Tracing the Milky Way warp and spiral arms with classical Cepheids

B. Lemasle, H. N. Lala, V. Kovtyukh, M. Hanke, Z. Prudil, G. Bono, V. F. Braga, R. da Silva, M. Fabrizio, G. Fiorentino, et al.

Astronomy & Astrophysics

Tracing the Milky Way warp and spiral arms with classical Cepheids

B. Lemasle[1], H. N. Lala[1], V. Kovtyukh[2], M. Hanke[1], Z. Prudil[1], G. Bono[3,4], V. F. Braga[4,5], R. da Silva[4,5], M. Fabrizio[4,5], G. Fiorentino[4], P. François[6,7], E. K. Grebel[1], and A. Kniazev[8,9,10]

[1] Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchhofstr. 12-14, 69120 Heidelberg, Germany
e-mail: `lemasle@uni-heidelberg.de`
[2] Astronomical Observatory, Odessa National University, Shevchenko Park, 65014 Odessa, Ukraine
[3] Dipartimento di Fisica, Universitá di Roma Tor Vergata, Via della Ricerca Scientifica 1, 00133 Rome, Italy
[4] INAF – Osservatorio Astronomico di Roma, Via Frascati 33, Monte Porzio Catone 00078, Rome, Italy
[5] Agenzia Spaziale Italiana, Space Science Data Center, Via del Politecnico snc, 00133 Rome, Italy
[6] GEPI, Observatoire de Paris, CNRS, Université Paris Diderot, Place Jules Janssen, 92190 Meudon, France
[7] UPJV, Université de Picardie Jules Verne, 33 Rue St. Leu, 80080 Amiens, France
[8] South African Astronomical Observatory, PO Box 9, 7935 Observatory, Cape Town, South Africa
[9] Southern African Large Telescope Foundation, PO Box 9, 7935 Observatory, Cape Town, South Africa
[10] Sternberg Astronomical Institute, Lomonosov Moscow State University, Universitetskij Pr. 13, Moscow 119992, Russia

ABSTRACT

Context. Mapping the Galactic spiral structure is a difficult task since the Sun is located in the Galactic plane and because of dust extinction. For these reasons, molecular masers in radio wavelengths have been used with great success to trace the Milky Way spiral arms. Recently, *Gaia* parallaxes have helped in investigating the spiral structure in the Solar extended neighborhood.
Aims. In this paper, we propose to determine the location of the spiral arms using Cepheids since they are bright, young supergiants with accurate distances (they are the first ladder of the extragalactic distance scale). They can be observed at very large distances; therefore, we need to take the Galactic warp into account.
Methods. Thanks to updated mid-infrared photometry and to the most complete catalog of Galactic Cepheids, we derived the parameters of the warp using a robust regression method. Using a clustering algorithm, we identified groups of Cepheids after having corrected their Galactocentric distances from the (small) effects of the warp.
Results. We derived new parameters for the Galactic warp, and we show that the warp cannot be responsible for the increased dispersion of abundance gradients in the outer disk reported in previous studies. We show that Cepheids can be used to trace spiral arms, even at large distances from the Sun. The groups we identify are consistent with previous studies explicitly deriving the position of spiral arms using young tracers (masers, OB(A) stars) or mapping overdensities of upper main-sequence stars in the Solar neighborhood thanks to *Gaia* data.

Key words. stars: variables: Cepheids – Galaxy: disk – Galaxy: structure

1. Introduction

Cepheids are massive or intermediate-mass pulsating variable stars; they are well-known as a calibrator of the extragalactic distance scale via their period-luminosity (PL) relations. Their ages range from a few tens to a few hundreds of megayears, which makes Cepheids excellent tracers of young stellar populations, for instance, in the Milky Way disk (e.g., Lemasle et al. 2013; Ripepi et al. 2021), in the Magellanic Clouds (e.g., Lemasle et al. 2017; Romaniello et al. 2022), and in nearby dwarf irregular galaxies (e.g., Neeley et al. 2021).

The Solar System is located close to the Milky Way plane ($z_\odot = 20.8^1$ pc, Bennett & Bovy 2019), at $R_\odot = 8.275$ kpc from the Galactic center (GRAVITY Collaboration 2021). Given the high extinction in the plane, mapping the Milky Way spiral structure is a difficult task. Maser sources associated with young

massive stars in high-mass star-forming regions are among the most reliable tracers since they are very young and their parallaxes can be measured with radio-interferometry (Reid et al. 2019). Although radio-interferometric measurements are limited to a couple hundred sources, they present the strong advantage of being unaffected by extinction, and, therefore, of tracing the spiral structure at large distances from the Sun. On the basis of these measurements, Reid et al. (2019) found that the Milky Way spiral structure consists of four arms, plus the Local arm, which they consider to be an isolated segment. However, alternative models exist, for instance, a two-(major)-arm model by Drimmel (2000). Hou & Han (2014) demonstrate the difficulty in determining the number of spiral arms in the Milky Way. In this context, *Gaia* constrained the location of several spiral arms in the fourth Galactic quadrant: within ≈5 kpc from the Sun, where its parallaxes remain accurate enough, it provided distances for thousands of OB(A) (Chen et al. 2019; Xu et al. 2021; Poggio et al. 2021; Zari et al. 2021) upper main-sequence stars or young open clusters (Castro-Ginard et al. 2021;

---

[1] We note in passing that $z_\odot$ values based on hydrogen radio emission are often smaller (e.g., ~4 pc, Blaauw et al. 1960) than those based on stellar tracers (see also Bland-Hawthorn & Gerhard 2016).

Hao et al. 2021; Monteiro et al. 2021), all enabling various teams to trace the spiral arms in the extended Solar neighborhood. In this paper we take advantage of the accurate distances of classical Cepheids to investigate the spiral structure of the Milky Way. Since some of these Cepheids are very distant, we need to take the Galactic warp into account.

The paper is organized as follows: in Sect. 2, we briefly explain how our catalog of classical Cepheids was gathered. In Sect. 2.3, we take advantage of the catalog (including *Gaia* Early Data Release 3 (EDR3) data, Gaia Collaboration 2021a) to determine new period-luminosity and period-Wesenheit relations in the WISE bands, and to derive a homogeneous set of distances for the Cepheids. In Sect. 3, we examine the properties of the Galactic warp. In Sect. 4 we inspect the Milky Way spiral arms as traced by classical Cepheids, and their impact on abundance gradients is discussed in Sect. 5. Section 6 provides our summary and conclusions.

2. The catalog

2.1. Input data: Variability catalogs

We have built a comprehensive catalog of pulsating variable stars, gathering the data published by many photometric surveys dedicated to variability, or having at least some time-domain capabilities. They are listed below. One of our main sources of classical Cepheids is the Optical Gravitational Lensing Experiment (OGLE) survey, which provides a two-decade-long monitoring of the Magellanic Clouds (Soszyński et al. 2015, 2017a) and of the Milky Way bulge and disk (Soszyński et al. 2017b, 2020; Udalski et al. 2018). The data from the *Gaia* satellite are an amazing all-sky tool to discover and monitor variable stars (Clementini et al. 2019). Due to the small number of observations covering their light curves, some variables were misclassified in *Gaia* DR2 (see Lemasle et al. 2018, for instance), which led Ripepi et al. (2019) to reclassify the *Gaia* DR2 Galactic classical Cepheids, which we also added to our list of Galactic Cepheids. The All-Sky Automated Survey for Supernovae (ASAS-SN, Jayasinghe et al. 2018, 2019b,a) surveys the entire sky down to $V \approx 18$ mag using a network of 24 small telescopes, making it particularly useful to follow (among others) the bright variables that are inaccessible to other surveys due to saturation. ASAS-SN draws from the All Sky Automated Survey (ASAS, Pojmanski 1997, 2002), and we used additional data from the machine-learned ASAS Classification Catalog (MACC, Richards et al. 2012). For bright Cepheids, we also used the list of classical Cepheids[2] provided by Skowron et al. (2019a). The Zwicky Transient Facility (ZTF, Chen et al. 2020), a recent time-domain survey, provided a large number of new targets, especially in the Northern Hemisphere, in general with an extremely good sampling of the period. Finally, the Wide-field Infrared Survey Explorer (WISE, Chen et al. 2018) discovered a large number of pulsating stars, and its observing window in the mid-infrared provides exquisite distances via period-luminosity or period-Wesenheit relations in a spectral domain where extinction is minimal. Since the classification of variable stars in the infrared is a challenging task (due to the similarity of the light curves of different classes of variable stars), we retained only those stars that could be identified as a classical Cepheid in at least one optical photometric survey.

2.2. Merging the data and quality control

Merging and quality control are important steps when collecting such an amount of data from various sources. We followed the procedure established for RR Lyrae and Type II Cepheids detailed in Lala et al. (in prep.). The catalog of classical Cepheids will be published in a forthcoming paper, which will also provide more details on the quality control process[3]. The main points are briefly listed below:

– Stars from each individual survey have been robustly cross-matched against *Gaia* EDR3, in order to recover their astrometry, and their photometry in the *Gaia* bands (if available). Stringent quality cuts regarding large-scale systematics, binarity, or crowded regions using keywords such as `ruwe`, `ipd_gof_harmonic_amplitude`, `astrometric_excess_noise`, etc were applied to *Gaia* astrometric and photometric data following the recommendations of Fabricius et al. (2021), Lindegren et al. (2021), and Riello et al. (2021).

– To avoid stars with spurious astrometric solutions, discarding stars with a fractional parallax uncertainty >0.15 (6.6σ) is common practice. However, we compared the parallax-based distances to those obtained from various period-luminosity or period-Wesenheit relations. The agreement is obviously better for stars with fractional uncertainties <0.15 (median difference ~0.2 kpc), but the agreement is still good for stars with fractional uncertainties between 3σ and 6.6σ (median difference ~0.6 kpc)[4]. Therefore, we considered stars with a fractional parallax uncertainty <0.33. We note that this cut was not applied when deriving period-Wesenheit (PW) relations as it would bias the input sample. With such a criterion, we loose only a few tens nearby Cepheids that passed the previous quality cuts.

– The classification in different subclasses of pulsating variables and the exact value of the period has been taken from OGLE, and if not available, then from other surveys. If a star was observed in more than one survey, priority was given to the survey with the largest number of data points in the light curve. The classification and the value of the period are in general an excellent match between different surveys, and most of the confusion arises because the very nature of some of the surveys does not allow them to discriminate, for instance, anomalous Cepheids, or stars pulsating in several modes simultaneously.

– In addition, we checked for aliased periods (when the coverage of the light curve is inadequate, the recovered period may be a multiple of the true period). Similarly, candidate variables with a period matching exactly the terrestrial rotation period have been removed.

– For surveys observing in the same photometric bands, we checked for possible zero-point offsets between their photometry and found them to be negligible, with the exception of the MACC catalog.

2.3. Period-Wesenheit relations and distances

We recovered unWISE photometry ($W1, W2$) for this catalog from IRSA[5] (Meisner et al. 2021). unWISE photometry consists of all-sky static coadds based on six years of WISE and NEOWISE operations, in contrast to the one-year ALLWISE

---

[2] https://www.astrouw.edu.pl/ogle/ogle4/OCVS/allGalCep.listID

[3] And a comparison with the new list of classical Cepheids from the OGLE team (Pietrukowicz et al. 2021).

[4] The extent of the agreement varies from star to star depending on the period-luminosity or period-Wesenheit relation used in the comparison.

[5] https://irsa.ipac.caltech.edu/cgi-bin/Gator/nph-scan?utf8=%E2%9C%93&mission=irsa&projshort=WISE
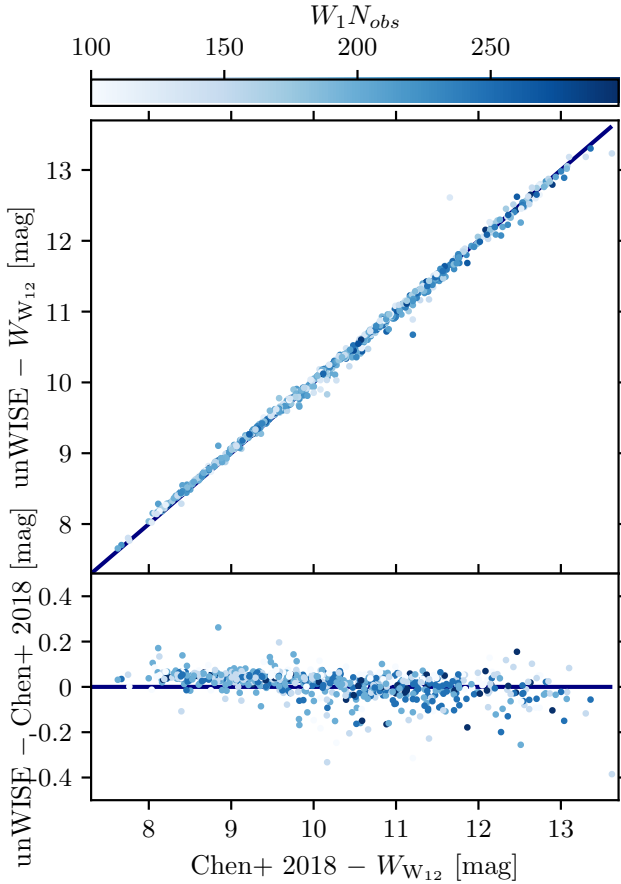
Fig. 1. Comparison between the unWISE photometry (static coadds, 6 yr of operations) and the WISE photometry in Chen et al. (2018) (light-curve fitting, 5 years of operations).

Table 1. Period-Wesenheit relations in WISE bands for fundamental mode (DCEP_F) and first-overtone (DCEP_1O) classical Cepheids located in the LMC.

Type	n	α	β	σ
DCEP_F	2326	-2.436 ± 0.013	-3.196 ± 0.019	0.149 ± 0.005
DCEP_1O	1591	-2.936 ± 0.015	-3.342 ± 0.037	0.181 ± 0.008

Notes. n denotes the total number of stars used to compute the law. The mean and standard deviations of posterior distributions of the model parameters are presented in the last three columns.

Table 2. Covariance matrix for the Bayesian robust regression of the period-Wesenheit relations.

	α	β	σ	ν
α	3.24967949e−04	−4.38328130e−04	1.46899174e−03	−6.11806477e−05
β	−4.38328130e−04	6.77959533e−04	−2.16090433e−03	2.92371896e−05
σ	1.46899174e−03	−2.16090433e−03	1.52551182e+00	1.10784521e−03
ν	−6.11806477e−05	2.92371896e−05	1.10784521e−03	2.26928615e−03
α	3.08465117e−04	−6.95175096e−04	−2.30024957e−04	−1.07183562e−04
β	−6.95175096e−04	2.01753200e−03	2.93887199e−03	1.97729713e−04
σ	−2.30024957e−04	2.93887199e−03	1.38204481e+00	2.73979159e−03
ν	−1.07183562e−04	1.97729713e−04	2.73979159e−03	2.53117985e−03

Notes. α and β are the zero-point and the slope of the relations, σ their standard deviation and ν their normality parameter. The *upper panel* is for fundamental mode classical Cepheids and the *lower panel* for first-overtone classical Cepheids.

data release (Cutri et al. 2013). Besides the longer observation baseline, unWISE photometry also holds the advantage of being more robust in crowded regions, on account of using crowdsource (Schlafly et al. 2019) cataloging software. Figure 1 compares unWISE photometry with that of the Chen et al. (2018) catalog. The latter presented an all-sky variable star catalog based on five years of WISE and NEOWISE data and their magnitudes were determined by Fourier-fitting the light curves. The greater number of measurements, coupled with the fact that the pulsation amplitude of classical Cepheids is roughly 0.2 mag (e.g., Chen et al. 2018) in mid-IR wavelenghts, results in unWISE photometry agreeing excellently with Fourier-fitted mean magnitudes. We obtained $W1, W2$ unWISE photometry for 3260 Cepheids (after taking all the processing flags[6] into account).

To determine distances using unWISE photometry, we computed new period-Wesenheit relations in WISE bands. We created a catalog of LMC classical Cepheids, similar to the one described here for the Milky Way. The apparent Wesenheit $W_{\rm W12}$ was calculated as:

$$W_{\rm W_{12}} = W2 - R_{\rm W_2,W_1} * (W1 - W2). \tag{1}$$

The multiplicative constant $R_{\rm W_2,W_1}$ is the total-to-selective extinction ratio:

$$R_{\rm W_2,W_1} = \frac{A_{\rm W_2}}{E(W1 - W2)}, \tag{2}$$

with $R_{\rm W_2,W_1} = 2.0$ (Wang & Chen 2019). Absolute Wesenheits were calculated using the LMC distance modulus (18.477 mag) from Pietrzyński et al. (2019). We used pymc3 (Salvatier et al. 2016) to perform a Bayesian robust regression (as described in Sect. 3.2) on the following model:

$$W_{\rm abs} \sim \mathcal{T}\left(\alpha + \beta \times \log_{10}\left\{\frac{\rm Period}{\rm 1\,d}\right\}, \sigma^2, \nu\right), \tag{3}$$

where α and β, the intercept and the slope of the model, are assumed to follow a normal distribution, while σ, the intrinsic scatter, is assumed to follow a half-normal distribution and ν, the normality parameter (degrees of freedom), is assumed to follow a Gamma distribution (Juárez & Steel 2010). $W_{\rm abs}$ is the absolute Wesenheit and we assume that the likelihood of the model follows a Student's T-distribution. The model parameters (and their uncertainties) for the PW relations in WISE bands are given in Table 1. The covariance matrix is provided in Table 2. From these relations, we obtained distances precise up to 3% (5%) in the low- (high-)extinction regions. The Wesenheit pseudo-magnitudes are unaffected by the extinction toward individual stars (or their uncertainties) by construction (Madore 1982). Their dependence on reddening only lies in the accuracy and potential nonuniversality of the $R_{\rm W_2,W_1}$ ratio.

In what follows, we used only Cepheids pulsating in the fundamental (F) or the first overtone (1O) mode: they are, by far, the most numerous, and Cepheids in other subclasses have slightly less accurate distances since their period-Wesenheit relations are calibrated with a smaller number of stars. We used distances determined using the best WISE data, as described above, for 2098 Cepheids for which such photometry was available. If not (586 Cepheids), we used as distance the inverse of the *Gaia* EDR3 parallax, provided this distance is less than 5 kpc. Uncertainties on the astrometric or photometric distances have been

propagated throughout the paper. This restricted catalog contains 2684 Cepheids (F,1O), while the catalogs of Chen et al. (2019) and Skowron et al. (2019b) contained 1339 and 2390 Cepheids, respectively.

3. The Galactic warp as traced by classical Cepheids

3.1. The Galactic warp

The Galactic warp (Kerr 1957; Oort et al. 1958) is a large-scale distortion of the Milky Way disk. It is caused by a torque exerted on the disk, whose origin has been suggested to result either from a misalignment between the rotation axis of the disk and of the halo (e.g., Sparke & Casertano 1988; Debattista & Sellwood 1999), or from the inner disk (e.g., Chen et al. 2019), or from material accreted to the halo (e.g., Ostriker & Binney 1989; Jiang & Binney 1999), or from tidal perturbations associated with nearby Milky Way satellites such as Sagittarius (e.g., Ibata & Razoumov 1998; Laporte et al. 2019) and the Magellanic Clouds (e.g., Weinberg & Blitz 2006; Garavito-Camargo et al. 2019).

Different tracers have been used to map the warp, from neutral hydrogen (e.g., Henderson et al. 1982) to molecular clouds (e.g., Wouterloot et al. 1990) and star counts (e.g., Reylé et al. 2009; Amôres et al. 2017). Individual stellar classes, namely OB stars (e.g., Miyamoto et al. 1988; Reed 1996; Yu et al. 2021), RGB or red clump stars (e.g., López-Corredoira et al. 2002; Momany et al. 2006; Wang et al. 2020), and even pulsars (e.g., Yusifov 2004) have also been used. It has been shown in the Milky Way (e.g., Chen et al. 2019; Romero-Gómez et al. 2019) and in external galaxies (e.g., Radburn-Smith et al. 2014) that only the youngest stellar objects accurately map the Galactic warp. The mismatch between the warp as traced by hydrogen and young stellar objects, and the one traced by older stars, suggests a large value for the warp's precession, leading, for instance, Poggio et al. (2020) and Cheng et al. (2020) to favor the tidal perturbation scenario. On the other hand, Chrobáková & López-Corredoira (2021) found no evidence of a warp precession.

Cepheids are an ideal tracer for studying the present-day warp because they are easily distinguished from other types of stars, and because their distances can be derived individually with great accuracy, even at large distances (>5 kpc) where *Gaia* parallaxes become uninformative. Chen et al. (2019), Skowron et al. (2019a,b) have used distances derived from mid-infrared (*Spitzer*: Benjamin et al. 2003; Churchwell et al. 2009, WISE: Chen et al. 2018) or near-infrared photometry (2MASS: Skrutskie et al. 2006). Dékány et al. (2019) traced the warp in highly reddened regions covered by the VVV survey. All these studies enabled the tracing of the Galactic warp and they confirmed that the Cepheids' warp follows the H I warp. Moreover, Skowron et al. (2019b) showed that the northern part of the warp is very prominent, with an amplitude 10% larger than that of the southern warp.

These studies differ regarding the analytical formula adopted for the warp, the input catalog of Cepheids (different surveys have different completeness and contamination levels), and the photometry used to compute the Cepheids' distances. With respect to those studies, we benefit from updated WISE data, and our sample is larger (by several hundreds of stars) without sacrificing purity. We adopt the definition of the warp proposed by Skowron et al. (2019b): the Milky Way warp starts at a given radius r_0 and its shape follows the

Table 3. Parameters of the Galactic warp as derived using a robust regression method.

	Mean	σ
r_0 (kpc)	4.8626	0.3136
z_0 (pc)	13.0	4.9
z_1 (pc)	8.9	0.6
z_2 (pc)	1.4	0.3
θ_1 (deg)	-13.48	1.92
θ_2 (deg)	-26.27	5.60
σ	0.052	0.003

equation:

$$z(r, \Theta) = \begin{cases} z_0 & r < r_0 \\ z_0 + (r - r_0)^2 \times [z_1 \ \sin(\Theta - \Theta_1) + z_2 \ \sin(2\,(\Theta - \Theta_2))] & r \geq r_0 \end{cases} \quad (4)$$

where z is the vertical distance from the Galactic plane, r is the distance from the Galactic center, and Θ is the Galactocentric azimuth. $\Theta = 0°$ points in the "Sun to Galactic center" direction, while $\Theta = 180°$ points toward the Galactic anticenter. Θ increases counterclockwise if the Galaxy is seen from above. We adjusted the radial, vertical, and angular parameters r_0, z_0, z_1, z_2, Θ_1, Θ_2, using a Bayesian robust regression method.

3.2. Tracing the warp

We estimated the parameters of the warp model using Bayesian robust regression. As mentioned above, we assume the warp formula by Skowron et al. (2019b) given in Eq. (4) for the likelihood of our model. We assume a Student's t-distribution for the likelihood of our model, as it is much less sensitive to outliers and, therefore, provides more robust estimates of the parameters than those from a normal distribution (in presence of outliers, other approaches often shift the mean toward the outliers and increase the standard deviation, while the Student's t-distribution decreases the weight of the outliers). We use the Hamiltonian MCMC sampler (Betancourt 2017) of `pymc3` to sample the posterior distribution. Uncertainties on (r, Θ, z) and their covariances have been propagated from the uncertainties and covariances on right ascension, declination, and distances using Jacobian matrices (ESA 1997; Price-Whelan 2017) to perform coordinates transformations (the uncertainties on the position of the Sun relative to the Galactic center have been ignored).

We run the analysis in two steps: first we assume a normal distribution for the priors, adopting for the mean value and the standard deviation $r_0 = 5 \pm 2$ kpc, $z_0, z_1, z_2 = 0.5 \pm 1$ kpc, $\Theta_1, \Theta_2 = \pi \pm \pi$ rad. From the posterior distributions of this first stage, we use the mean values and 10× the standard deviations as the input (normal) priors for the second stage. In both steps, we run 4 chains and use the first 10 000 samples of each chain to tune the multidimensional posterior and accept the next 5000 draws as our posterior distribution (we checked that the auto-correlation is low for each individual chain). From the posterior distributions obtained in the second step, we adopt the mean values and the standard deviations as the parameters of the warp and their uncertainties, respectively. They are listed in Table 3. The posterior distributions of the parameters, as well as the standard deviation of the regression model σ are shown in Fig. A.2, which was drawn using the ArviZ package (Kumar et al. 2019). The full covariance matrix is provided in Table A.1.
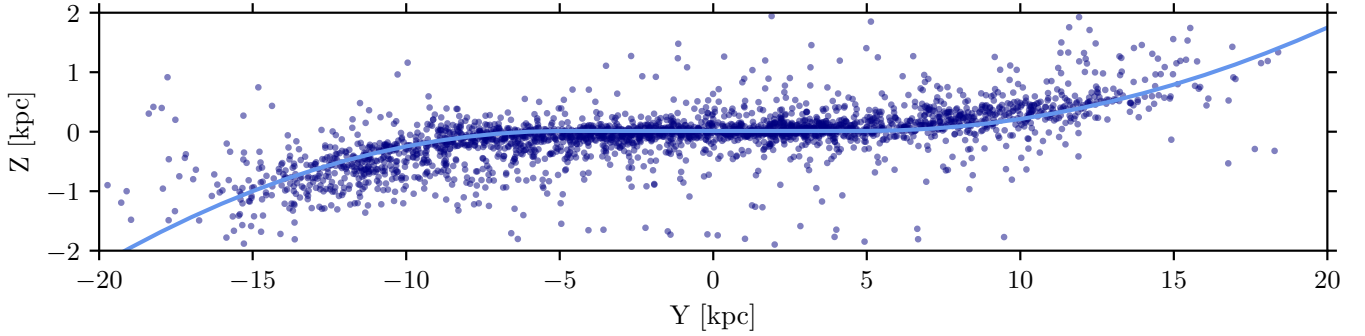
Fig. 2. Milky Way warped disk computed with the parameters obtained from the robust regression method (light blue) for $X = 0$ kpc. Individual Cepheids are over-plotted in dark blue, with Galactocentric distances computed using WISE mid-infrared photometry and PW relations. See also Fig. A.3.

The model of the warp computed with these parameters is displayed in Fig. 2, where the scale of the vertical axis is strongly enhanced. This figure, as well as Fig. A.3, shows that our model reproduces closely the vertical distribution of the Galactic Cepheids. The warp is more pronounced than the one provided by Chen et al. (2019), and this was already noted by Skowron et al. (2019b). The reason is simply that both our study and the one by Skowron et al. (2019b) rely on a larger number of Cepheids covering the four Galactic quadrants (although the sample is clearly incomplete in Q1 and Q4), while the Cepheids in Chen et al. (2019) mostly belong to Q2 and Q3.

The onset radius r_0 of the warp in our study is in fairly good agreement with the value reported by Skowron et al. (2019b) (4.86 ± 0.31 kpc vs. 4.23 ± 0.12 kpc). We note that if the formal value of the onset radius of the warp is small, the influence of the warp on the vertical position of Cepheids starts to be noticeable only at roughly the Solar radius.

The vertical parameters z_1 and z_2 are very similar in both studies: ~7 and ~1 pc in our case vs ~8 and ~2 pc for Skowron et al. (2019b). Our value for z_0 (~26 pc) is smaller than the ~44 pc reported by Skowron et al. (2019b), but in good agreement with recent literature values, for instance, the G+early K stars in the nearby sample of Gaia Collaboration (2021b).

Due to differences in their definition, Galactocentric azimuths are shifted by 180° between our work and Skowron et al. (2019b). Our value of $\Theta_1 = -13.48°$ must then be compared to their $\Theta_1 = 158.3 - 180 = -21.7°$, while the values of Θ_2, $-26.27°$ and $-13.6°$, respectively can be compared directly given the factor 2 in the second sine term of Eq. (4). The angular values are not exact matches, they remain however close to each other and lead to a very similar description of the Galactic warp.

It is not a surprise that our model resembles the warp model by Skowron et al. (2019b): we adopted their analytical relation, and even if we added a few hundreds of stars, the overall coverage of the disk is similar. However, the technique to derive the warp parameters is completely independent[7]. We believe that the small differences originate from this somewhat larger stellar sample combined to the updated WISE photometry. Moreover, uncertainties on coordinates and distances are included in the fitting procedure in our study. Although not formally compatible with those of Skowron et al. (2019b), the uncertainties on the warp parameters are extremely small. Considering the spa-

tial extent of the disk, they do not impact the determination of the Cepheids' Galactocentric distance.

Our results are also in excellent agreement with studies of the warp using H I data. For instance, Nakanishi & Sofue (2003) found that the warping in H I is the strongest for $\theta = +80°$ and $\theta = +260°$. We find similar angular values (see Fig. A.3). They report that the warping starts at $R_G = 10-12$ kpc and reaches ~1.5 kpc at $R_G = 16$ kpc ($\theta = +80°$) and ~-1 kpc at $R_G = 16$ kpc ($\theta = +260°$). Levine et al. (2006a) focused on the H I outer disk, well beyond the stellar disk. They found the maximum warping at $\theta = +90°$ and $\theta = 270°$. The height of the warp reaches ~4 kpc at $R_G = 22$ kpc and ~5.5 kpc at $R_G = 28$ kpc. At $R_G = 16$ kpc, they report a maximum height of ~1.3 kpc and a maximum depth of ~-0.8 kpc. These values, later confirmed by Kalberla et al. (2007), are slightly lower than those provided earlier by Nakanishi & Sofue (2003) and in better agreement with our own findings. Levine et al. (2006a) found that the H I warp can be approximated with a superposition of three vertical harmonics of the disk. Interestingly, these modes grow linearly in the outer disk. The mode $m = 1$, which dominates the warp for the range of Galactocentric distances covered by our Cepheids' sample, is linear between $R_G = 10$ kpc and $R_G = 25$ kpc with a slope of 0.197 kpc kpc^{-1}, which leads to a height of $\simeq 1$ kpc at $R_G = 15$ kpc. The other two modes do not influence the warp below $R_G = 15$ kpc, and they grow linearly with similar slopes until $R_G = 22$ kpc.

A comparison of the warp altitudes above and below the Galactic planes reached by various tracers at a Galactocentric distance of $R = 14$ kpc is shown in Fig. 3. It indicates that the warp becomes more pronounced for older tracers, as already mentioned, for instance, by Romero-Gómez et al. (2019).

3.3. Unwarping the Milky Way

In abundance gradient studies, the abundance of a given element is plotted against the radial Galactocentric distance of the stars composing the sample, or against their distance to the Galactic plane. However, in regions where the disk is strongly warped, the Galactocentric distances of stars end up being shorter than if the star was located in a nonwarped Milky Way disk. Because of the scarcity of spectroscopic data, very few studies have analyzed the spatial distribution of abundances (see Kovtyukh et al. 2022, for instance). Instead, it has been customary to collate Cepheids located at different Galactocentric azimuths in a unique 2D plane, where [Fe/H] is displayed as a function of the Galactocentric distance of the Cepheids in the sample. This necessary shortcoming implies ignoring the warping of the disk. It might
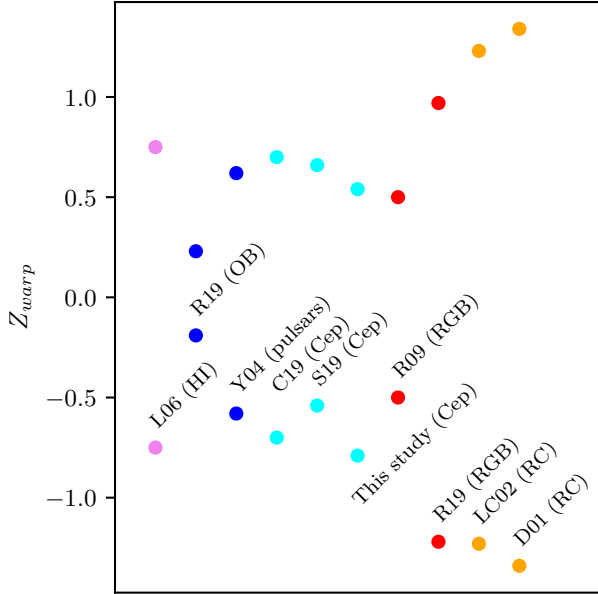
---

[7] To derive the warp parameters, Skowron et al. (2019a,b) simply mention that they minimize the sum of squares of orthogonal distances between individual Cepheids and the model, with the squared distances modulated by an exponential term penalizing outliers.

Fig. 3. Values of the warp altitude above and below the Galactic plane at the Galactocentric radius $R = 14$ kpc for different tracers ordered by approximate age. The values have been taken from Table 1 in Romero-Gómez et al. (2019) or computed by us. The tracers used to investigate the warp cover neutral hydrogen H I (Levine et al. 2006b, L06), OB stars (Romero-Gómez et al. 2019, R19), pulsars (Yusifov 2004, Y04), Cepheids (Chen et al. 2019, C19), (Skowron et al. 2019b, S19), this study, RGB (Reylé et al. 2009, R09), (Romero-Gómez et al. 2019, R19), and red clump (RC) stars (Drimmel & Spergel 2001, D01), (López-Corredoira et al. 2002, LC02).

be (at least partially) responsible for the increased dispersion of abundances in the outer disk since it brings together Cepheids located in warped and nonwarped regions. In order to investigate this issue, we have compared the length of a bow between the Galactic Center and a Cepheid located at the Galactocentric distance d_{GC} with d_{GC} itself. The details of the calculation are given in Appendix A.1. As can be seen in Fig. 4, the difference between Galactocentric distances computed on a flat and on a warped disk are negligible below 10 kpc. They can reach 100 pc at $R_G = 18$ kpc, but only for the stars located in regions where the warp is pronounced, while other Cepheids are barely affected. The differences become larger only in the outer regions of the disk, where only a small number of Cepheids have been reported until now. They remain, however, too small to explain the larger dispersion around the mean metallicity gradient reported, for instance, by Genovali et al. (2014) in the outer disk.

4. Tracing the spiral arms with classical Cepheids

Since they provide accurate distances, there have been many investigations of the Galactic structure using Cepheids, often with the primary goal to derive the Galactocentric distance of the Sun or to determine the Milky Way rotation curve (e.g., Caldwell & Coulson 1987; Pont et al. 1997; Metzger et al. 1998; Mróz et al. 2019, and references therein). Regarding the spiral arms, Dambis et al. (2015) matched their Cepheid data to a four-armed pattern with a pitch angle of $9.5 \pm 0.1°$. Using Cepheids in the far side of the disk, Minniti et al. (2021) favor instead a two-arms model expanding into four arms for $R_G \gtrsim 5-6$ kpc. However, the spiral structure has mostly been traced by younger tracers (for instance, H II regions, Georgelin & Georgelin 1976), although their distances were in general less accurate because

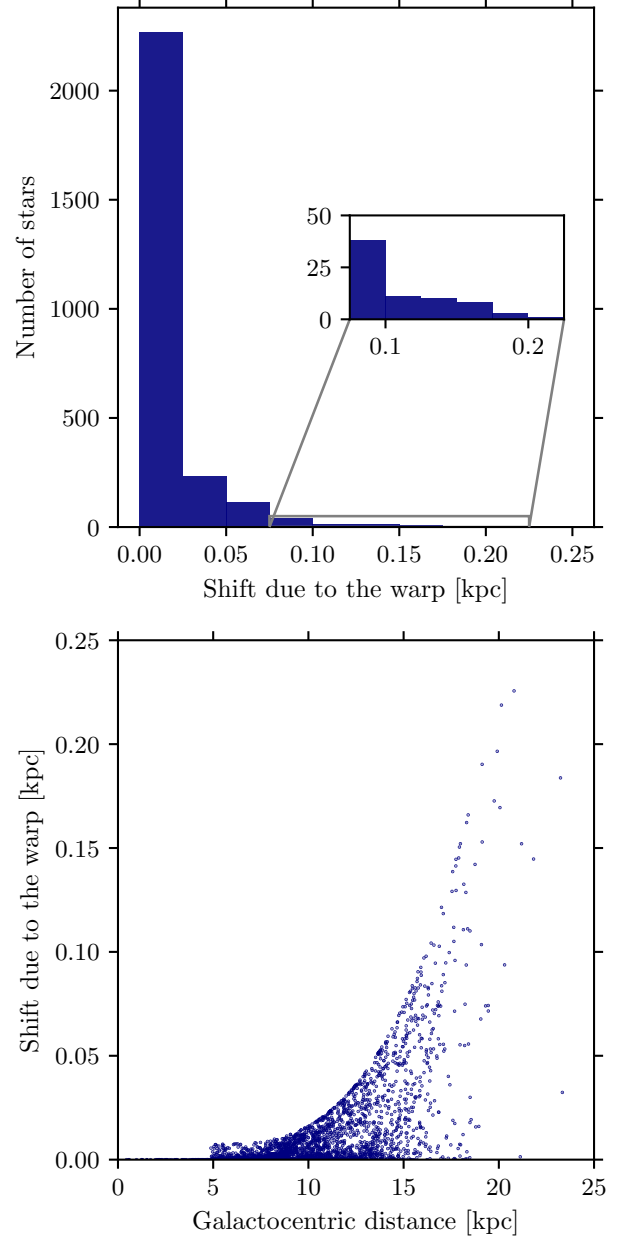**Fig. 4.** Impact of the warp on Galactocentric distances. *Top panel*: distribution of the difference between Galactocentric distances computed on a flat and on a warped disk for our entire sample of Cepheids. *Bottom panel*: difference between Galactocentric distances computed on a flat and on a warped disk, as a function of the Galactocentric distance.

they rely on kinematical models. Indeed, in a traditional textbook picture of a spiral arm, a shock wave concentrates material in the so-called dust lane. Toward the outer disk, one then encounters masers associated to protostars, followed by H II regions, and further on by stars having reached (e.g., OB stars) or evolved off the main sequence (e.g., Roberts 1969; Vallée 2020).

4.1. The age question and the choice of spiral arms tracers

It was quickly suggested (Fernie 1958; Kraft & Schmidt 1963) that the brighter, longer-period Cepheids match the spiral arms as traced by atomic hydrogen better than their fainter, shorter-period counterparts. It was also correctly conjectured that such Cepheids are younger and, therefore, had less time to drift
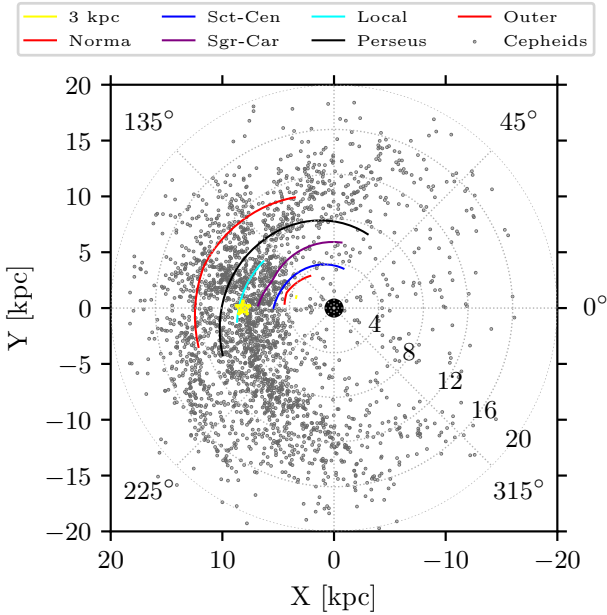
Fig. 5. Classical Cepheids (black dots) in the Galactic plane. The Galactocentric distances are derived from mid-infrared photometry to minimize the effect of reddening. The spiral arms delineated by Reid et al. (2019) are over-plotted using the same color-coding as in the original paper. Two segments are plotted in red as they presumably belong to the same Norma-Outer ring. Concentric circles are shown every 4 kpc to guide the eye. The Galactic center (black filled circle) is at (0,0) and the Sun (yellow star) at (8.275,0).

away from their birthplace. Indeed the age of a Cepheid is inversely correlated with its period via period-age relations (see Efremov 1978, and references therein). Two ways can be envisioned to overcome the issue of tracers of the spiral arms having evolved off them: either selecting truly young tracers (H II regions, O stars), or selecting only the youngest objects for tracers spanning a larger age range.

For instance, Castro-Ginard et al. (2021) restricted their sample of open clusters to those younger than 80 Myr, while Hao et al. (2021) used 100 Myr. Selecting young stellar groups within 3 kpc from the Sun, Kounkel et al. (2020) report that the separation between spiral arms remains visible up to 63 Myr (log(age) = 7.8). Their scenario favors transient arms, and indicates that the Sagittarius arm has moved toward the Galactic center by 0.5 kpc in the last 100 Myr. Using a small number of Cepheids in the Solar neighborhood, also split into a younger and an older group, Veselova & Nikiforov (2020) found 7 and 8 spiral arm segments, respectively. For a given segment, they report that the parameters retrieved for the young and the old objects differ, especially in the case of the Sagittarius and the Perseus arms (see also Bobylev et al. 2021).

In this context it is worth mentioning that Cepheids' ages are not extremely accurate since they can vary by a factor up to 2 depending on whether stellar rotation is included (Anderson 2014) or not (Bono et al. 2005) in the evolutionary models (see also the recent paper by De Somma et al. 2020, with no rotation). However, their ranking by age is very reliable since their period, which can be measured with great accuracy, is the driving parameter via period-age relations. In an attempt to constrain these relations using Cepheids in open clusters, Medina et al. (2021) noted that ages of young open clusters potentially hosting Cepheids (and, therefore, younger than ≈300 Myr) are quite inaccurate given the paucity or even the absence of evolved stars

and the stochastic sampling of their initial mass function (IMF). Such uncertainties affect not only the absolute ages of young clusters but also their ranking by age.

Both Poggio et al. (2021) and Zari et al. (2021) found overdensities of upper main-sequence stars in *Gaia* data, which could be associated with the Sagittarius-Carina and the Scutum-Centaurus arms. Poggio et al. (2021) found no obvious match between their overdensities and Cepheids with age <100 Myr, while in contrast the agreement was good with open clusters of similar ages. They noted, however, that the young Cepheids (<100 Myr) overlap quite well with the spiral structure proposed by Taylor & Cordes (1993) or the Perseus arm characteristics proposed by Levine et al. (2006b). Gaia Collaboration (2022) reached the same conclusions, we note in passing that they accepted Cepheids up to 200 Myr old in their sample. Similarly, Majaess et al. (2009) indicate that Cepheids younger than 80 Myr in their sample are good tracers of the spiral structure. Although the same level of detail cannot be reached at large distances, it is worth mentioning that in M 31, the sample of classical Cepheids of Kodric et al. (2018) closely follows the position of the ring structures rich in dust and star-forming regions. Finally, Minchev et al. (2013, 2014) coupled their chemo-dynamical model to high-resolution simulations tailored to the Milky Way in a cosmological context. They concluded that the oldest stars are the most affected by stellar radial migration, while the young stars are found near their birth radii. Recent studies (e.g., Frankel et al. 2020; Lian et al. 2022, and references therein) confirmed that radial migration is inefficient over short time-scales[8].

Notwithstanding, over-plotting the spiral arms delineated by Reid et al. (2019) on top of the entire sample of Cepheids without any age restriction, as shown in Fig. 5, indicates that the Cepheids' overdensities match the spiral arms well. Inter-arm regions have lower densities of Cepheids, as can be seen in the radial distribution of Cepheids located in a Galactocentric angular sector around 160° displayed in Fig. 6.

In what follows, we use the analytical period-age relations provided by Bono et al. (2005) to derive the Cepheids' ages, and we restrict our sample to stars pulsating in the fundamental or the first-overtone mode for which such relations are available[9].

4.2. Locating groups of Cepheids

To identify the spiral arms, we used `t-SNE` (t-distributed Stochastic Neighbor Embedding, van der Maaten & Hinton 2008), a nonlinear dimensionality reduction technique. Although `t-SNE` is often used to visualize high-dimensional data in a lower-dimensional space, we only used as input the coordinates $(\theta, \ln r)$ of the Cepheids in our dataset, where r has been corrected from the effects of the warp (see Sect. 3.3). Since the algorithm uses a Student's t-distribution to compute the similarity between two data points in the `t-SNE` output space, it performs very well in keeping similar input data points close together in the output space, even if they come from crowded regions. The downside is that `t-SNE` performs poorly when data are sparse. After the data were standardized, `t-SNE` was initialized using a principal component analysis and run for 6000 iterations in a 2D

---

[8] Lian et al. (2022) report, for instance, average migration distances of 0.5–1.6 kpc after 2 Gyr and 1.0–1.8 kpc after 3 Gyr (roughly 20 to 150 times more than the ages of Cepheids considered here).
[9] Second-overtone and multimode Cepheids are quite uncommon in the Milky Way, see Bono et al. (2002), Smolec & Moskalik (2010), Lemasle et al. (2018).
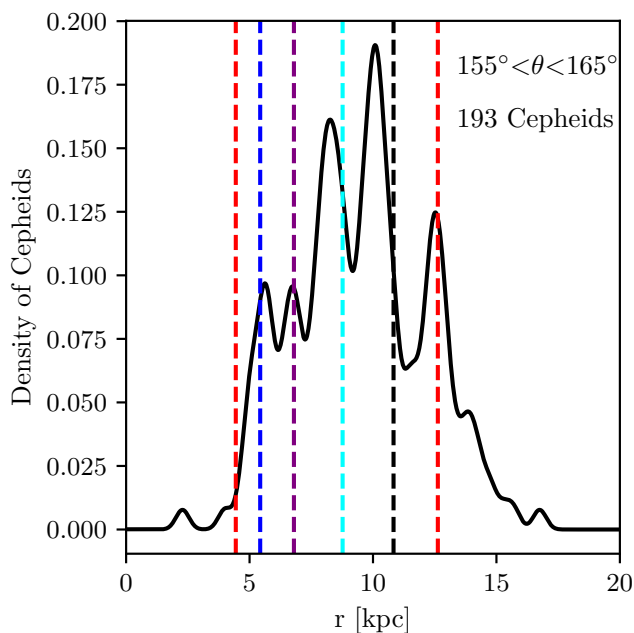
Fig. 6. Kernel density estimation (with a kernel bandwidth of 0.1) of the radial distribution of Cepheids located in a Galactocentric angular sector around 160°. The location of the spiral arms of Reid et al. (2019) in this sector are shown using the same color-coding as in the original paper. The angular sector around 160° intercepts the spiral arms in a region where the completeness of the data is not hindered by the two shadow cones visible in Fig. 5 which hamper the detection of Cepheids beyond nearby regions with strong extinction.

space. The perplexity (the effective number of neighbors considered by `t-SNE` for any given data point) was set to 90. For our dataset, the topology of the outcome in the `t-SNE` space is robust to the choice of the perplexity value, as well as to the value of the early exaggeration (set to 5), which ensures that tight clusters in the data will not overlap in the `t-SNE` space. Individual groups are then identified using the clustering algorithm `HDBSCAN` (see details below).

To ascertain that our `t-SNE+HDBSCAN` algorithm is working, we have run several tests where the mock spiral structure is based on the Reid et al. (2019) model. They are described in detail in Appendix B. Tests show that the algorithm recovers the mock spiral arms very well, even in the presence of large amounts of "inter-arm" Cepheids, that can be considered as noise. A fraction of these "inter-arm" Cepheids is then included in the nearest arm, but this impacts the recovered location of the spiral arm only marginally. We note that the algorithm is sensitive to small gaps (regions without stars) in individual spiral arms. A given spiral arm may then be split in several segments limited by those gaps. This is more likely to occur when two spiral arms are very close to each other. In such a case, it might even happen that two segments from two different arms are wrongly joined within the same group (and the recovered arm location wrongly falls at a median distance between the two segments).

Coming back to real data, the top-left panel in Fig. 7 shows how Cepheids younger than 150 Myr are distributed in the `t-SNE` space. In this plot, the color-coding indicates groups identified by `HDBSCAN`, a clustering algorithm using unsupervised learning to identify clusters in a distribution of data points (Campello et al. 2015; McInnes et al. 2017). `HDBSCAN` was run with hyperparameters imposing a minimum of 5 groups, well below the number of clusters actually found, a minimum of

20 members per group in order to avoid spurious detections of tiny groups, and assuming Euclidean distances between individual points in the embedded space.

Using the same color-coding, the bottom panel of Fig. 7 shows the Cepheids in the $(\theta, \ln r)$ space. The groups identified by `t-SNE+HDBSCAN` form narrow, linear sequences in this plane, as is expected under the common assumption that spiral arms follow a logarithmic spiral. The top-right panel of Fig. 7, showing the spatial distribution of the identified groups in the Milky Way plane suggests that each group forms indeed a section of a given spiral arm. A comparison with the spiral structure obtained via other tracers (see Sect. 4.5) confirms that our method allows us to trace the Milky Way spiral arms using Cepheids.

As already mentioned, `t-SNE` does not perform well in the case of sparse data, and the large groups (1, 3) gathering distant Cepheids reflect this weakness. They do not trace reliable spatial structures, but emerge only because the search of pulsating variable stars is still largely incomplete (and their classification uncertain) at large distances in the disk, especially toward regions of high extinction like the far side of the disk. Since they do not correspond to real features, these groups will not be discussed further in the paper. Similarly, a few isolated Cepheids in the outer disk are attributed to likely unreliable groups, for instance, to groups 2 or 15.

From tests (see Fig. C.1) where the sample of Cepheids considered is restricted to those younger than 100, 120, 150, and 250 Myr, respectively, we draw several conclusions:

1. Whatever the age cut, the groups identified have the same morphology in the `t-SNE` space, translating into similar spiral arms in the Galactic plane. Increasing the age limit, hence the number of stars, enables us to split the larger groups into subgroups.

2. Increasing the age limit also enables us to better identify spiral features toward the outer disk. This is not a surprise since, for instance, Skowron et al. (2019a) already mentioned that younger Cepheids are observed preferentially in the inner disk. Such a trend is counter-intuitive in the context of an inside-out formation of the disk (Matteucci & Francois 1989), it is in fact a combined manifestation of the Milky Way's radial metallicity gradient (e.g. Lemasle et al. 2008), where stellar metallicities decrease from the inner to the outer disk, together with the metallicity-dependence of the Cepheids' instability strip (IS, e.g., Fiorentino et al. 2013; De Somma et al. 2020), where the age at which a star of a given mass reaches the IS (or possibly does not even cross it) depends on its metallicity.

3. If the age limit is set too high, features start to blur again.

In the rest of the paper, we work with a sample restricted to Cepheids younger than 150 Myr (figures regarding the samples with different age limits are provided in Appendix C). The arbitrary selection of 150 Myr as an age limit is a compromise in order to identify a good number of spiral features, bearing in mind the earlier discussion on age in Sect. 4.1. The groups identified here lead us to characterize spiral arms by segments. In the next subsections, we put such a definition in context, we provide the characteristics of each individual segment, we compare our spiral pattern to some of the most commonly used spiral models, and we investigate the age distribution of Cepheids within individual segments.

4.3. Defining spiral arms by segments

It is not entirely clear whether the definition of spiral arms by segment is a consequence of the algorithm employed (`t-SNE` focuses on local similarities in the data), of inhomogeneous
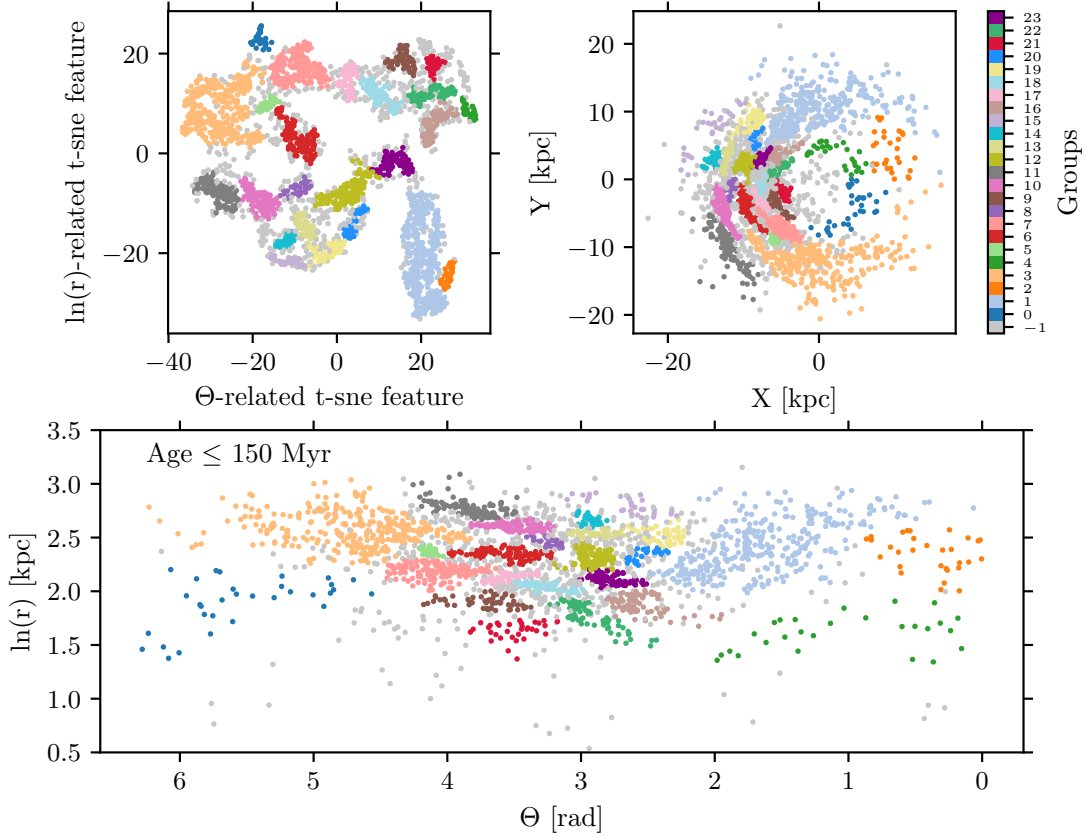
Fig. 7. Groups of Cepheids identified by `HDBSCAN` in the `t-SNE` space (*top left panel*). The same groups are presented in the θ, $\ln r$ space (*bottom panel*), where θ is the Galactocentric azimuth and $\ln r$ the logarithm of the Galactocentric radius (corrected from the warp), and in the Galactic plane (*top right panel*). All the Cepheids shown in this figure are younger than 150 Myr.

completeness of the Cepheid data, or simply a natural outcome of the mechanisms driving the formation of spiral arms in the Milky Way. Poggio et al. (2021) mention that a good fraction of the clumpiness they see in their data (possibly even some larger-scale structures) are caused by foreground extinction. Zari et al. (2021) note, however, that some low-density features are not located in regions of strong extinction and are detected in the spatial distributions of many young tracers, a possible indication that those are not artifacts in the data. It is plausible that young stellar tracers have a clumpy distribution, either because they trace the clumpiness of giant molecular clouds (assuming a stationary spiral pattern), or because star formation is associated with the kinematics and the density distribution across spiral arms (if one considers instead that they are the aftermath of a transient phenomenon).

Studies of external galaxies have also shown that spiral arms are not necessarily homogeneous structures, but can present under-/over-densities, or even be defined by contiguous segments in the most pronounced cases (e.g., Chernin 1999; Kendall et al. 2011; Honig & Reid 2015). Similarly, Reid et al. (2019) introduced kinks in their logarithmic spiral arms to better fit the spatial distribution of their tracers. Spiral segments are also a natural outcome of theoretical models (e.g., Grand et al. 2012; D'Onghia et al. 2013; Mel'nik & Rautiainen 2013) and it has been proposed that they are the response of the stellar disk to the growth of overdensities corotating with the disk (see e.g., Sellwood & Masters 2022, for a detailed discussion on these topics).

Udalski et al. (2018) mention that OGLE can detect a $P = 3$ d Cepheid at ~20 kpc from the Sun, even with an extinction

reaching $A_I \approx 4$ mag, and estimate the completeness of the OGLE survey to be on the order of 90% for classical Cepheids. It seems then reasonable to discard a significant incompleteness of the Cepheids' catalogs. Another aspect to consider is the number of Cepheids. Their progenitors, late B-type stars, are not extremely numerous given the structure of the IMF, and the brevity of the Cepheid phase, a few tens to a few hundreds megayears (depending on their mass and metallicity) makes them rare objects and, as such, likely not the best tool to discriminate between a multiarm and a flocculent Milky Way. The algorithm developed by Veselova & Nikiforov (2020) relies only on the tracers considered to determine the properties of spiral arms/segments, without any assumption on the total number of segments or the membership of a given tracer in a specific segment. Using a sample of nearby Cepheids, they found seven spiral segments using the youngest part of the sample and eight using the oldest Cepheids.

4.4. Parameters of spiral segments traced by Cepheids

In order to determine the properties of a given individual segment, we fit a linear relation in the $(\theta, \ln r)$ space through all group members identified by `t-SNE+HDBSCAN`:

$$\ln r = a \times \theta + b, \tag{5}$$

where a is the slope and b the intercept. From the slope, we derive the pitch angle. The minimal and maximal Galactocentric azimuths covered by a given group are also recorded. The midpoint of these two values is used as the reference angle, and the corresponding radius, calculated using the fitted linear relation,

Table 4. Characteristics of individual segments of spiral arms as identified by `t-SNE+HDBSCAN` for Cepheids younger than 150 Myr.

Group	Slope (kpc rad^{-1})	σ_{slope} (kpc rad^{-1})	Intercept (kpc)	$\sigma_{intercept}$ (kpc)	Ref. angle (rad)	Ref. radius (kpc)	ln(Ref. radius) (kpc)	Pitch angle (rad)	Min. angle (rad)	Max. angle (rad)
[4]	−0.090	0.050	1.705	0.060	1.068	4.997	1.609	0.090	0.154	1.982
[21]	−0.082	0.087	1.922	0.304	3.502	5.129	1.635	0.082	3.177	3.827
[22]	0.575	0.042	0.087	0.124	2.835	5.568	1.717	−0.522	2.436	3.233
[16]	0.187	0.047	1.399	0.121	2.379	6.321	1.844	−0.185	1.950	2.807
[9]	0.052	0.032	1.707	0.120	3.776	6.708	1.903	−0.052	3.375	4.177
[0]	−0.255	0.073	3.304	0.407	5.424	6.827	1.921	0.250	4.566	6.282
[18]	0.163	0.019	1.512	0.061	3.253	7.708	2.042	−0.162	2.980	3.526
[23]	0.145	0.038	1.711	0.107	2.759	8.257	2.111	−0.144	2.511	3.007
[17]	−0.212	0.026	2.882	0.093	3.527	8.451	2.134	0.209	3.317	3.736
[7]	0.039	0.025	2.031	0.101	4.044	8.924	2.189	−0.039	3.639	4.450
[12]	−0.023	0.062	2.381	0.181	2.945	10.107	2.313	0.023	2.691	3.198
[6]	0.048	0.020	2.176	0.072	3.602	10.474	2.349	−0.048	3.220	3.984
[5]	0.357	0.124	0.900	0.511	4.120	10.705	2.371	−0.343	4.031	4.208
[8]	0.289	0.072	1.520	0.235	3.274	11.777	2.466	−0.281	3.141	3.406
[13]	−0.006	0.029	2.554	0.082	2.889	12.637	2.537	0.006	2.684	3.094
[10]	0.041	0.020	2.460	0.069	3.516	13.520	2.604	−0.041	3.207	3.826
[14]	0.290	0.121	1.814	0.355	2.919	14.302	2.660	−0.282	2.806	3.031
[11]	0.212	0.033	1.962	0.125	3.868	16.150	2.782	−0.209	3.480	4.256

Notes. The slope and intercept have been computed in the $(\theta, \ln r)$ space. The reference angle and radius have no physical meaning, they are simply selected as the midpoint of the Galactocentric azimuth range covered by a given group.

is the reference (logarithm of the) radius. These values, listed in Table 4, are then used to trace the spiral segments displayed in Fig. 8. As can be seen on the figure or in the tabulated data, several segments are located at a similar reference radius and can be interpreted as different sections of the same spiral arm.

4.5. Comparison with spiral arms models

Figure 8 displays the spiral segment we derived for Cepheids younger than 150 Myr over-plotted on our data or on various spiral models, namely those of Levine et al. (2006b), Reid et al. (2019), and Hou (2021). Figures similar to Fig. 8 for the other age ranges are provided in Appendix D.

Several large groups (1, pale blue; 2, orange; 3, pale orange) are not resolved by our algorithm. They are located at large distances from the Sun in the first and in the fourth quadrant. In addition to sparser data in these regions, it is possible that even only slightly larger uncertainties on the distances, and/or a larger number of contaminants, blur the spiral arms signal. Having no physical meaning, these groups are not considered further.

Two other groups (0, dark blue; 4, green) trace long (several kiloparsecs) spiral segments and are defined by a relatively small number of Cepheids. It could be that the stars were connected simply from the lack of further data, but is possible that some of them actually trace real features. They are located in the far side of the (inner) disk. Minniti et al. (2021) have reported that these Cepheids are compatible with a two-arm model (Perseus and Sct-Cen) branching out into four arms for $R_G \gtrsim 5-6$ kpc.

Group 21 (red) seems to prolong the Sct-Cen arm from the model by Reid et al. (2019). However, it better follows the Norma arm as charted by Hou (2021). Group 9 (brown) is also an excellent match to the Sct-Cen arm by Hou (2021), while it would appear to be more of a continuation of the Sgr-Car as defined by Reid et al. (2019). Group 22 (green) seems to bridge the Sct-Cen and the Sgr-Car spiral arms, according to both models by Reid et al. (2019) and Hou (2021). It may be that here the algorithm is unable to separate two closely parallel segments. If

so, this group should actually be split in two parts, which would then follow the Sct-Cen and the Sgr-Car spiral arm, respectively.

Group 16 (light brown) is an almost exact match to the Sgr-Car arm from Reid et al. (2019), and remains in reasonable agreement with the definition of this arm by Hou (2021). Group 18 (light cyan) constitutes a plausible continuation of the Sgr-Car arm from Reid et al. (2019), and indeed it partially overlaps this arm in the model by Hou (2021). Closer to the Sun, this segment, however, approaches the Local arm.

Group 17 (light pink) continues the Local arm as traced by Reid et al. (2019), but it would reach the Sgr-Car arm if the latter were prolonged with the same pitch angle. And indeed this feature overlaps the Sgr-Car arm in the model by Hou (2021). Group 23 (violet) overlaps very well with the Local arm when compared to both models. Group 7 (coral pink) is located at the same Galactocentric distance as the Local arm (~8 kpc), but it could equally be an extension of the Sgr-Car or of the Local arm, especially if we adopt the location proposed by Hou (2021) for the latter.

Both group 12 (khaki) and group 6 (red) are good matches to the Perseus arm as defined by Reid et al. (2019) and especially Hou (2021). Group 5 (pale green) further extends the Perseus arm toward the fourth quadrant.

Similar conclusions can be reached when comparing our segments to the spatial overdensities reported by Poggio et al. (2021) in their sample of upper main sequence stars, thus comparable objects as those used by Hou (2021) to trace spiral arms. The outcome of this exercise is displayed in Fig. 9. It suggests that groups 18 (light cyan), 17 (light pink) and 7 (coral pink) are all related to the Sgr-Car arm, while group 23 (violet) is the Local arm. It remains unclear whether group 6 (red) also belongs to the Local arm, as suggested by Poggio et al. (2021), or whether it is a continuation of the Perseus arm, as suggested by Fig. 8 from its pitch angle as well as from the good match with the models by Reid et al. (2019) and Hou (2021). In the latter case, the Local arm might end close to the present location of the Sun.
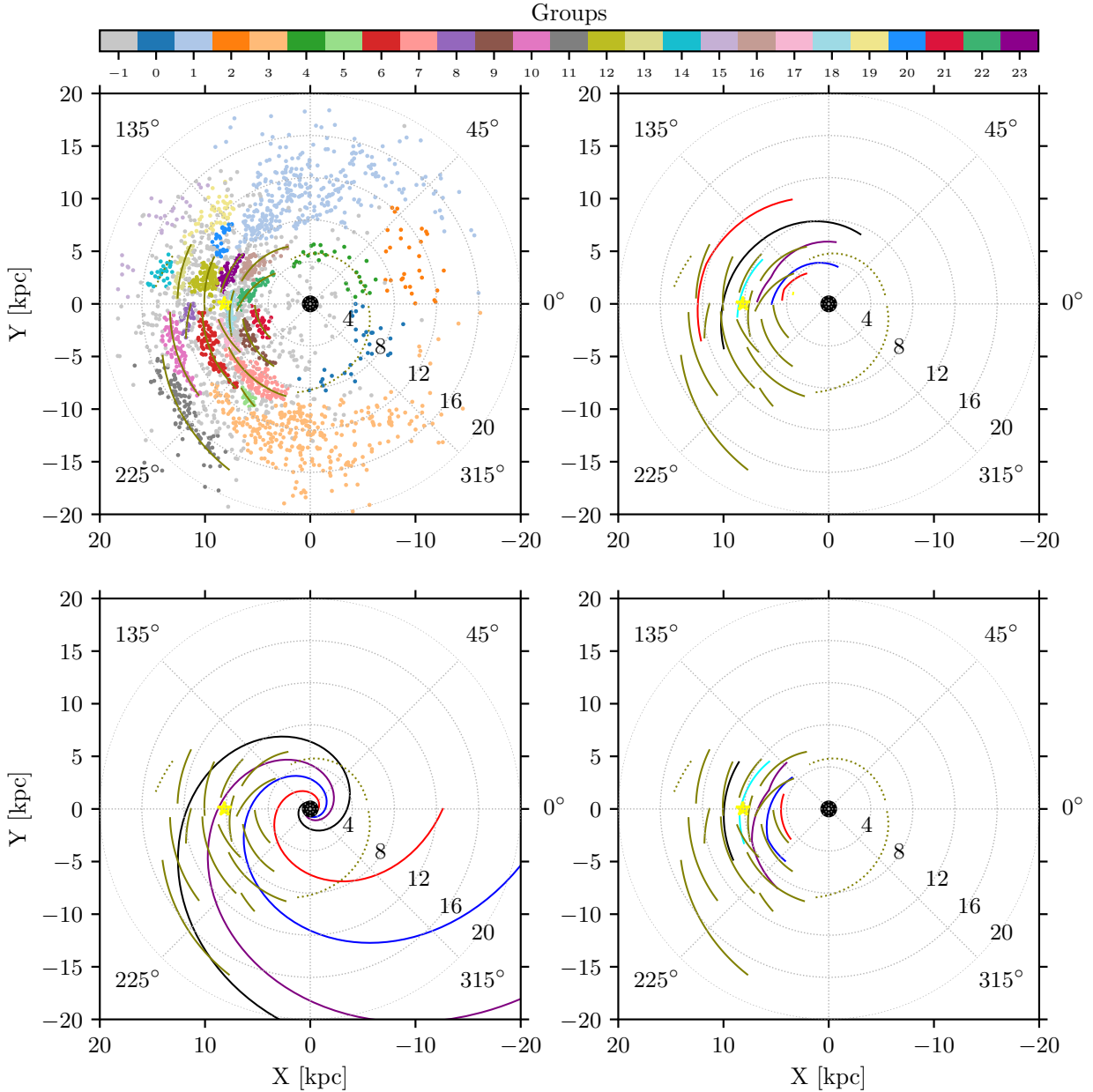
Fig. 8. Comparison with previous models. *Top left*: spiral segments (olive) over-plotted on Cepheids identified as members of a group. Groups are color-coded as in Fig. 7. *Top right*: spiral segments and the model of Reid et al. (2019). The spiral arms delineated by Reid et al. (2019) are over-plotted using the same color-coding as in the original paper. *Bottom left*: spiral segments and the model of Levine et al. (2006b), with the same color-coding as in Reid et al. (2019). *Bottom right*: spiral segments and the model of Hou (2021), with the same color-coding as in Reid et al. (2019).

Our results underline the difficulty to unravel the Milky Way spiral structure in the Solar vicinity[10]. They confirm that the Local arm is not a short segment or a spur emanating from another arm (van de Hulst et al. 1954) but an independent, elongated structure (Reid et al. 2019; Xu et al. 2021), extending at least from ≈135° to ≈180°, and possibly approaching the Sgr-Car arm at Galactocentric azimuths ≈200°.

Group 13 (pale khaki) is an excellent match to the Outer arm as sketched by Reid et al. (2019), and group 10 (pink) seems to constitute its natural continuation, although it is located

at a somewhat larger Galactocentric distance than modeled by Reid et al. (2019). Group 14 (cyan) is identified as a potential spur extending out of the Outer arm in the anticenter direction. Group 8 appears as an isolated segment in between the Perseus and the Outer arms.

Groups 19 (pale yellow) and 20 (blue) are not a reliable feature. In the range 135°–150°, at least a fraction of the Cepheids attributed to these groups should arguably belong to group 12 (khaki) or group 13 (pale khaki). In this region, the algorithm is strongly affected by a shadow cone[11] already reported by Poggio et al. (2021, their Fig. 1), which is also quite obvious in our data (another shadow cone splits group 1 and probably

---

[10] We note in passing that 16 Cepheids within 1 kpc from the Sun have been discarded, most of them because they did not fulfill the fractional parallax uncertainty criterion. Including them would increase the amount of nearby Cepheids by 30%.

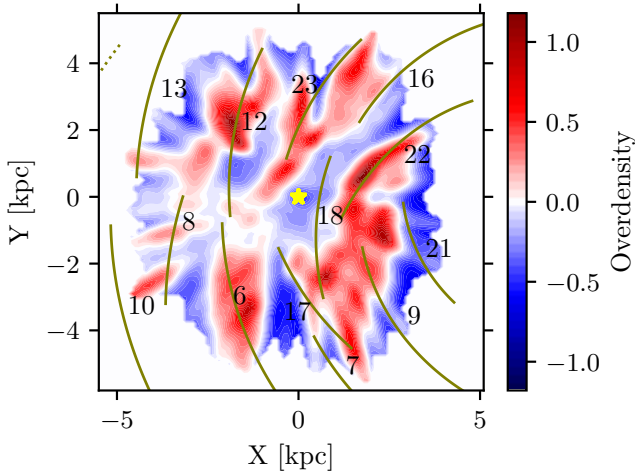[11] Where the detection of targets is hampered beyond nearby regions with strong extinction.

Fig. 9. Overdensities in the spatial distribution of upper main sequence stars (Poggio et al. 2021), based on a local density scale length of 0.3 kpc. The Sun is represented by the yellow star at (0,0). Our spiral segments (olive) are over-plotted.

prevents the algorithm to identify groups in these distant regions of the first quadrant). The reality of group 15 (pale violet) cannot be assessed due to the paucity of Cepheids in the far outer disk, group 11 (gray) can be attributed to the Outer arm only if we accept a drastic change in its pitch angle. This feature agrees quite well with the model by Levine et al. (2006b), where it represents, however, the Perseus arm, as already noted by Poggio et al. (2021). The data are too sparse to draw firm conclusions, and especially to trust that this feature indeed extends up to ~250°.

We also compared our findings to the study by Veselova & Nikiforov (2020). It relies on a much smaller sample of Cepheids (636, from Mel'nik et al. 2015) with overall less accurate distances (but well-constrained radial velocities). They consider segments as a section of a logarithmic spiral. The membership of Cepheids in segments and the properties of each segment are determined simultaneously. Matching spiral segments in Veselova & Nikiforov (2020) to our groups was carried out via a visual inspection of Fig. 8 in this study and Fig. 5 in theirs. When possible, we included in the comparison some of our segments lying outside the (small) range of Galactocentric azimuths encompassed by the study of Veselova & Nikiforov (2020). We note that group 19 (red), which we associate with their Sagittarius-2 segment, only seems to have a limited length in our study (but one could argue that the small spur visible at the extremity of group 4 (green) might be a continuation of this feature). In Table 5, we compare the Galactocentric distance of the spiral arms in their study and in ours. The results are in very good agreement until the (first) Outer arm. We note that the two outermost features are only defined by a select number of stars in Veselova & Nikiforov (2020), while our comparison groups contain much more stars, from which, however, only a few overlap with the spatial extent of the sample by Veselova & Nikiforov (2020). The consistency between the two studies suggests that our detection of spiral arms is robust, even outside the comparison region. We also find a fairly good agreement with the groups of Cepheids identified by Genovali et al. (2014), without considerations on age. The identification was based on a clustering algorithm (Path Linkage Criterion, Battinelli 1991) already applied to Galactic Cepheids by Ivanov (2008), and on a stellar density

threshold between the candidate groups and their immediate neighborhood.

4.6. Age distribution of Cepheids across spiral segments

To investigate the age distribution of Cepheids in our individual spiral segments, we use the data tabulated in Table 4 to rotate each data point by the pitch angle value around the reference point in the $(\theta, \ln r)$ space. It then becomes easy to derive the (logarithmic) distance of a given Cepheid from the (logarithmic) reference radius, and to convert it into real spatial distances. We then plot the ages of the Cepheids attributed to a given group against their distance to the reference radius. Age gradients across individual segments are shown in Fig. C.2, together with linear fits to the data.

The period-age relations by Bono et al. (2005) and Anderson et al. (2014) do not allow us to derive uncertainties on the age of individual Cepheids. The standard deviation of the period-age relation by Bono et al. (2005), which mainly reflects the finite width of the instability strip, can be used as a (loose) proxy for the 1-σ uncertainty on age. We find that it can reach 15 Myr for nonrotating Cepheids. Including rotation in models increases the ages of Cepheids by 50 to 100%, depending on the amount of rotation and the period of the Cepheid (Anderson et al. 2014). Finally, the helium and metal contents, a possible core convective overshooting during the core H-burning stage, and the mass loss efficiency (mostly) during the red giant branch phase, all affect the model predictions of the Cepheids' individual ages (see e.g., De Somma et al. 2020). The theoretical period-age relations are provided only for selected values of these parameters. The true values of these physical quantities likely vary from star to star, and they are anyway not available to us at this time, a fortiori for large samples. It is important for the current analysis to mention these caveats, but they should not hide the fact that Cepheids are among the stars with the best age determinations.

Beyond the uncertainties on the Cepheids' ages already mentioned above, and the underlying assumption that our spiral segments can be defined by sections of a logarithmic spiral arm, we note that the birth location of the Cepheids is still unknown. Medina et al. (2021) report only a relatively small number of Cepheids confidently associated with open clusters, which may indicate that Cepheids are born elsewhere, or alternatively that their birth cluster or association dissolved quickly. The dynamical evolution of Cepheids in star clusters has been investigated theoretically by Dinnbier et al. (2022). Moreover, the recent discovery of numerous spurs and feathers (e.g., Kuhn et al. 2021; Veena et al. 2021, and references therein), sometimes extending quite far away from the estimated locus of the corresponding spiral arm, hints at a more complicated picture.

Still, our analysis already provides a gross estimate of the age gradient across spiral arms. Numerous studies (e.g., Shabani et al. 2018; Castro-Ginard et al. 2021, to quote only a few recent ones) have searched for age gradients in our or in external galaxies in order to test the spiral density wave theory (see e.g., Dobbs & Pringle 2010), but the majority of them simply report their absence or their detection (but see Vallée 2020). The values we obtain range from 0 to $\approx$15 Myr kpc^{-1}. They agree quite well with the age gradients of 12 ± 2 Myr kpc^{-1} reported by Vallée (2021). We note in passing that given the age limit we set at 150 Myr, a few age distributions for individual segments are potentially truncated at higher ages and may provide unreliable values.

Table 5. Comparison of the Galactocentric distances of spiral arms in this study with the Galactocentric distances of spiral arms in Veselova & Nikiforov (2020, V20) and with the Galactocentric distances of Cepheids groups identified by Genovali et al. (2014, G14).

Arm	G14 (kpc)	Group(s)	V20 "Young" (kpc)	V20 "Old" (kpc)	This study (kpc)	Group(s)
Scutum			5.94	6.07	5.35	21, 22
Sagittarius-1	6.41	I, II	6.74	6.79	6.62	0, 9, 16
Sagittarius-2	7.45	III, IV	7.41	7.52	7.71	18
Local	8.10	V	8.31	8.19	8.35	17, 23
	8.99	VI			8.92	7
Perseus	9.54	VII, VIII	10.13	9.91	10.11	12
	10.13	IX, X			10.59	5, 6
Outer-1			12.59	12.49	13.08	10, 13
Outer-1a				14.68	14.30	14
Outer-2			16.61	16.79	16.15	11

Notes. In Veselova & Nikiforov (2020), the Cepheids are split into an old ($P \leq 9$ d) and a young ($P \geq 9$ d) sample. Heliocentric distances of the spiral arms of Veselova & Nikiforov (2020) are taken from their Table 1, and converted into Galactocentric distances by adding 7.33 kpc and multiplying the outcome by a correction factor of 1.117, as indicated by those authors. The Galactocentric distances of the spiral arms in this study are the average of the reference radius of individual groups (see Table 4). Groups representative of a given spiral arm have been selected by comparing Fig. 8 in this study and Fig. 5 in Veselova & Nikiforov (2020). For the Cepheids groups identified by Genovali et al. (2014), we use directly the tabulated Galactocentric distances or their average value if we combine two of their groups. We note that Genovali et al. (2014) used 7.94 kpc (Groenewegen et al. 2008) as the Galactocentric distance of the Sun.
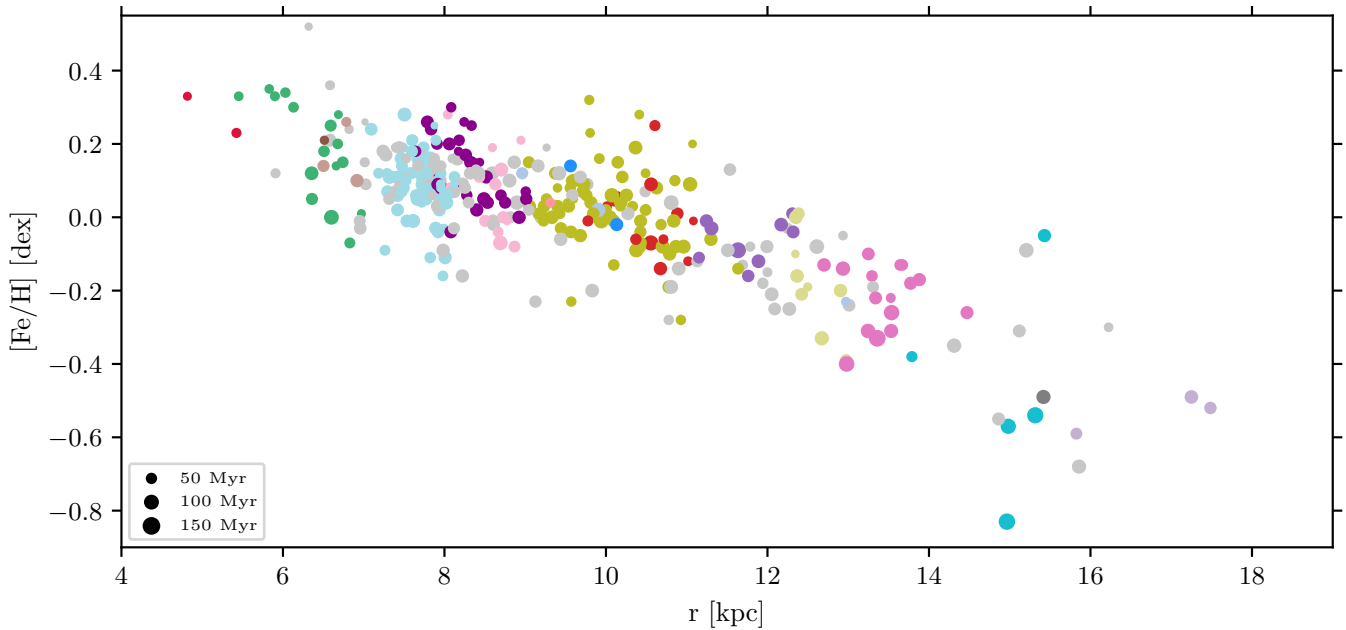

Fig. 10. Radial metallicity gradient, where [Fe/H] derived from high-resolution spectroscopic observations (Genovali et al. 2014) is plotted against Galactocentric distances r determined using a period-luminosity relation (this study) in the mid-infrared. The different spiral segments are color-coded as in Fig. 7, and the size of the data-points correlates with the age of the Cepheids (larger points for higher ages). As in previous figures, only Cepheids with an age ≤ 150 Myr are considered.

5. Spiral arms and abundance gradients

In this section, we investigate how the spiral arms may impact radial and azimuthal[12] metallicity gradients, using literature values from Genovali et al. (2014) and references therein. We emphasize that the conclusions drawn in this section are only tentative since we list below a number of important caveats that should be kept in mind.

5.1. Caveats

The first caveat we want to mention is that the spectroscopic analysis of Cepheids is not immune from (phase-dependent) NLTE effects (see the series of papers by Vasilyev et al. 2017, 2018, 2019, for instance). NLTE effects are more important for more massive Cepheids, which are also the longer period, younger ones that are presumed to be the best tracers of spiral arms.

The catalog of Cepheid metallicities by Genovali et al. (2014) contains several tens of Cepheids analyzed in their paper, using the same method as for the stars in Lemasle et al. (2007, 2008), Romaniello et al. (2008), Pedicelli et al. (2010),
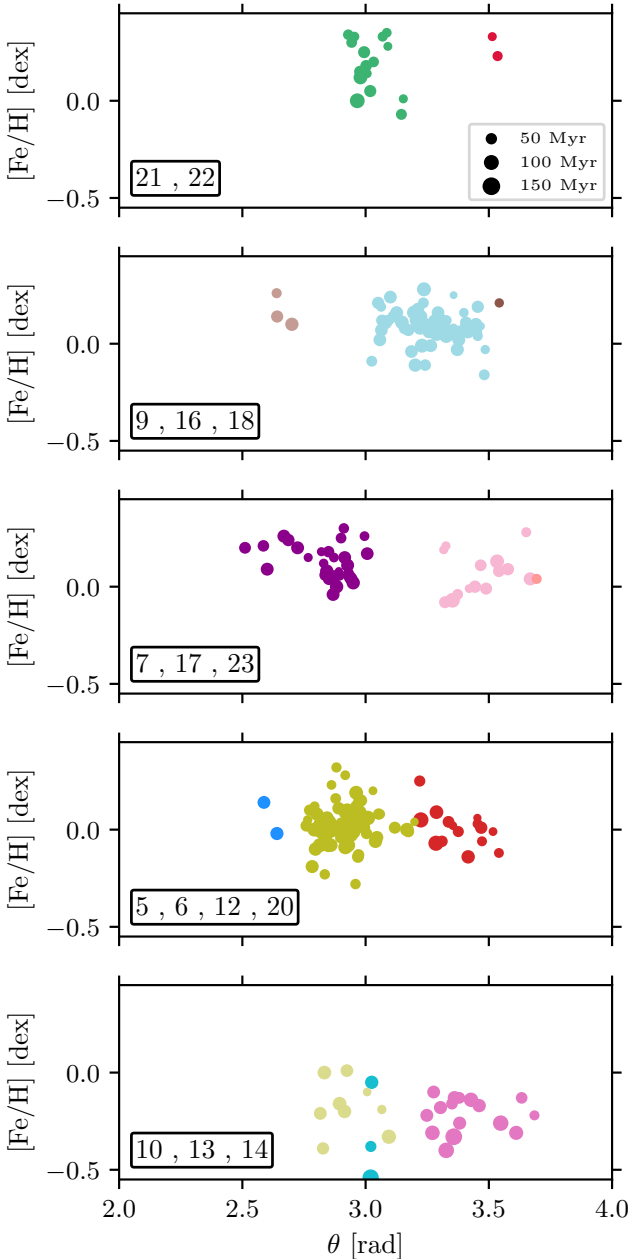
---

[12] By azimuthal metallicity gradient, we mean the variation of the metallicity with the Galactocentric azimuth, within a given radial annulus at fixed Galactocentric distance.

Fig. 11. Longitudinal metallicity gradient, where [Fe/H] derived from high-resolution spectroscopic observations (Genovali et al. 2014) is plotted against the Galactocentric azimuth θ. The different spiral segments are color-coded as in Fig. 7, and the size of the data-points correlates with the age of the Cepheids (larger points for larger ages). As in previous figures, only Cepheids with an age $\leq$ 150 Myr are considered.

Genovali et al. (2013) that they also included in their catalog. They complemented those studies with literature data from other groups, namely Sziládi et al. (2007), Yong et al. (2006), Luck et al. (2011), Luck & Lambert (2011), which were rescaled to take systematic differences between these studies into account. Although these differences are $\lesssim$0.1 dex (with the only exception of Yong et al. 2006 who report notably lower metallicities for the stars in common), we are still dealing with inhomogeneous data.

The strongest caveat is related to line depth ratios (LDR, see Kovtyukh & Gorlova 2000; Kovtyukh 2007; Proxauf et al. 2018), the traditional method to determine the effective temperature of Cepheids, prior to a canonical spectroscopic analysis. Even though the temperature scale derived from line depth ratios is quite precise (better than 100 K), several concerns have arisen regarding the accuracy of the scale, and possible systematic errors depending on the metallicity of the star and the phase of the observation have been suggested (e.g., Mancino et al. 2021). Such potential systematics have motivated the series of papers initiated in Lemasle et al. (2020), where we aim at developing an unbiased method for deriving the chemical composition of Cepheids. Finally, we note that the spectroscopic data available in the literature contain mostly stars within a few kiloparsecs from the Sun, and, therefore, cover only a fraction of the spiral segments identified in this study, some of them very partially (in terms of their longitudinal extension).

5.2. The radial gradient of iron

Assuming a logarithmic spiral structure, one could expect that stars located in different spiral arms may overlap in a 2D representation of the radial metallicity gradient (where the Galactocentric azimuth is not considered), as shown in Fig. 10, For the data we have at hand, this is not the case, but this most likely simply reflects the limited longitudinal extension of the spectroscopic data. A gap at $r \approx 11.5$ kpc indicates a lack of young Cepheids at this radius, but it is already vanishing when considering slightly older Cepheids (see Fig. C.1).

An interesting feature of Fig. 10 concerns the observed scatter in [Fe/H] at a given radius. It is not clear yet whether this scatter is real or only a consequence of uncertainties in the chemical analysis, but the [Fe/H] scatter for Cepheids attributed to a given segment or for unclassified (potentially inter-arms') Cepheids is on the same order of magnitude. In a companion paper from our group (da Silva et al. 2022), we show that the dispersion around the mean gradient can be reduced by tackling some of the systematics (line list, atomic parameters) affecting abundance estimates.

Given the caveats exposed above, we did not try to investigate whether a simple linear gradient or more complicated features, including, for instance, breaks in the slope, would best represent the data. Such an investigation is possible using modern regression methods, but it would require robust estimates of the accuracy and precision of the abundance determination, which we have set as a goal for the series mentioned earlier, and a larger sample of Cepheids, which will soon become available from WEAVE (Dalton 2016) and 4MOST (Chiappini et al. 2019; Bensby et al. 2019). For the same reasons, we did not yet consider other elements except iron.

5.3. The azimuthal gradient of iron

Figure 11 shows the metallicity distribution within spiral segments corresponding to several spiral arms. Keeping the caveats mentioned in Sect. 5.1 in mind, the metallicity excursion barely exceeds 0.4 dex within a given spiral arm, and there are some hints of a metallicity trend with the Galactocentric azimuth (but the range of Galactocentric azimuths covered is relatively small), possibly increasing toward the outer disk. Such results are in line with the conclusions drawn by Kovtyukh et al. (2022). Excellent precision and accuracy are required since model predictions indicate that the expected effects are modest. For instance, Spitoni et al. (2019) indicate that azimuthal variations in the oxygen abundance gradient are on the order of 0.1 dex. Mollá et al. (2019) reach similar conclusions. Within a given segment or spiral arm, we also see no metallicity trend with age.

6. Conclusions

In this paper we took advantage of updated mid-infrared photometry and of the most complete (to date) catalog of Galactic Cepheids to determine the shape of the Milky Way warp using a Bayesian robust regression method. We have derived the Galactocentric distances of individual Cepheids in a nonwarped Milky Way and we concluded that the warp cannot be responsible for the increased dispersion of abundance gradients in the outer disk.

Thanks to a clustering algorithm, we have identified groups of Cepheids in the (θ, ln r) space, where θ and r are the Galactocentric azimuth and distance (corrected from the effects of the warp) of individual Cepheids. Assuming different values for the oldest Cepheid considered, we have fit these groups with segments in the (θ, ln r) space, which translate into portions of spiral arms in the (θ, r) space. These groups are consistent with previous studies mapping the density of young tracers in the Solar neighborhood (e.g., Poggio et al. 2021; Zari et al. 2021), or using them to derive explicitly the location of the spiral arms (e.g., Reid et al. 2019; Hou 2021).

References

Amôres, E. B., Robin, A. C., & Reylé, C. 2017, A&A, 602, A67
Anderson, R. I. 2014, A&A, 566, L10
Anderson, R. I., Ekström, S., Georgy, C., et al. 2014, A&A, 564, A100
Battinelli, P. 1991, A&A, 244, 69
Benjamin, R. A., Churchwell, E., Babler, B. L., et al. 2003, PASP, 115, 953
Bennett, M., & Bovy, J. 2019, MNRAS, 482, 1417
Bensby, T., Bergemann, M., Rybizki, J., et al. 2019, The Messenger, 175, 35
Betancourt, M. 2017, arXiv e-prints [arXiv:1701.02434]
Blaauw, A., Gum, C. S., Pawsey, J. L., & Westerhout, G. 1960, MNRAS, 121, 123
Bland-Hawthorn, J., & Gerhard, O. 2016, ARA&A, 54, 529
Bobylev, V. V., Bajkova, A. T., Rastorguev, A. S., & Zabolotskikh, M. V. 2021, MNRAS, 502, 4377
Bono, G., Groenewegen, M. A. T., Marconi, M., & Caputo, F. 2002, ApJ, 574, L33
Bono, G., Marconi, M., Cassisi, S., et al. 2005, ApJ, 621, 966
Caldwell, J. A. R., & Coulson, I. M. 1987, AJ, 93, 1090
Campello, R. J. G. B., Moulavi, D., Zimek, A., & Sander, J. 2015, ACM Trans. Knowl. Discov. Data, 10, 1
Castro-Ginard, A., McMillan, P. J., Luri, X., et al. 2021, A&A, 652, A162
Chen, X., Wang, S., Deng, L., de Grijs, R., & Yang, M. 2018, ApJS, 237, 28
Chen, X., Wang, S., Deng, L., et al. 2019, Nat. Astron., 3, 320
Chen, X., Wang, S., Deng, L., et al. 2020, ApJS, 249, 18
Cheng, X., Anguiano, B., Majewski, S. R., et al. 2020, ApJ, 905, 49
Chernin, A. D. 1999, MNRAS, 308, 321
Chiappini, C., Minchev, I., Starkenburg, E., et al. 2019, The Messenger, 175, 30
Chrobáková, Ž., & López-Corredoira, M. 2021, ApJ, 912, 130
Churchwell, E., Babler, B. L., Meade, M. R., et al. 2009, PASP, 121, 213
Clementini, G., Ripepi, V., Molinaro, R., et al. 2019, A&A, 622, A60
Cutri, R. M., Wright, E. L., Conrow, T., et al. 2013, Explanatory Supplement to the AllWISE Data Release Products
da Silva, R., Crestani, J., Bono, G., et al. 2022, A&A, 661, A104
Dalton, G. 2016, in ASP Conf. Ser., 507, 97
Dambis, A. K., Berdnikov, L. N., Efremov, Y. N., et al. 2015, Astron. Lett., 41, 489
De Somma, G., Marconi, M., Cassisi, S., et al. 2020, MNRAS, 496, 5039
Debattista, V. P., & Sellwood, J. A. 1999, ApJ, 513, L107

Dékány, I., Hajdu, G., Grebel, E. K., & Catelan, M. 2019, ApJ, 883, 58
Dinnbier, F., Anderson, R. I., & Kroupa, P. 2022, A&A, 659, A169
Dobbs, C. L., & Pringle, J. E. 2010, MNRAS, 409, 396
D'Onghia, E., Vogelsberger, M., & Hernquist, L. 2013, ApJ, 766, 34
Drimmel, R. 2000, A&A, 358, L13
Drimmel, R., & Spergel, D. N. 2001, ApJ, 556, 181
Efremov, I. N. 1978, Sov. Astron., 22, 161
ESA 1997, The Hipparcos and Tycho Catalogues. Astrometric and Photometric Star Catalogues Derived from the ESA Hipparcos Space Astrometry Mission (Noordwijk, Netherlands: ESA Publications Division), 1200
Fabricius, C., Luri, X., Arenou, F., et al. 2021, A&A, 649, A5
Fernie, J. D. 1958, AJ, 63, 219
Fiorentino, G., Musella, I., & Marconi, M. 2013, MNRAS, 434, 2866
Frankel, N., Sanders, J., Ting, Y.-S., & Rix, H.-W. 2020, ApJ, 896, 15
Gaia Collaboration (Brown, A. G. A., et al.) 2021a, A&A, 649, A1
Gaia Collaboration (Smart, R. L., et al.) 2021b, A&A, 649, A6
Gaia Collaboration (Drimmel, R., et al.) 2022, A&A, in press https://doi.org/10.1051/0004-6361/202243797
Garavito-Camargo, N., Besla, G., Laporte, C. F. P., et al. 2019, ApJ, 884, 51
Genovali, K., Lemasle, B., Bono, G., et al. 2013, A&A, 554, A132
Genovali, K., Lemasle, B., Bono, G., et al. 2014, A&A, 566, A37
Georgelin, Y. M., & Georgelin, Y. P. 1976, A&A, 49, 57
Grand, R. J. J., Kawata, D., & Cropper, M. 2012, MNRAS, 421, 1529
GRAVITY Collaboration (Abuter, R., et al.) 2021, A&A, 647, A59
Groenewegen, M. A. T., Udalski, A., & Bono, G. 2008, A&A, 481, 441
Hao, C. J., Xu, Y., Hou, L. G., et al. 2021, A&A, 652, A102
Henderson, A. P., Jackson, P. D., & Kerr, F. J. 1982, ApJ, 263, 116
Honig, Z. N., & Reid, M. J. 2015, ApJ, 800, 53
Hotelling, H. 1931, Ann. Math. Stat., 2, 360
Hou, L. G. 2021, Front. Astron. Space Sci., 8, 103
Hou, L. G., & Han, J. L. 2014, A&A, 569, A125
Ibata, R. A., & Razoumov, A. O. 1998, A&A, 336, 130
Ivanov, G. R. 2008, Bulg. Astron. J., 10, 15
Jayasinghe, T., Kochanek, C. S., Stanek, K. Z., et al. 2018, MNRAS, 477, 3145
Jayasinghe, T., Stanek, K. Z., Kochanek, C. S., et al. 2019a, MNRAS, 485, 961
Jayasinghe, T., Stanek, K. Z., Kochanek, C. S., et al. 2019b, MNRAS, 486, 1907
Jiang, I.-G., & Binney, J. 1999, MNRAS, 303, L7
Juárez, M. A., & Steel, M. F. J. 2010, J. Bus. Econ. Stat., 28, 52
Kalberla, P. M. W., Dedes, L., Kerp, J., & Haud, U. 2007, A&A, 469, 511
Kendall, S., Kennicutt, R. C., & Clarke, C. 2011, MNRAS, 414, 538
Kerr, F. J. 1957, AJ, 62, 93
Kodric, M., Riffeser, A., Hopp, U., et al. 2018, AJ, 156, 130
Kounkel, M., Covey, K., & Stassun, K. G. 2020, AJ, 160, 279
Kovtyukh, V. V. 2007, MNRAS, 378, 617
Kovtyukh, V. V., & Gorlova, N. I. 2000, A&A, 358, 587
Kovtyukh, V., Lemasle, B., Bono, G., et al. 2022, MNRAS, 510, 1894
Kraft, R. P., & Schmidt, M. 1963, ApJ, 137, 249
Kuhn, M. A., Benjamin, R. A., Zucker, C., et al. 2021, A&A, 651, L10
Kumar, R., Carroll, C., Hartikainen, A., & Martin, O. 2019, J. Open Sour. Softw., 4, 1143
Laporte, C. F. P., Minchev, I., Johnston, K. V., & Gómez, F. A. 2019, MNRAS, 485, 3134
Lemasle, B., François, P., Bono, G., et al. 2007, A&A, 467, 283
Lemasle, B., François, P., Piersimoni, A., et al. 2008, A&A, 490, 613
Lemasle, B., François, P., Genovali, K., et al. 2013, A&A, 558, A31
Lemasle, B., Groenewegen, M. A. T., Grebel, E. K., et al. 2017, A&A, 608, A85
Lemasle, B., Hajdu, G., Kovtyukh, V., et al. 2018, A&A, 618, A160
Lemasle, B., Hanke, M., Storm, J., Bono, G., & Grebel, E. K. 2020, A&A, 641, A71
Levine, E. S., Blitz, L., & Heiles, C. 2006a, ApJ, 643, 881
Levine, E. S., Blitz, L., & Heiles, C. 2006b, Science, 312, 1773
Lian, J., Zasowski, G., Hasselquist, S., et al. 2022, MNRAS, 511, 5639
Lindegren, L., Klioner, S. A., Hernández, J., et al. 2021, A&A, 649, A2
López-Corredoira, M., Cabrera-Lavers, A., Garzón, F., & Hammersley, P. L. 2002, A&A, 394, 883
Luck, R. E., & Lambert, D. L. 2011, AJ, 142, 136
Luck, R. E., Andrievsky, S. M., Kovtyukh, V. V., Gieren, W., & Graczyk, D. 2011, AJ, 142, 51
Madore, B. F. 1982, ApJ, 253, 575
Majaess, D. J., Turner, D. G., & Lane, D. J. 2009, MNRAS, 398, 263
Mancino, S., Romaniello, M., Anderson, R. I., & Kudritzki, R. P. 2021, in ASP Conf. Ser., 529, 265
Matteucci, F., & Francois, P. 1989, MNRAS, 239, 885
McInnes, L., Healy, J., & Astels, S. 2017, J. Open Sour. Softw., 2, 205
Medina, G. E., Lemasle, B., & Grebel, E. K. 2021, MNRAS, 505, 1342
Meisner, A. M., Lang, D., Schlafly, E. F., & Schlegel, D. J. 2021, Res. Notes Am. Astron. Soc., 5, 168
Mel'nik, A. M., & Rautiainen, P. 2013, Astron. Nachr., 334, 777

Mel'nik, A. M., Rautiainen, P., Berdnikov, L. N., Dambis, A. K., & Rastorguev, A. S. 2015, Astron. Nachr., 336, 70
Metzger, M. R., Caldwell, J. A. R., & Schechter, P. L. 1998, AJ, 115, 635
Minchev, I., Chiappini, C., & Martig, M. 2013, A&A, 558, A9
Minchev, I., Chiappini, C., & Martig, M. 2014, A&A, 572, A92
Minniti, J. H., Zoccali, M., Rojas-Arriagada, A., et al. 2021, A&A, 654, A138
Miyamoto, M., Yoshizawa, M., & Suzuki, S. 1988, A&A, 194, 107
Mollá, M., Wekesa, S., Cavichia, O., et al. 2019, MNRAS, 490, 665
Momany, Y., Zaggia, S., Gilmore, G., et al. 2006, A&A, 451, 515
Monteiro, H., Barros, D. A., Dias, W. S., & Lépine, J. R. D. 2021, Front. Astron. Space Sci., 8, 62
Mróz, P., Udalski, A., Skowron, D. M., et al. 2019, ApJ, 870, L10
Nakanishi, H., & Sofue, Y. 2003, PASJ, 55, 191
Neeley, J. R., Monelli, M., Marengo, M., et al. 2021, ApJ, 920, 152
Oort, J. H., Kerr, F. J., & Westerhout, G. 1958, MNRAS, 118, 379
Ostriker, E. C., & Binney, J. J. 1989, MNRAS, 237, 785
Pedicelli, S., Lemasle, B., Groenewegen, M., et al. 2010, A&A, 518, A11
Pietrukowicz, P., Soszyński, I., & Udalski, A. 2021, Acta Astron., 71, 205
Pietrzyński, G., Graczyk, D., Gallenne, A., et al. 2019, Nature, 567, 200
Poggio, E., Drimmel, R., Andrae, R., et al. 2020, Nat. Astron., 4, 590
Poggio, E., Drimmel, R., Cantat-Gaudin, T., et al. 2021, A&A, 651, A104
Pojmanski, G. 1997, Acta Astron., 47, 467
Pojmanski, G. 2002, Acta Astron., 52, 397
Pont, F., Queloz, D., Bratschi, P., & Mayor, M. 1997, A&A, 318, 416
Price-Whelan, A. M. 2017, J. Open Sour. Softw., 2, 388
Proxauf, B., da Silva, R., Kovtyukh, V. V., et al. 2018, A&A, 616, A82
Radburn-Smith, D. J., de Jong, R. S., Streich, D., et al. 2014, ApJ, 780, 105
Reed, B. C. 1996, AJ, 111, 804
Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2019, ApJ, 885, 131
Reylé, C., Marshall, D. J., Robin, A. C., & Schultheis, M. 2009, A&A, 495, 819
Richards, J. W., Starr, D. L., Miller, A. A., et al. 2012, ApJS, 203, 32
Riello, M., De Angeli, F., Evans, D. W., et al. 2021, A&A, 649, A3
Ripepi, V., Molinaro, R., Musella, I., et al. 2019, A&A, 625, A14
Ripepi, V., Catanzaro, G., Molinaro, R., et al. 2021, MNRAS, 508, 4047
Roberts, W. W. 1969, ApJ, 158, 123
Romaniello, M., Primas, F., Mottini, M., et al. 2008, A&A, 488, 731
Romaniello, M., Riess, A., Mancino, S., et al. 2022, A&A, 658, A29
Romero-Gómez, M., Mateu, C., Aguilar, L., Figueras, F., & Castro-Ginard, A. 2019, A&A, 627, A150
Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. 2016, PeerJ Comput. Sci., 2, e55
Schlafly, E. F., Meisner, A. M., & Green, G. M. 2019, ApJS, 240, 30
Sellwood, J. A., & Masters, K. L. 2022, ARA&A, 60, 73

Shabani, F., Grebel, E. K., Pasquali, A., et al. 2018, MNRAS, 478, 3590
Skowron, D. M., Skowron, J., Mróz, P., et al. 2019a, Science, 365, 478
Skowron, D. M., Skowron, J., Mróz, P., et al. 2019b, Acta Astron., 69, 305
Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163
Smolec, R., & Moskalik, P. 2010, A&A, 524, A40
Soszyński, I., Udalski, A., Szymański, M. K., et al. 2015, Acta Astron., 65, 297
Soszyński, I., Udalski, A., Szymański, M. K., et al. 2017a, Acta Astron., 67, 103
Soszyński, I., Udalski, A., Szymański, M. K., et al. 2017b, Acta Astron., 67, 297
Soszyński, I., Udalski, A., Szymański, M. K., et al. 2020, Acta Astron., 70, 101
Sparke, L. S., & Casertano, S. 1988, MNRAS, 234, 873
Spitoni, E., Silva Aguirre, V., Matteucci, F., Calura, F., & Grisoni, V. 2019, A&A, 623, A60
Sziládi, K., Vinkó, J., Poretti, E., Szabados, L., & Kun, M. 2007, A&A, 473, 579
Taylor, J. H., & Cordes, J. M. 1993, ApJ, 411, 674
Udalski, A., Soszyński, I., Pietrukowicz, P., et al. 2018, Acta Astron., 68, 315
Vallée, J. P. 2020, New Astron., 76, 101337
Vallée, J. P. 2021, Int. J. Astron. Astrophys., 11, 445
van de Hulst, H. C., Muller, C. A., & Oort, J. H. 1954, Bull. Astron. Inst. Neth., 12, 117
van der Maaten, L., & Hinton, G. 2008, J. Mach. Learn. Res., 9, 2579
Vasilyev, V., Ludwig, H.-G., Freytag, B., Lemasle, B., & Marconi, M. 2017, A&A, 606, A140
Vasilyev, V., Ludwig, H.-G., Freytag, B., Lemasle, B., & Marconi, M. 2018, A&A, 611, A19
Vasilyev, V., Amarsi, A. M., Ludwig, H.-G., & Lemasle, B. 2019, A&A, 624, A85
Veena, V. S., Schilke, P., Sánchez-Monge, Á., et al. 2021, ApJ, 921, L42
Veselova, A. V., & Nikiforov, I. 2020, Res. Astron. Astrophys., 20, 209
Wang, S., & Chen, X. 2019, ApJ, 877, 116
Wang, H.-F., López-Corredoira, M., Huang, Y., et al. 2020, ApJ, 897, 119
Weinberg, M. D., & Blitz, L. 2006, ApJ, 641, L33
Wouterloot, J. G. A., Brand, J., Burton, W. B., & Kwee, K. K. 1990, A&A, 230, 21
Xu, Y., Hou, L. G., Bian, S. B., et al. 2021, A&A, 645, L8
Yong, D., Carney, B. W., Teixera de Almeida, M. L., & Pohl, B. L. 2006, AJ, 131, 2256
Yu, Y., Wang, H.-F., Cui, W.-Y., et al. 2021, ApJ, 922, 80
Yusifov, I. 2004, in The Magnetized Interstellar Medium, eds. B. Uyaniker, W. Reich, & R. Wielebinski (Katlenburg-Lindau: Copernicus GmbH), 165
Zari, E., Rix, H. W., Frankel, N., et al. 2021, A&A, 650, A112

Appendix A: The Milky Way warp

A.1. Length of a bow following the warp

We adopt the definition of the warp proposed by Skowron et al. (2019b), where the Milky Way is warped at radii exceeding a given radius r_0. At such distances, the warp is given by the equation:

$$z(r, \Theta) = z_0 + (r - r_0)^2 \times [z_1 \, sin(\Theta - \Theta_1) + z_2 \, sin(2\,(\Theta - \Theta_2))]$$

where z is the vertical distance from the Galactic plane, r is the distance from the Galactic center, and Θ is the Galactocentric azimuth. The constants r_0, z_0, z_1, z_2, Θ_1, Θ_2 are radial, vertical, and angular parameters describing the surface.

With such a definition, the warp has a parabolic shape, which varies with the Galactocentric azimuth Θ. Indeed, if we set the position of a given Cepheid as $(r_{cep}, \Theta_{cep}, z_{cep})$ in Galactocentric coordinates, the equation of the warp is of the form:

$$z = z_0 + C \times (r - r_0)^2 \tag{A.1}$$

where

$$C = z_1 \, sin(\Theta_{cep} - \Theta_1) + z_2 \, sin(2\,(\Theta_{cep} - \Theta_2))$$

A.2. The warp model

in the vertical plane containing the Galactic center and the Cepheid at the Galactocentric azimuth Θ_{cep}.

The length of the bow between the Galactic center and a given Cepheid located at the distance d_{GC} from the Galactic center is given by the integral equation:

$$l_{bow} = \int_0^{d_{GC}\cos(\phi)} \sqrt{1 + \left(\frac{dz}{dr}\right)^2} \, dr \tag{A.2}$$

where $z = f(r)$ is the function describing the warp and $D = d_{GC} \cos \phi$ is the Galactocentric distance of the Cepheid projected on the Galactic plane.

Given the definition of the warp given by Skowron et al. (2019b), we have to integrate:

$$l_{bow} = \int_0^D \sqrt{4C^2\,(r - r_0)^2 + 1} \, dr \tag{A.3}$$

with $D = d_{GC} \cos \phi$. After a series of substitutions of variables, it is possible to analytically determine the value of l_{bow}. In this work we derived it numerically using SCIPY.
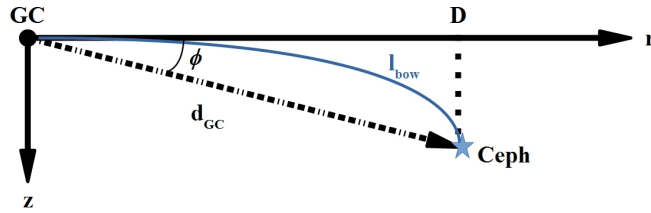

Fig. A.1. Slice of the (warped) Galactic disk in the Galactocentric azimuth θ containing a Cepheid, where d_{GC} is the Galactocentric distance of the Cepheid, D its projection on the Galactic plane, and ϕ its Galactocentric latitude.
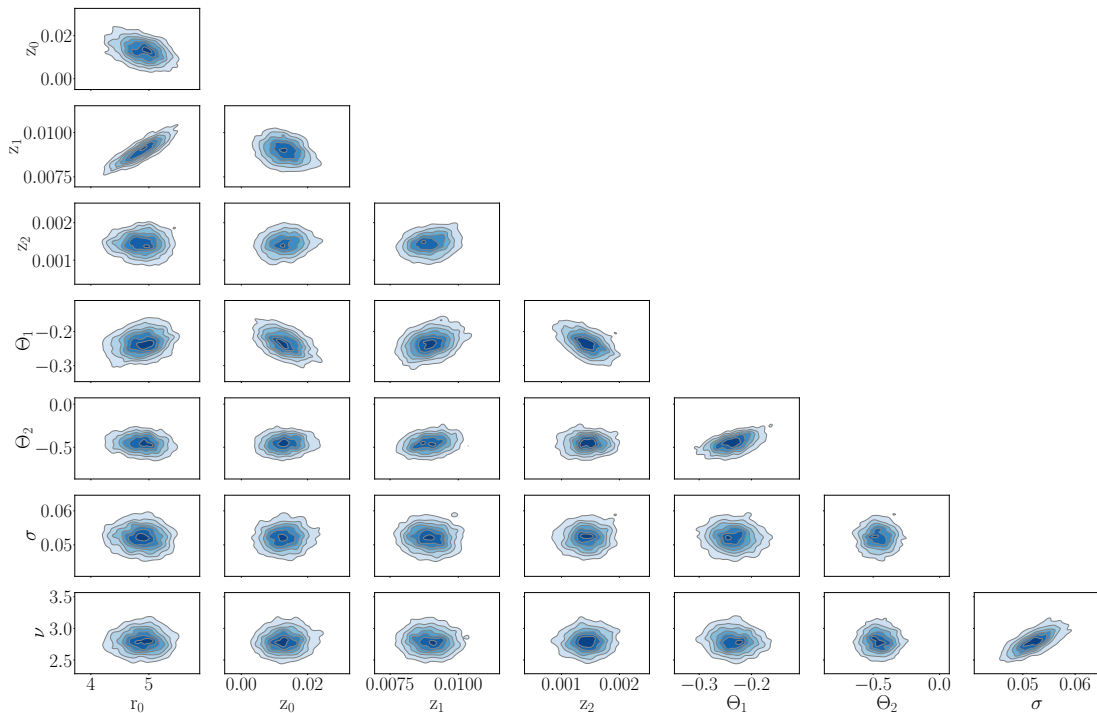

Fig. A.2. Pairwise correlations of the posterior distributions of the structural parameters r_0, z_0, z_1, z_2, Θ_1, Θ_2 of the warp model, together with the standard deviation σ and the normality parameter ν of the model.
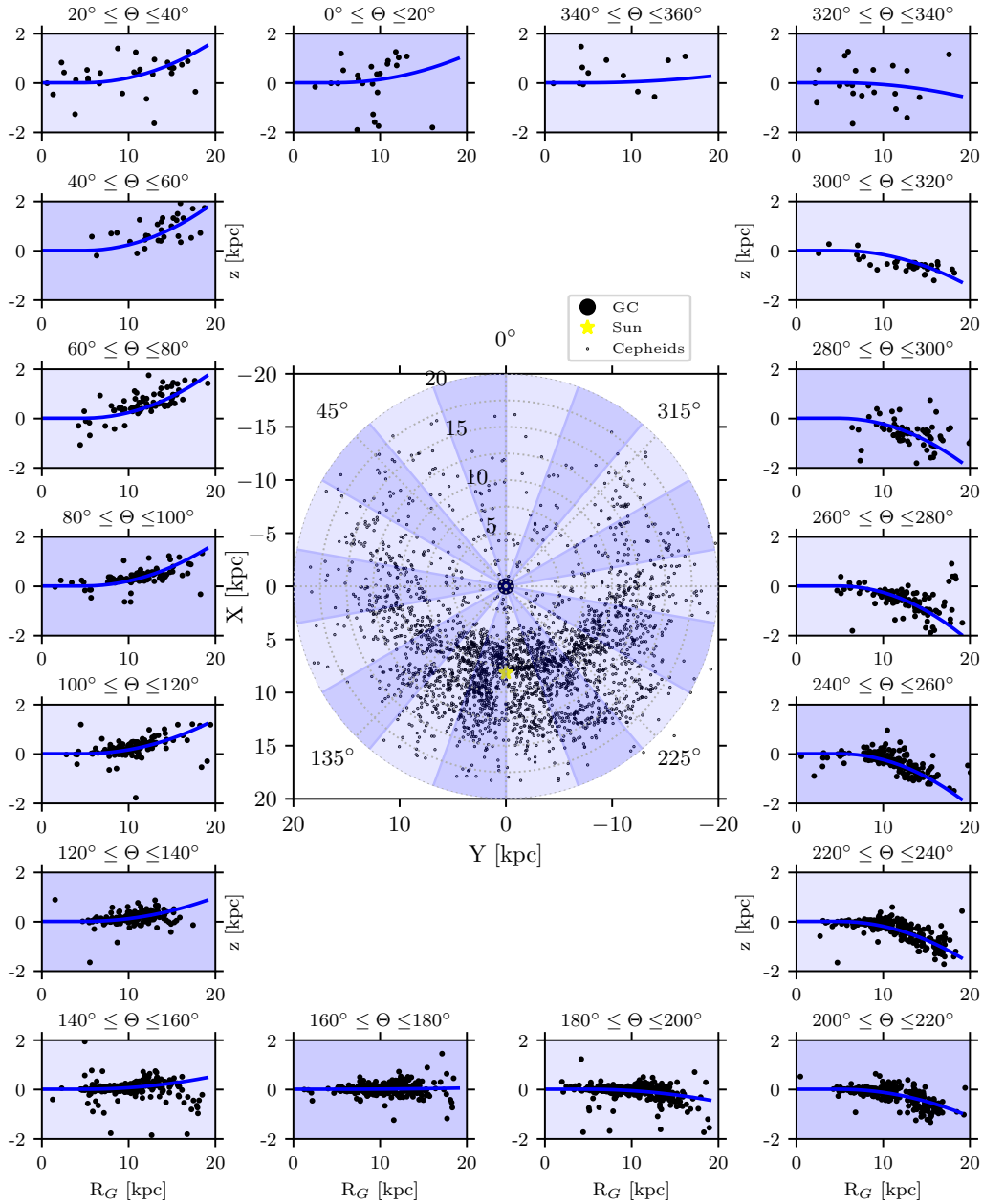
Fig. A.3. Galactic warp traced by classical Cepheids. The Milky Way disk is divided into 18 slices. The vertical distribution of Cepheids within each slice is displayed in adjacent panels, showing the distance from the Galactic plane z as a function of the Galactocentric distance r. The black dots represent individual Cepheids while the blue line is the warp model, computed with the parameters listed in Table 3 and for the median Galactocentric azimuth of a given slice.

Table A.1. Covariance matrix for the Bayesian robust regression of the warp model.

	r_0	z_0	z_1	z_2	Θ_1	Θ_2	σ	ν
r_0	0.0983427	-0.000584572	0.000173638	4.72231e-07	0.00244786	-0.00333678	-7.20902e-05	-0.00166704
z_0	-0.000584572	2.41349e-05	-8.84631e-07	1.83132e-07	-8.14438e-05	3.50388e-05	1.55684e-06	5.25924e-05
z_1	0.000173638	-8.84631e-07	4.20396e-07	3.95521e-08	7.20249e-06	1.64115e-05	-1.53575e-07	-5.55467e-06
z_2	4.72231e-07	1.83132e-07	3.95521e-08	7.70687e-08	-4.86087e-06	8.99776e-08	3.91017e-08	1.83853e-06
Θ_1	0.00244786	-8.14438e-05	7.20249e-06	-4.86087e-06	0.00112674	0.00154508	-7.66386e-06	-0.000389352
Θ_2	-0.00333678	3.50388e-05	1.64115e-05	8.99776e-08	0.00154508	0.00955416	-9.70985e-07	-0.000574705
σ	-7.20902e-05	1.55684e-06	-1.53575e-07	3.91017e-08	-7.66386e-06	-9.70985e-07	1.09651e-05	0.000380751
ν	-0.00166704	5.25924e-05	-5.55467e-06	1.83853e-06	-0.000389352	-0.000574705	0.000380751	0.0283286

Notes. $r_0, z_0, z_1, z_2, \Theta_1, \Theta_2$ are the structural parameters of the warp model, σ its standard deviation and ν its normality parameter.

Appendix B: Testing the algorithm

B.1. Creation of mock spiral arms

To create mock spiral arms, we rely on the Reid et al. (2019) model. We selected it over other possible choices (e.g., Hou 2021) because the immediate proximity of the Norma and Sct-Cen arms, especially in the vicinity of the kink in the Norma arm, is very challenging for our algorithm. We want to stress here that only a few Cepheids are currently known at the expected locations of the 3-kpc and Norma arms (see Fig. 5). We then select N equidistant points along the spiral arms traced by Reid et al. (2019) to populate the spiral arms with mock Cepheids by using N bivariate normal distributions. Such a distribution is defined by its mean and covariance matrix, which are analogous to the mean and variance of a 1D normal distribution. In our case, the means are the (x,y) positions of each of the N reference points and the covariance matrix is $\begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$.

The values for the parameters of the bivariate distribution are tabulated in Table B.1. The values of σ are proportional to the widths of the individual spiral arms proposed by Reid et al. (2019). With these parameters, several ten thousands mock Cepheids are created. For each individual arm, a very small number of them lies outside the angular domain encompassed by the Reid et al. (2019) model, they are discarded for comparison purposes (see Sect. B.2). This sharpens the extremities of the mock spiral arms, which would otherwise appear rounded because of stars generated by the very first and very last bivariate distributions. Finally, N_{spiral} mock Cepheids are randomly selected as our spiral arms Cepheids sample.

B.2. Comparison between the mock and the retrieved sample: An ideal test case

To quantify the comparison between the original mock data and the retrieved sample, we use Hotelling's t-Squared (t^2, Hotelling 1931) statistics. It is the multivariate generalization of the Student's t-distribution, and it is specifically useful when comparing two distributions of unequal sizes, mean values and variances. Our null hypothesis is that the original and retrieved samples for a given spiral arm are drawn from the same parent distribution.

Fig. B.1 displays the original mock data (N_{spiral}=1500 Cepheids) in the Milky Way plane and in the t-SNE space. Given its small angular extension in the Reid et al. (2019) model, the 3-kpc arm is usually populated with only a handful of Cepheids (or less) and will be ignored in what follows. In all the tests, we did not try to adjust the hyper-parameters of t-SNE+HDBSCAN to the specifics of the mock data. and we kept the same values as for the real data. The very high values for the p-values and the recovery fractions in Table B.2 certify that the spiral arms are almost perfectly recovered by the algorithm. This can be seen as well in Fig. B.2. The Norma and the Sct-Cen arms are recovered via two segments resulting from under-densities or gaps in the spatial distribution of Cepheids.

Increasing the number of stars in the mock catalog reduces the number and extension of such gaps, leading to the recovery of mock spiral arms within a single structure. In Fig. B.1, we note that the other arms (e.g., the Perseus arm) also show gaps in the distribution of mock Cepheids, without being split into segments. From this, we conclude that segments are more likely to occur in crowded regions of the t-SNE space.
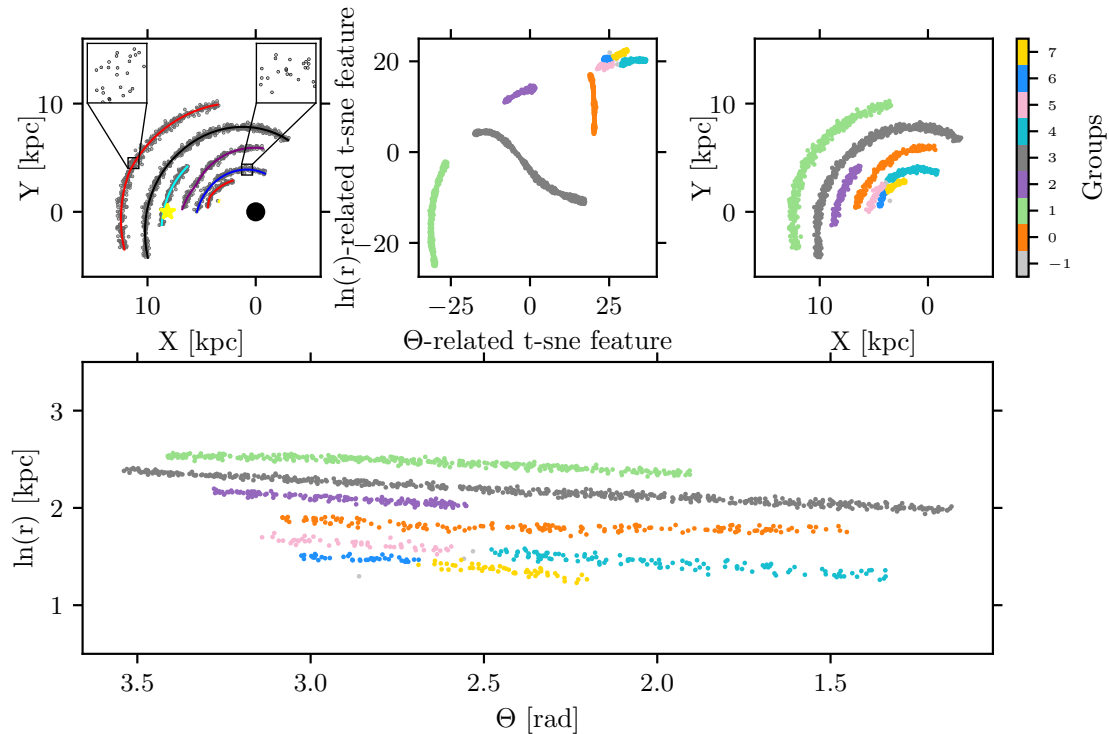

Fig. B.1. Mock sample of N_{spiral}=1500 Cepheids in spiral arms following the model by Reid et al. (2019) (top left panel). Groups of Cepheids identified by HDBSCAN in the t-SNE space (top middle panel). The same groups are presented in the (Θ, $\ln r$) space (bottom panel), where Θ is the Galactocentric azimuth and $\ln r$ the logarithm of the Galactocentric radius (corrected from the warp), and in the Galactic plane (top right panel).
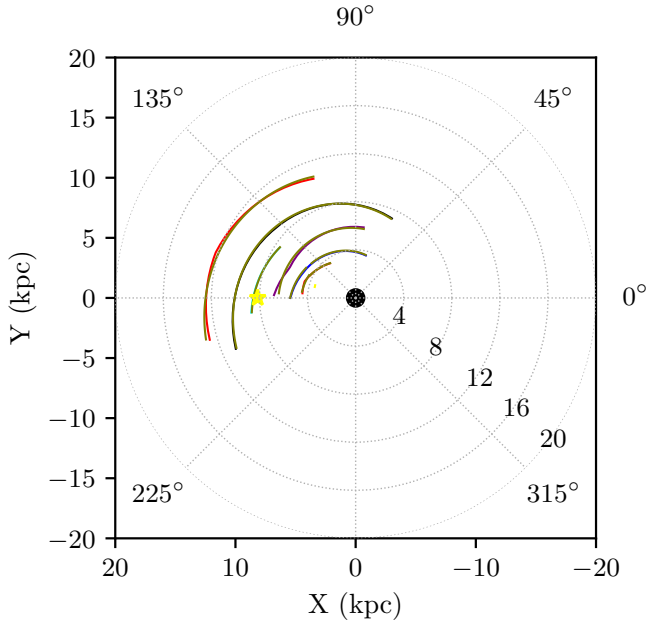
Fig. B.2. Retrieved spiral segments (olive) over-plotted on the original spiral arms delineated by Reid et al. (2019) used to define the mock sample of stars. The color-coding of the spiral arms is the same as in the original paper.
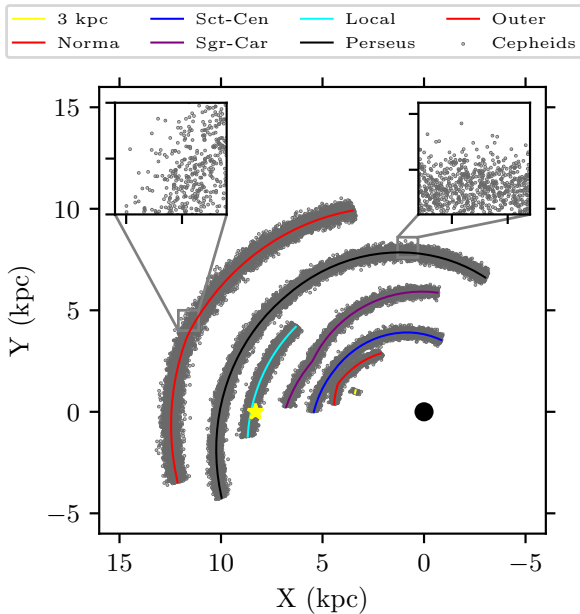

Fig. B.3. Full sample of mock Cepheids in spiral arms following the model by Reid et al. (2019). The noise parameter in this example is 0.02. Cepheids created outside of the angular range covered by a given spiral arm have been discarded. From this sample, N_{spiral} mock Cepheids are randomly selected.

B.3. Test case 2: Noisy spiral arms

In this test case, we increase the dispersion of mock Cepheids in spiral arms by adding random noise to their (x,y) coordinates. To achieve this, we added random values to them, which were sampled from a univariate standard normal distribution scaled by the "noise" parameter (see Table B.1). In this example, the noise parameter is 0.03.

Table B.1. Parameters for the creation of mock spiral arms based on the Reid et al. (2019) model.

	σ^2	N	Noise parameter		
			Ideal	Noisy	Inter arms
3 kpc	0.014	20	0.0	0.03	0.01
Norma	0.011	20	0.0	0.03	0.01
Sct-Cen	0.018	20	0.0	0.03	0.01
Sgr-Car	0.021	20	0.0	0.03	0.01
Local	0.024	50	0.0	0.03	0.01
Perseus	0.027	50	0.0	0.03	0.01
Outer	0.050	50	0.0	0.03	0.01

Notes. N is the number of centers for bivariate normal distributions with variance σ^2 in their covariance matrix. Noise is the scaling factor for random noise added to the coordinates of the mock stars. Random noise is drawn from an univariate normal distribution of mean 0 and variance 1.

Table B.2. Groups retrieved by the algorithm that correspond to a given spiral arm.

Arm	Group	p-value	N_{spiral}	Retrieved	% retrieved
Norma	[7, 6]	0.98	93	92	98.92
Sct-Cen	[4, 5]	1.00	169	168	99.41
Sgr-Car	[0]	1.00	167	167	100.00
Local	[2]	1.00	158	158	100.00
Perseus	[3]	1.00	550	550	100.00
Outer	[1]	1.00	362	362	100.00

Notes. The number and percentage of retrieved stars compared to the original number N_{spiral} of mock stars in the spiral arm is provided, as well as the p-value testing the null hypothesis that the mock and retrieved spiral arms are identical using Hotelling's t-Squared statistics.

Table B.3. Same as Table B.2.

Arm	Group	p-value	N_{spiral}	Retrieved	% retrieved
Norma	[4, 8]	0.94	92	91	98.91
Sct-Cen	[5, 6, 7, 8]	0.01	150	142	94.67
Sgr-Car	[1]	1.00	159	159	100.00
Local	[2]	1.00	161	161	100.00
Perseus	[3]	1.00	553	553	100.00
Outer	[0]	1.00	384	384	100.00

Notes. In this test case, random noise is added to the (x,y) coordinates of the stars in the mock spiral arms.

Figs. B.4 and B.5 display the outcome of this test performed on a sample of N_{spiral}=1500 Cepheids. The algorithm performs similarly well as in the ideal case, and the mock spiral arms are very well reproduced. However, despite a recovery rate approaching 95%, the p-value for the Sct-Cen strongly drops. A careful inspection of Fig. B.4 indicates that two stars attributed to group 8 rather belong to group 7. This misclassification leads to the inclusion of group 8 in the comparison to the original mock data for the Sct-Cen, and hence to the low p-value. We also note that a few stars close to the gap in the Sct-Cen arm remain unclassified (probably due to their proximity to the Norma arm),

This example remains a bit unrealistic because mock Cepheids stay confined to the spiral arms, and because the proximity of the Norma and Sct-Cen spiral arms prevents us from increasing the dispersion of individual spiral arms to very large values: Fig. B.3 shows that the mock Norma and Sct-Cen spi-
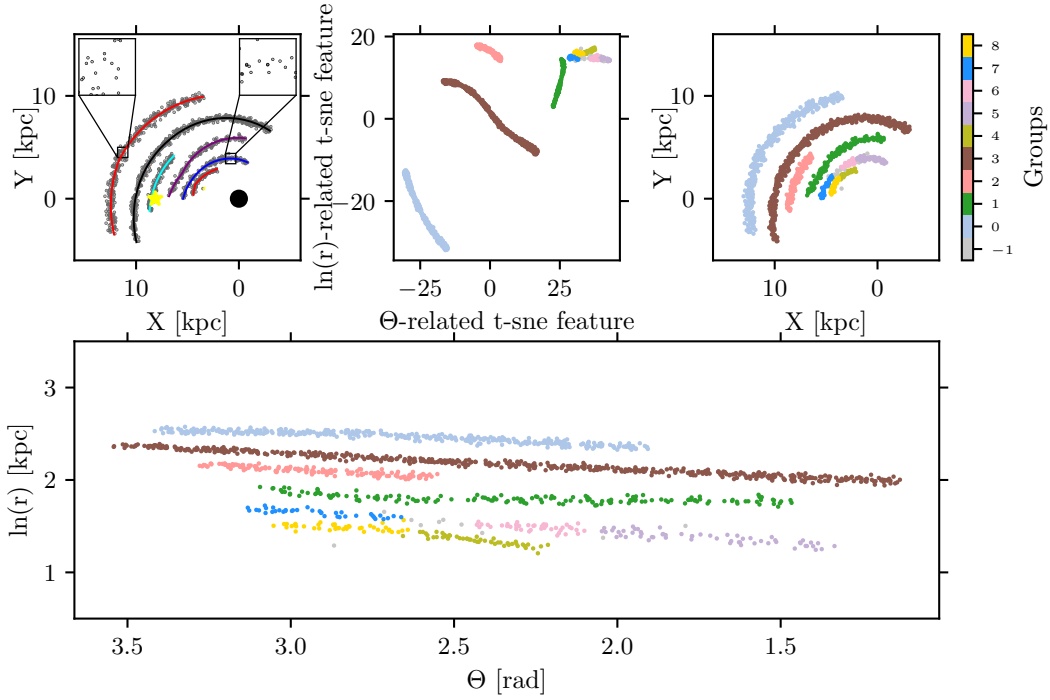
Fig. B.4. Same as Fig. B.1, but with random noise added to the (x,y) coordinates of the stars in the mock spiral arms.
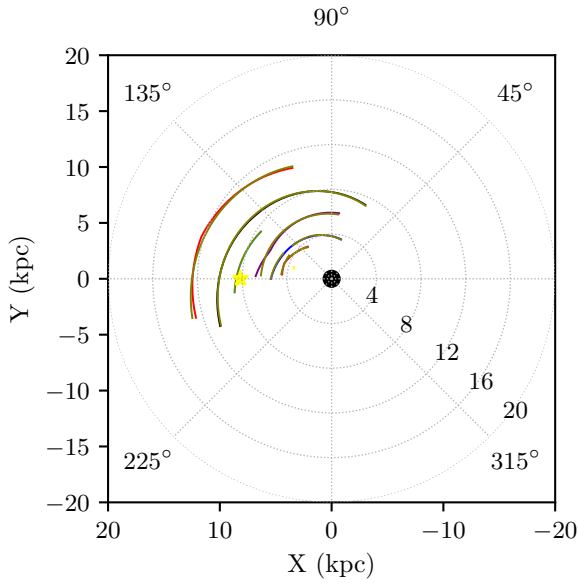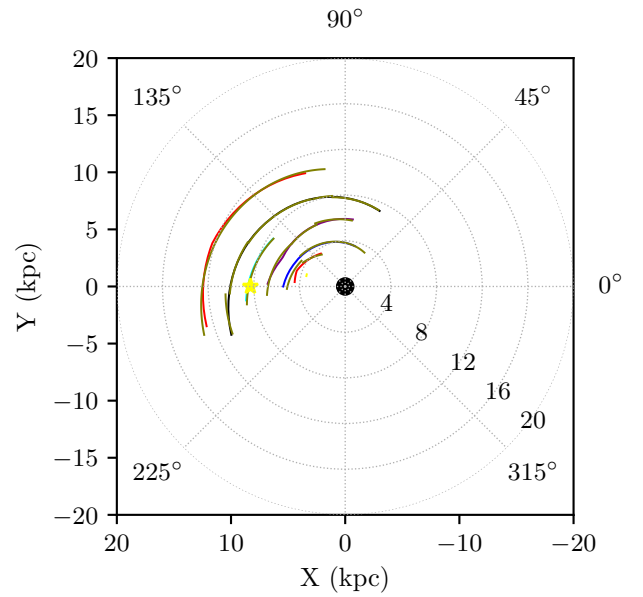

Fig. B.5. Same as Fig. B.2.

Fig. B.6. Same as Fig. B.2.

ral arms come close to touching each other already with a noise parameter equals to 0.02.

B.4. Test case 3: Inter-arms' Cepheids

To provide a more realistic test case, we add to the mock sample of Cepheids in spiral arms (N_{spiral}=900, noise=0.01) a large collection of N_{other}=1500 Cepheids, whose coordinates are drawn from a bivariate normal distribution with a mean centered on the Sun and a variance of 80. These Cepheids can be considered as inter-arm Cepheids, although by construction some of them may overlap the spiral arms or be located in their immediate proximity.

In this example, almost all spiral arm members are retrieved, but the t-SNE groups also wrongly include ≈15% to ≈55% non-members, leading to lower p-values, and, in the case of the Norma and Sct-Cen arm, to a partial mismatch between the input and retrieved spiral arms (also leading to low p-values for these two arms). Indeed, Fig. B.6 shows that in the region where they are closest to each other, Norma and Sct-Cen are merged into a single structure. Fig. B.6 also shows that the other spiral arms are well retrieved, sometimes by the means of several segments. An artificial structure (group 9) emerges from the noise because it is relatively isolated in the t-SNE space and could have been falsely identified as a real structure beyond the outer arm.

Table B.4. Same as Table B.2.

Arm	Group	p-value	N_{spiral}	Retrieved	% retrieved	Extra	% extra
Norma	[7, 4]	0.00	56	56	100.00	11	19.64
Sct-Cen	[8, 4]	0.47	108	108	100.00	25	23.15
Sgr-Car	[2, 6, 5]	0.02	98	98	100.00	35	35.71
Local	[0]	0.27	88	88	100.00	24	27.27
Perseus	[14, 16, 15, 13, 12, 17]	0.43	341	337	98.83	49	14.37
Outer	[1]	0.00	208	208	100.00	113	54.33

Notes. In this test case, random noise is added to the (x,y) coordinates of the stars in the mock spiral arms, and a large number of inter-arms' Cepheids is included in the sample. The table also indicates the number and percentage of inter-arms' stars wrongly included in the spiral arms.
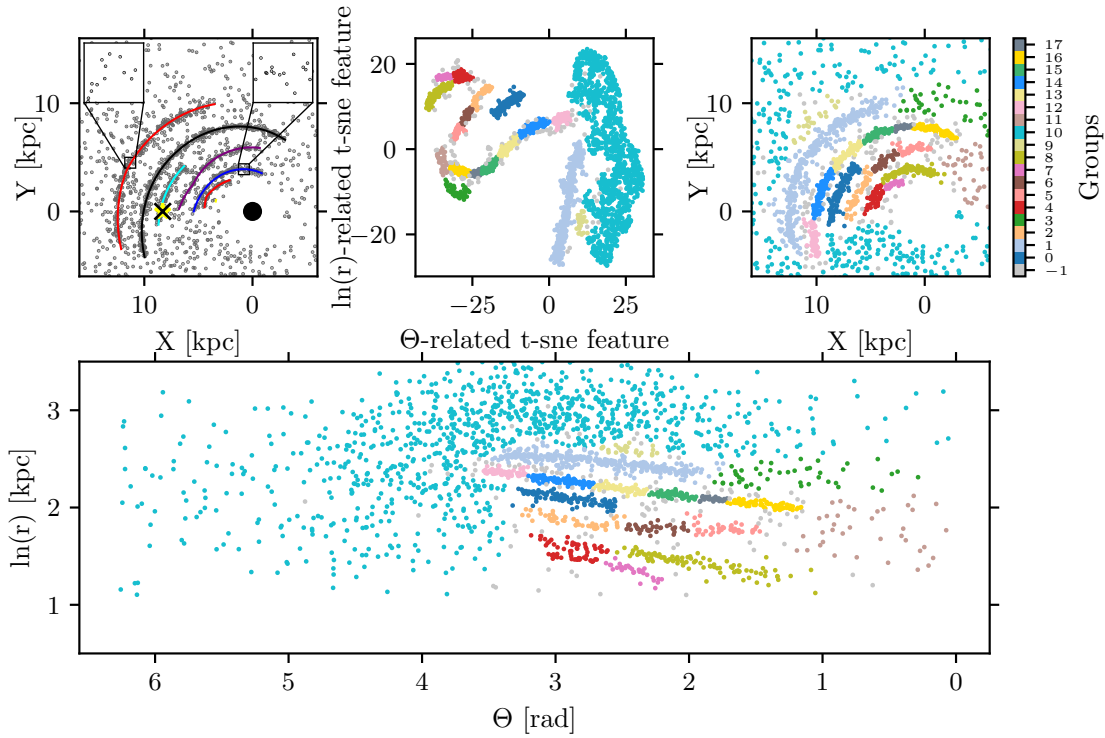

Fig. B.7. Same as Fig. B.1, but with random noise added to the (x,y) coordinates of the N_{spiral}=900 stars in the mock spiral arms, and N_{other}=1500 inter-arms' Cepheids.
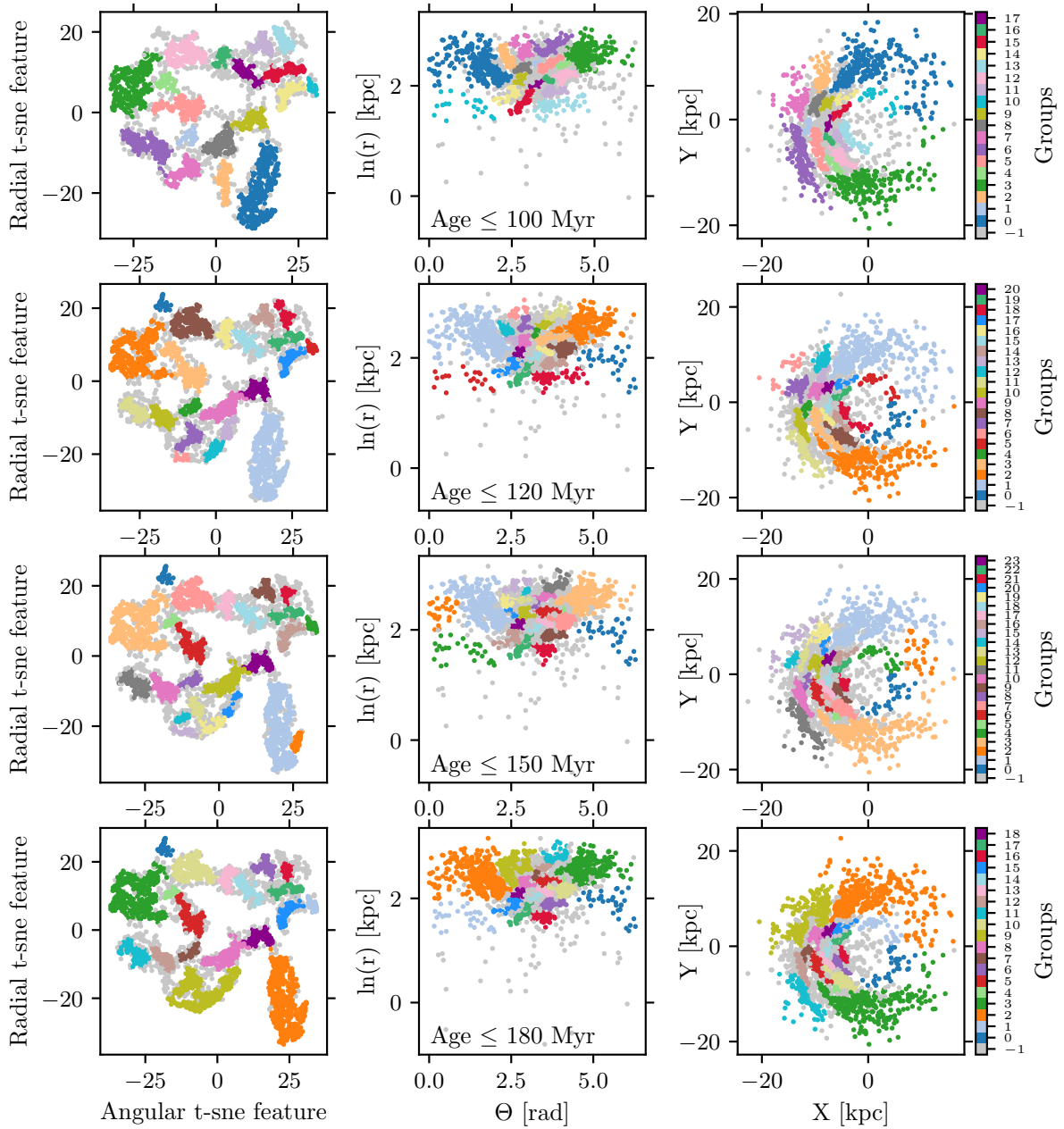
Appendix C: Characterization of spiral arms

Fig. C.1. Outcome of the algorithm for different age cuts. The left panels show groups of Cepheids identified via `HDBSCAN` in the `t-SNE` space. These groups are then displayed in the $(\theta, \ln r)$ plane (middle panels), where θ is the Galactocentric azimuth and $\ln r$ the logarithm of the Galactocentric radius (corrected from the warp), and in the Galactic plane (right panels), where the Galactic center is located at (0,0). The groups have been identified in different subsamples of the original Cepheid catalog by applying different age cuts, from Cepheids younger than 100 Myr (top row) to Cepheids younger than 180 Myr (bottom row).

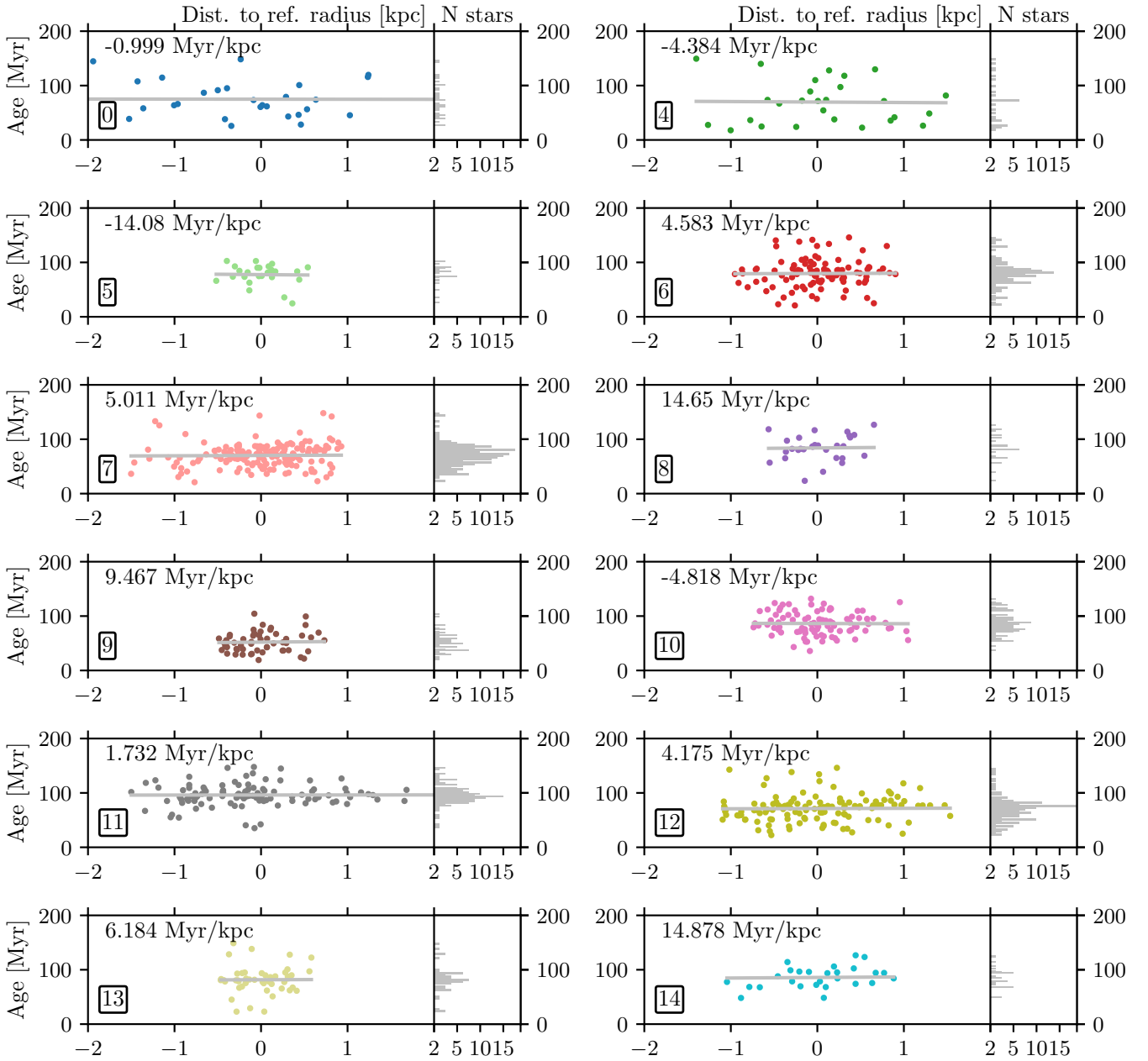**Fig. C.2.** Age distribution of Cepheids in individual spiral segments or groups of segments. The age (in megayears) is plotted against the distance (in kiloparsecs) to the reference radius of the segment. A linear fit to the data indicates the age gradient across the considered segment (the slope of the gradient is provided). As in Figs. 7–11, only Cepheids younger than 150 Myr are considered.
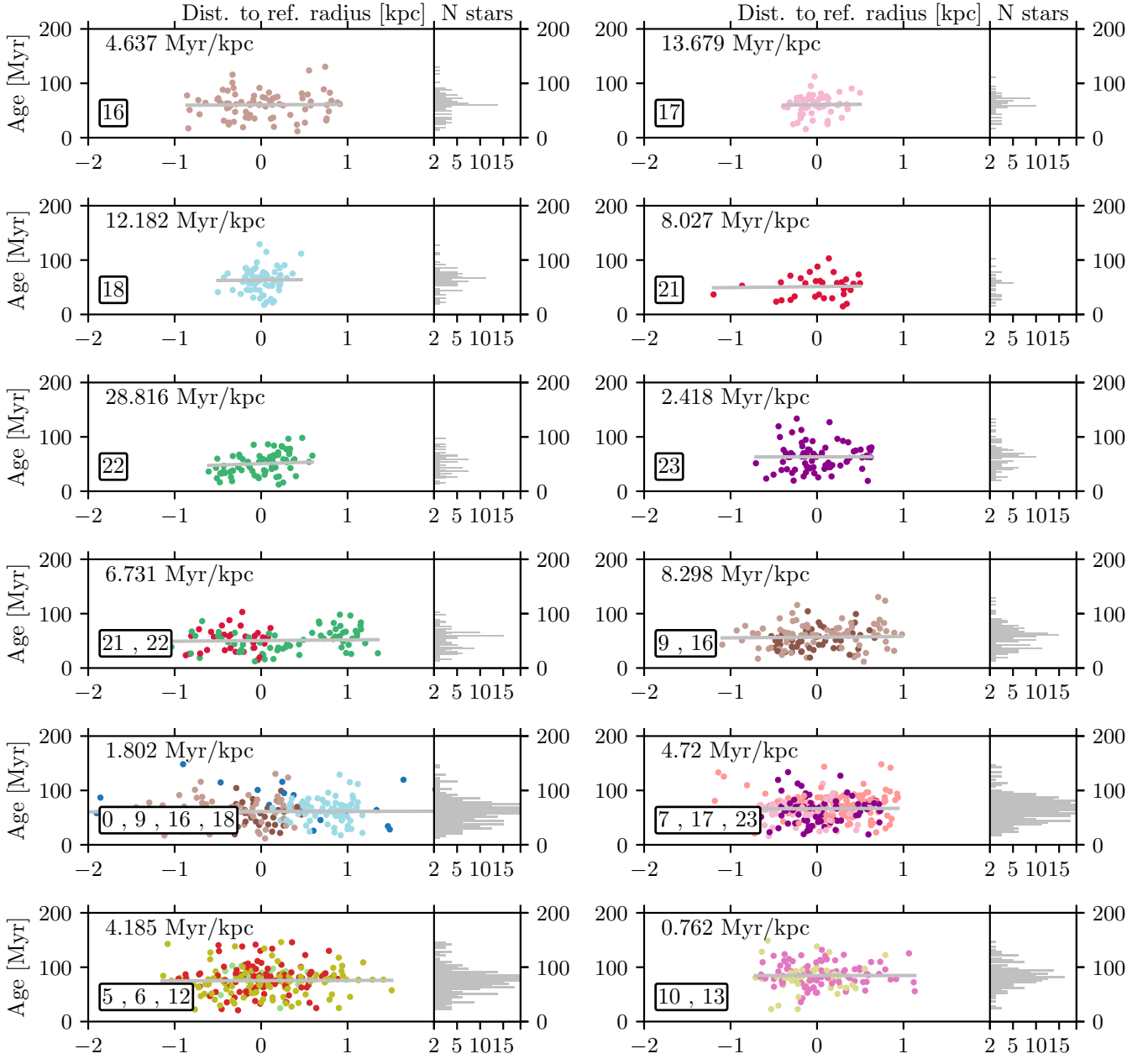
Fig. C.2. continued.

Appendix D: Spiral segments: Alternative age cuts
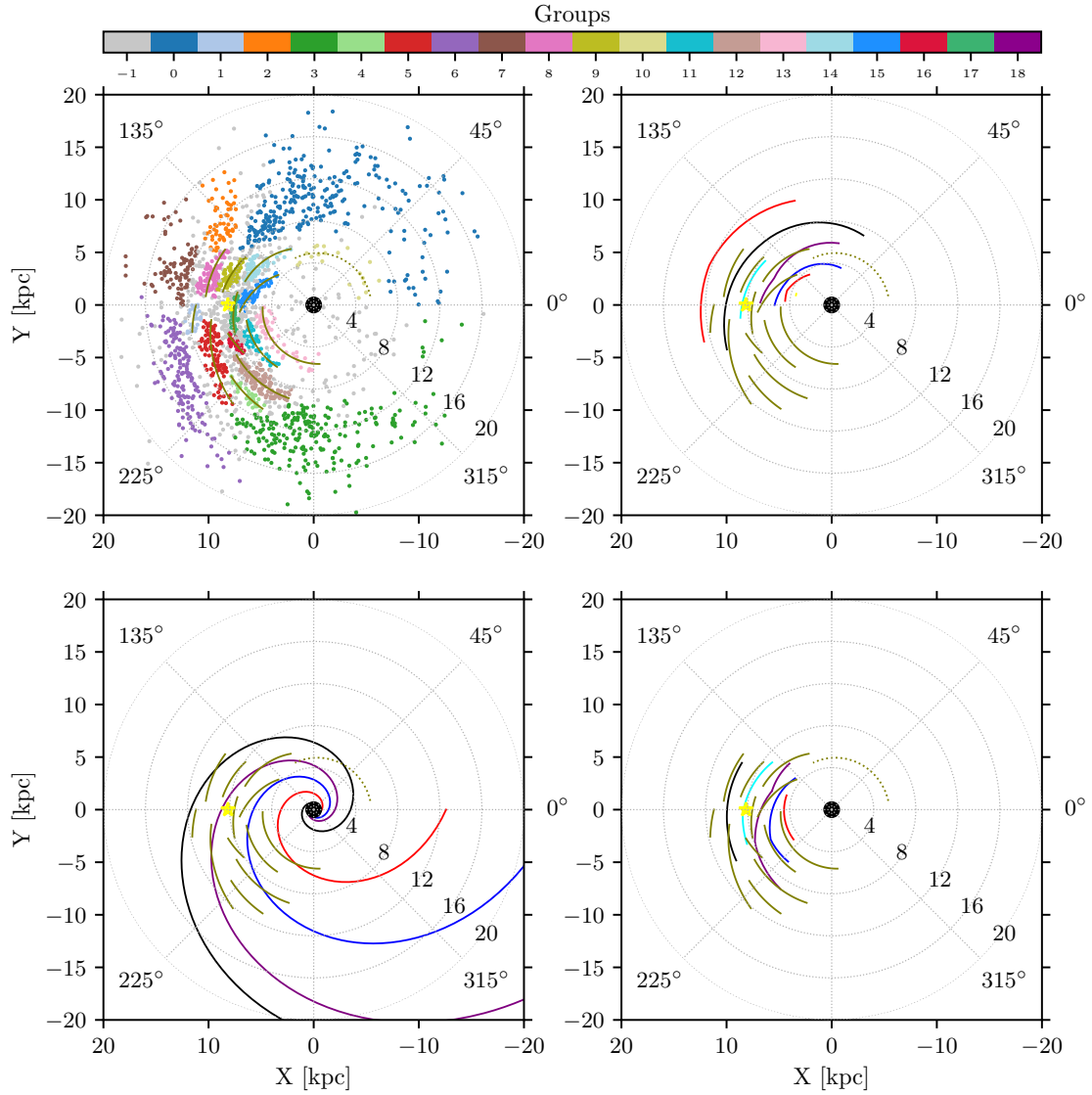
D.1. Cepheids younger than 100 Myr

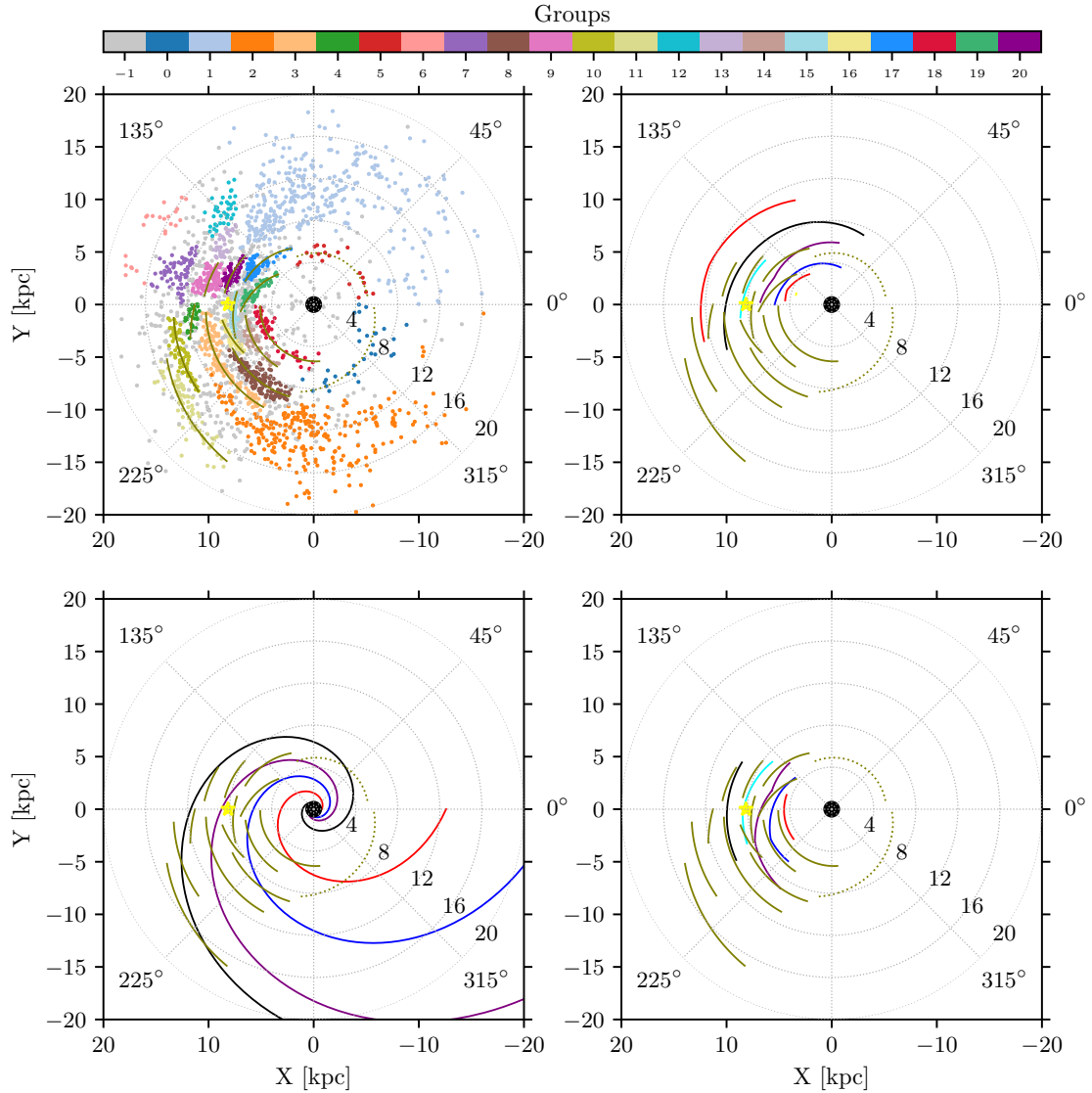**Fig. D.1.** Same as Fig. 8, but for Cepheids younger than 100 Myr.

Table D.1. Same as Table 4, but for Cepheids younger than 100 Myr.

Group	Slope (kpc.rad^{-1})	σ_{slope} (kpc.rad^{-1})	Intercept (kpc)	$\sigma_{intercept}$ (kpc)	Ref. angle (rad)	Ref. radius (kpc)	ln(Ref. radius) (kpc)	Pitch angle (rad)	Min. angle (rad)	Max. angle (rad)
[10]	-0.071	0.060	1.706	0.073	1.053	5.110	1.631	0.071	0.154	1.951
[13]	0.092	0.027	1.287	0.105	4.001	5.234	1.655	-0.092	3.190	4.812
[15]	0.591	0.042	0.043	0.123	2.835	5.575	1.718	-0.534	2.436	3.233
[14]	0.215	0.049	1.327	0.126	2.384	6.294	1.840	-0.212	1.961	2.807
[11]	0.052	0.033	1.706	0.122	3.776	6.702	1.902	-0.052	3.375	4.177
[17]	0.166	0.018	1.500	0.060	3.232	7.664	2.037	-0.164	2.980	3.485
[9]	0.211	0.047	1.535	0.132	2.765	8.318	2.118	-0.208	2.534	2.996
[16]	-0.242	0.059	2.990	0.212	3.602	8.318	2.118	0.237	3.467	3.736
[12]	0.065	0.025	1.927	0.100	4.084	8.957	2.192	-0.065	3.718	4.450
[8]	0.073	0.042	2.108	0.120	2.821	10.114	2.314	-0.073	2.583	3.060
[4]	0.148	0.045	1.768	0.183	4.055	10.677	2.368	-0.147	3.855	4.255
[5]	0.283	0.028	1.356	0.102	3.656	10.921	2.391	-0.276	3.288	4.025
[1]	0.252	0.086	1.624	0.281	3.251	11.510	2.443	-0.247	3.141	3.360

D.2. Cepheids younger than 120 Myr

Table D.2. Same as Table 4, but for Cepheids younger than 120 Myr.

Group	Slope (kpc.rad^{-1})	σ_{slope} (kpc.rad^{-1})	Intercept (kpc)	$\sigma_{intercept}$ (kpc)	Ref. angle (rad)	Ref. radius (kpc)	ln(Ref. radius) (kpc)	Pitch angle (rad)	Min. angle (rad)	Max. angle (rad)
[5]	-0.039	0.051	1.650	0.063	1.053	4.998	1.609	0.039	0.154	1.951
[18]	0.039	0.025	1.503	0.094	3.994	5.253	1.659	-0.039	3.177	4.812
[19]	0.570	0.043	0.099	0.125	2.835	5.556	1.715	-0.518	2.436	3.233
[17]	0.212	0.049	1.333	0.125	2.384	6.287	1.838	-0.209	1.961	2.807
[14]	0.046	0.031	1.731	0.116	3.743	6.707	1.903	-0.046	3.309	4.177
[0]	-0.228	0.073	3.169	0.405	5.424	6.906	1.932	0.224	4.566	6.282
[15]	0.157	0.018	1.533	0.059	3.260	7.728	2.045	-0.156	2.980	3.541
[20]	0.134	0.042	1.740	0.118	2.770	8.258	2.111	-0.133	2.534	3.007
[16]	-0.206	0.026	2.859	0.093	3.527	8.436	2.133	0.203	3.317	3.736
[8]	0.039	0.024	2.030	0.098	4.044	8.915	2.188	-0.039	3.639	4.450
[9]	0.152	0.073	1.878	0.211	2.896	10.157	2.318	-0.151	2.733	3.060
[3]	0.054	0.013	2.156	0.048	3.737	10.568	2.358	-0.054	3.220	4.255
[4]	0.281	0.072	1.541	0.233	3.274	11.716	2.461	-0.274	3.141	3.406
[10]	0.046	0.021	2.444	0.073	3.511	13.538	2.605	-0.046	3.237	3.785
[11]	0.190	0.028	2.036	0.107	3.851	15.923	2.768	-0.188	3.498	4.205

Fig. D.2. Same as Fig. 8, but for Cepheids younger than 120 Myr.
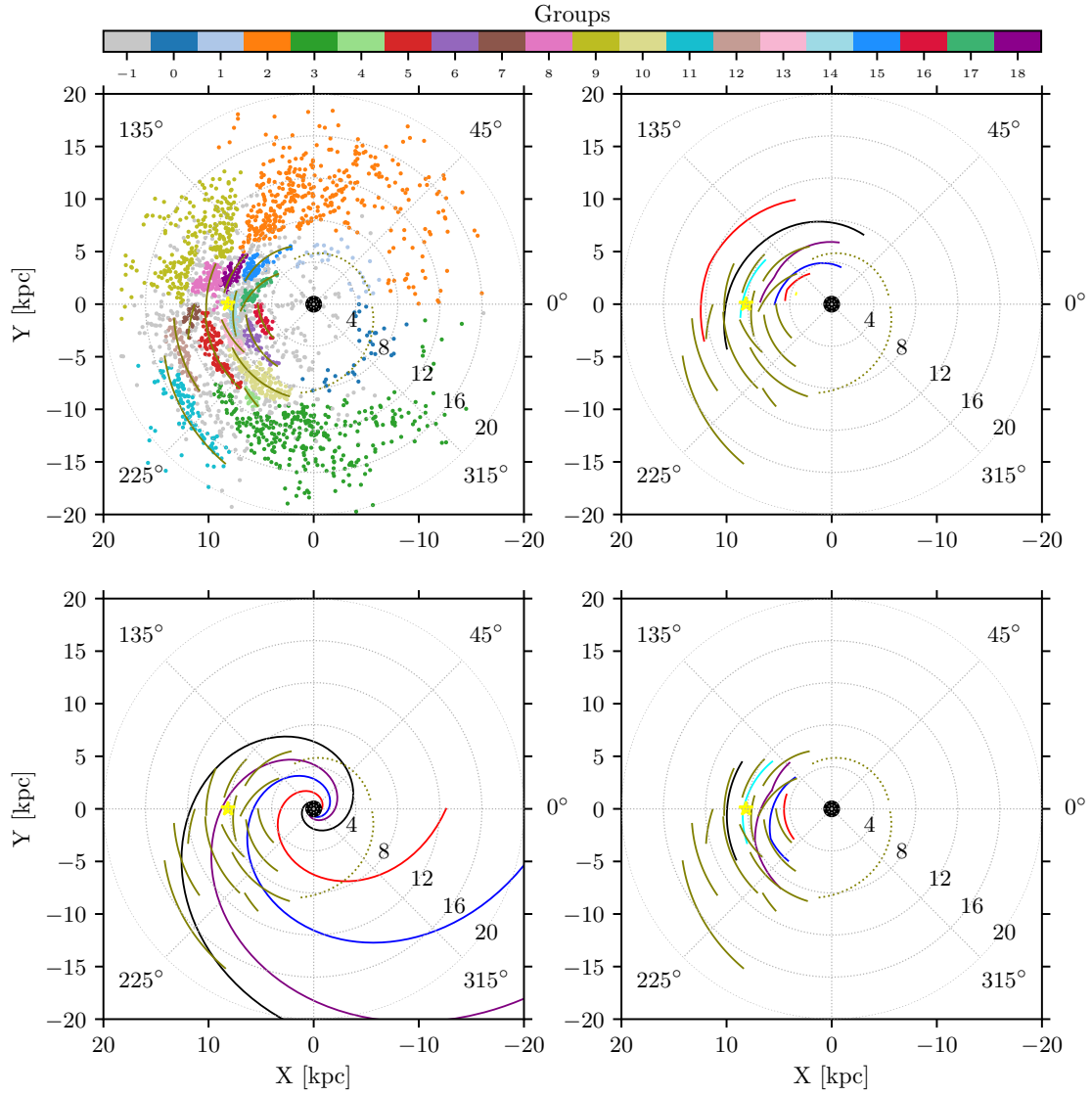
D.3. Cepheids younger than 180 Myr

Fig. D.3. Same as Fig. 8, but for Cepheids younger than 180 Myr.

Table D.3. Same as Table 4, but for Cepheids younger than 180 Myr.

Group	Slope (kpc.rad^{-1})	σ_{slope} (kpc.rad^{-1})	Intercept (kpc)	$\sigma_{intercept}$ (kpc)	Ref. angle (rad)	Ref. radius (kpc)	ln(Ref. radius) (kpc)	Pitch angle (rad)	Min. angle (rad)	Max. angle (rad)
[1]	-0.087	0.051	1.712	0.062	1.068	5.048	1.619	0.087	0.154	1.982
[16]	-0.088	0.058	1.944	0.201	3.486	5.141	1.637	0.088	3.145	3.827
[17]	0.570	0.041	0.100	0.121	2.835	5.561	1.716	-0.518	2.436	3.233
[15]	0.182	0.046	1.417	0.118	2.379	6.359	1.850	-0.180	1.950	2.807
[6]	0.053	0.032	1.700	0.119	3.743	6.675	1.898	-0.053	3.309	4.177
[0]	-0.256	0.071	3.312	0.391	5.424	6.845	1.923	0.251	4.566	6.282
[14]	0.128	0.018	1.627	0.059	3.253	7.717	2.043	-0.127	2.980	3.526
[18]	0.148	0.039	1.703	0.108	2.759	8.259	2.111	-0.147	2.511	3.007
[13]	-0.227	0.028	2.934	0.101	3.527	8.444	2.133	0.223	3.317	3.736
[10]	0.038	0.023	2.035	0.096	4.044	8.924	2.189	-0.038	3.639	4.450
[8]	0.038	0.064	2.205	0.186	2.977	10.157	2.318	-0.038	2.756	3.198
[5]	0.038	0.020	2.214	0.072	3.602	10.495	2.351	-0.038	3.220	3.984
[4]	0.357	0.124	0.900	0.511	4.120	10.705	2.371	-0.343	4.031	4.208
[7]	0.344	0.096	1.344	0.313	3.278	11.841	2.472	-0.331	3.149	3.406
[12]	0.037	0.020	2.472	0.072	3.517	13.492	2.602	-0.037	3.248	3.785
[11]	0.197	0.038	2.024	0.146	3.842	16.134	2.781	-0.195	3.480	4.205