to evaluate the selection function. We convolve $S(K_s)$ with $\sqrt{\sigma^2_{K_s,\mathrm{NSD}} + \sigma^2_{A_{K_s}}}$ to account for the scatter in both the interstellar and residual extinction. The properties of the selection function are shown in Fig. A1. For fixed period, the selection function biases the sample towards more distant stars, particularly those in the background disc. Due to the scatter in the extinction, the differential selection effects across the NSD region are minimal, although this is likely also a reflection of the shortcomings of the low-resolution 3D extinction maps we are limited to using. When comparing different stars of different periods, we see that stars with periods around 300 day are most strongly affected by selection effects (this is approximately the turnover in $M_{K_s,\mathrm{NSD}}$ in panel b of Fig. A1) with shorter and longer period stars more readily detected in VVV.

The selection effects do not appear to significantly affect the conclusions of our work. However, there is certainly scope to further understand the interstellar and circumstellar extinction effects on this sample. This may be possible through a more detailed study of the Gaia sample located at higher latitudes and lower extinctions.

APPENDIX B: ADDITIONAL POSTERIOR PREDICTIVE CHECKS

In addition to posterior predictive checks split by period in the main body of the paper (see Fig. 6), we show in Fig. B1 the mock samples split by on-sky location. This demonstrates the validity of the model fits to the data and demonstrates the decreasing dominance of the NSD with increasing $|\ell|$ and $|b|$. There is weak evidence of the model underpredicting the peak in μ_b which could arise in part from missing background disc stars in the model (possibly from slight oversmoothing in the kernel density estimates).

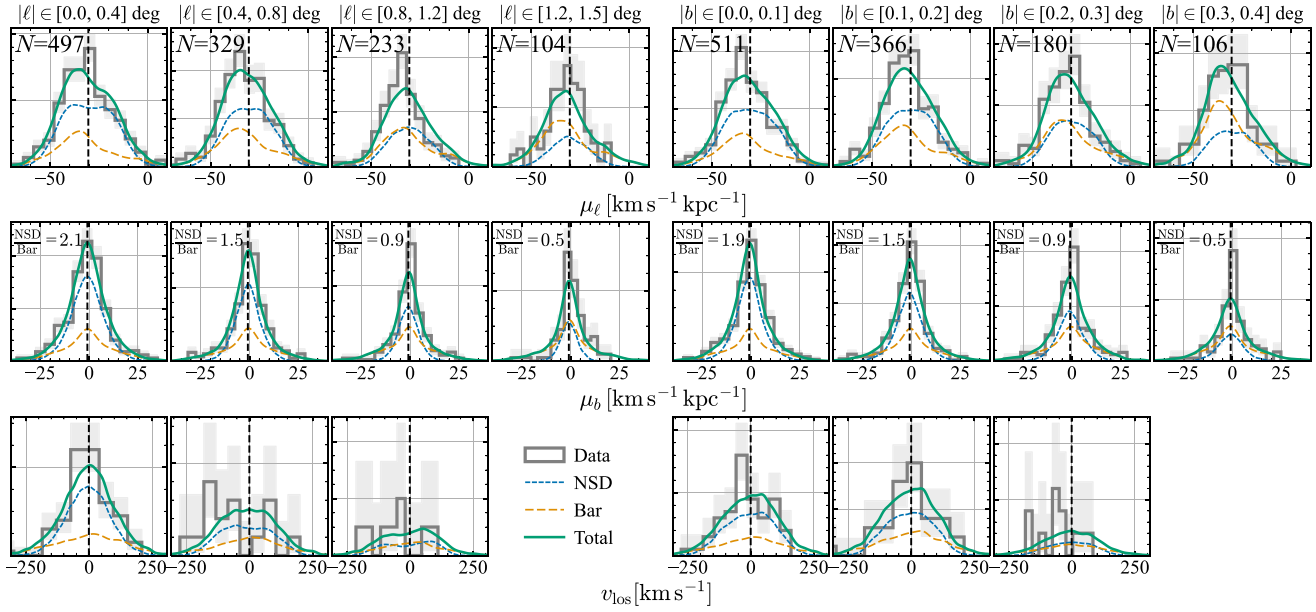

Figure B1. Velocity distributions for the data (grey with Poisson uncertainty shaded interval) compared to the best-fitting model split by Galactic longitude (left) and Galactic latitude (right). See the caption of Fig. 6 for more information.

APPENDIX C: RELATION TO BAR AGE ESTIMATES FOR EXTERNAL GALAXIES

We can connect our results to studies of measuring the bar formation epochs in other galaxies (Gadotti et al. 2015, 2018, 2020; De Sá-Freitas et al. 2023a, b). In Fig. C1, we display a mass-weighted age map of the inner few kpc as seen by an observer looking down on the disc of the Galaxy from infinity, the type of data one might obtain from an IFU. The NSD stands out as a slightly younger structure within the old bar. The difference in mean age is quite subtle (~ 1 Gyr) and only apparent within 2–3$R_{\rm NSD}$. De Sá-Freitas et al. (2023a, b) present a methodology for constraining a bar's age from IFU data. As in our methodology, the bar contamination within the NSD region is significant, so De Sá-Freitas et al. (2023b) give good evidence of the self-consistency of this method of estimating the bar age.

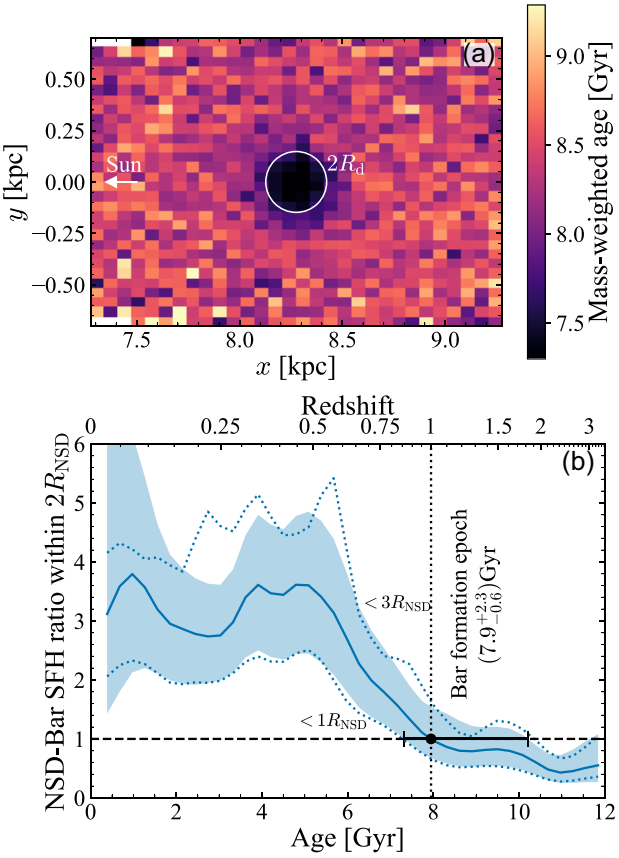

Figure C1. The Milky Way NSD as seen by an external observer: Panel (a) shows the mass-weighted age map of the inner regions of the Galaxy as seen by an observer located at infinity towards the Galactic North Pole. The white circle shows the region encompassing two scale lengths of the NSD and the Sun is located to the left outside the diagram. No extinction is considered. Panel (b) shows the ratio of the NSD to Bar star formation histories within a cylinder of $2R_{\rm NSD}$ (with uncertainty given by the shaded bracket and the two dotted lines the results within 1 and $3R_{\rm NSD}$. Redshift is computed using the cosmological parameters from Planck Collaboration (2020). Following De Sá-Freitas et al. (2023b), we define the bar formation epoch as the age at which the NSD is the dominant star-forming component within $2R_{\rm NSD}$ leading to the error bar and annotated age for the Milky Way bar as inferred by an outside observer.

This paper has been typeset from a TEX/LATEX file prepared by the author.