



Universiteit
Leiden
The Netherlands

Stellar radio beacons for Galactic astrometry

Quiroga Nunez, L.H.

Citation

Quiroga Nunez, L. H. (2020, March 12). *Stellar radio beacons for Galactic astrometry*. Retrieved from <https://hdl.handle.net/1887/86289>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/86289>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden

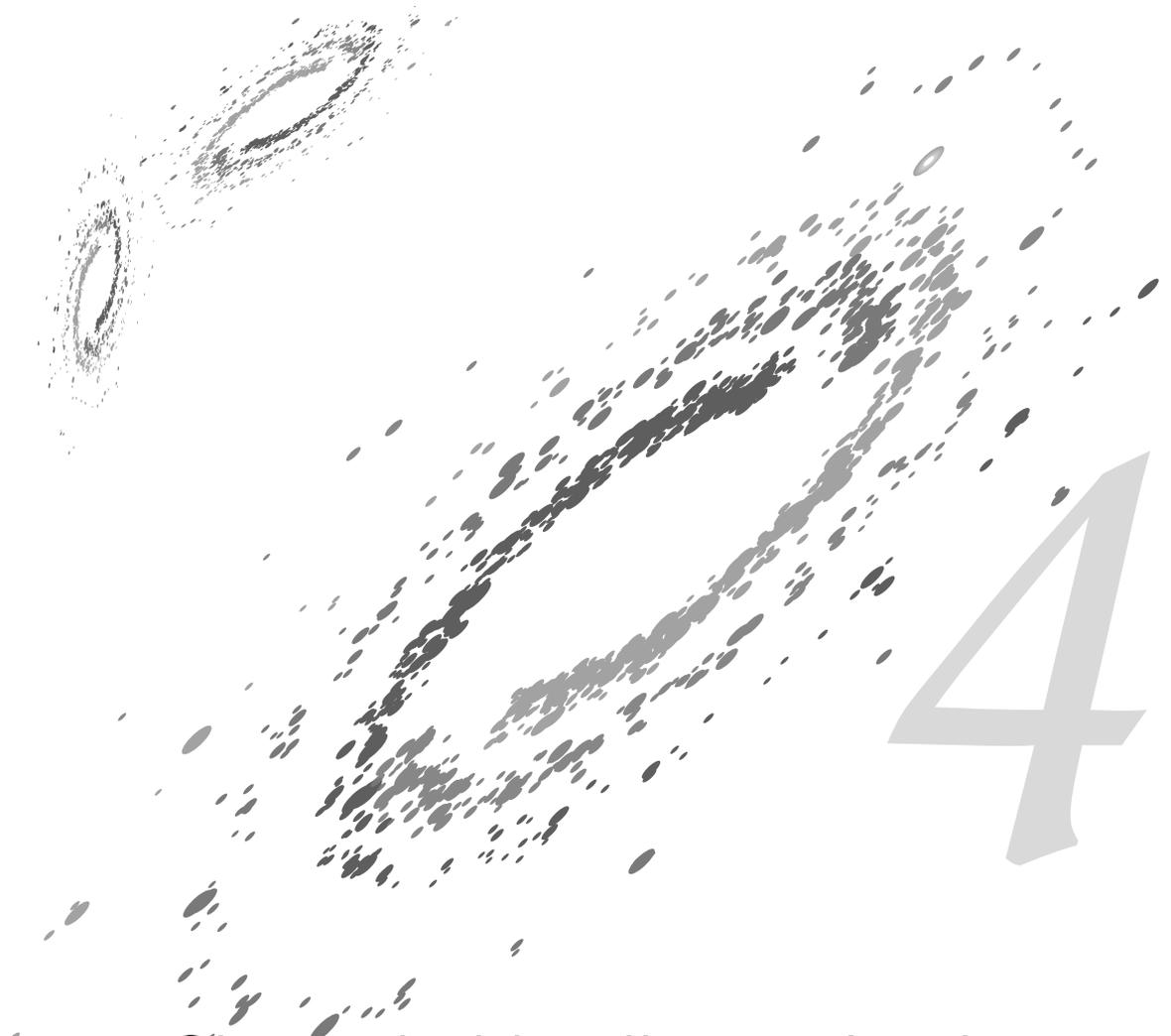


The handle <http://hdl.handle.net/1887/86289> holds various files of this Leiden University dissertation.

Author: Quiroga Nuñez, L.H.

Title: Stellar radio beacons for Galactic astrometry

Issue Date: 2020-03-12



Characterizing the evolved stellar population in the Galactic foreground

Quiroga-Núñez, L. H.; van Langevelde; Sjouwerman, L. O.; Pihlström, Y. M.; Brown, A. G. A.; Stroh, M.C.; H. J.; Rich, M. R. & Habing, H. J. *Characterizing the BAADE evolved stars in the Galactic foreground with Gaia 2020*, in preparation.

Abstract

Radio campaigns using maser stellar beacons have provided crucial information to characterize Galactic stellar populations. Currently, the Bulge Asymmetries and Dynamical Evolution (BAaDE) project is surveying infrared color-selected targets for SiO masers. This provides a sample of evolved stars that can be used to study the inner, optically obscured, Galaxy using line of sight velocities and possibly VLBI proper motions. In order to use the BAaDE sample for kinematic studies in the inner Galactic region, the stellar population should be characterized in terms of stellar evolutionary stage. In particular, it has been found before that BAaDE targets in the local disk and inner Galaxy originate from different stellar populations. The BAaDE targets, based on *MSX* colors, have been cross-matched with optical *Gaia*, infrared 2MASS, and preliminary radio (BAaDE, SiO maser) samples. By exploring the synergies of this cross-match together with *Gaia* parallaxes and extinction maps, the local ($d < 2$ kpc) BAaDE targets can be characterized. Additionally, the *Gaia* DR2 tables provide additional data on stellar classification and variability. We have defined a *BAaDe-Gaia* sample of 20,111 sources resulting from cross-matching BAaDE targets with IR and optical surveys. From this sample, a *local* sample of 2,060 stars with precise parallax measurements and within 2 kpc distance around the Sun was selected, for which absolute (bolometric) magnitudes are estimated. For 653 stars, *Gaia* also provides periods. These intrinsic properties are compared with samples of AGB stars from the literature finding that the original BAaDE color selection contains a contamination of low luminosity foreground objects. Objects with *Gaia* counterparts that are variable seem to be predominantly associated with AGB stars with moderate luminosity ($1,250 L_{\odot}$). The stars in the *local* sample for which there are SiO masers detected, are consistent with oxygen rich Mira stars with periods between 250 and 750 days.

4.1 Introduction

The characterization of Galactic stellar populations is a key ingredient to understand the structural (see e.g., Reid et al. 2019), chemical (see e.g., Ibata et al. 2017) and dynamical (see e.g., Martinez-Medina et al. 2017) evolution of the Milky Way, and indeed, its assembly through past merger events (e.g., Gómez et al. 2012). Typically, this is done by combining information on the spatial and kinematic distribution of a stellar population with an assessment of its age and origin (e.g., Mackereth et al. 2017). As the *Gaia* mission (Gaia-Collaboration et al. 2018; Lindegren et al. 2018) delivers more accurate, reliable data, it is revolutionizing our understanding of the assembly of the Galaxy. Many recent results demonstrate that mergers have been frequent over the history of the Milky Way and seem to occur also in the present era (Antoja et al. 2018; Helmi et al. 2018; Belokurov et al. 2019; Bland-Hawthorn et al. 2019).

Starting with the discovery of its HI spiral arms (Oort et al. 1958, and references therein), it has been clear that the Sun is a star in a spiral galaxy. In the inner region, the Milky Way seems to be dominated by a massive bar (e.g., Dwek et al. 1995) and an X-shaped structure (e.g., Wegg & Gerhard 2013), similar to what it is seen in extragalactic edge-on boxy bulges. As these are the most prominent dynamic features in the inner Galaxy, research of the kinematics and stellar populations that compromise the bar and the bulge is necessary to understand the morphology, structure and evolution of the Milky Way (Bland-Hawthorn & Gerhard 2016). Evolved stars, that are prominent in the mid-Infrared (IR), are possibly the best targets for such studies (Kunder et al. 2012). Indeed, the bar and bulge have been probed by counting IR stellar densities (Blitz & Spergel 1991; Babusiaux & Gilmore 2005), studying their metallicities and sometimes their variability, which for some stars can be used to obtain distance estimates.

Typically, these stars are too distant to measure proper motions or direct parallax distances from their stellar photosphere, as their Spectral Energy Distributions (SED) peak in IR, while their optical images are hidden behind circumstellar and interstellar dust. However, the most extreme of these evolved stars harbor circumstellar masers (Höfner & Olofsson 2018). Circumstellar masers are useful as they are bright beacons of a specific evolutionary stage in which evolved stars develop a thick circumstellar shell with specific molecular content and exceptional physical conditions. Moreover, the masers deliver accurate line of sight velocities through the Doppler effect. Finally, stellar maser emission reaches high brightness temperatures, allowing in principle Very Long Baseline Interferometry (VLBI) astrometry with micro-arcsecond accuracy (Reid & Honma 2014).

Previous surveys first focused on OH masers (Sevenster et al. 2001; Fish et al. 2006) and later targeted SiO masers with single dish telescopes (Messineo et al. 2018). When it was realized that the new capabilities at 7mm of the VLA and 3mm of ALMA offer efficient ways to study SiO masers, the Bulge Asymmetries and Dynamical Evolution project (BAaDE¹) was proposed. Using *MSX* infrared color selections, many thousands of SiO masers are found (Sjouwerman et al. 2017; Stroh et al. 2019). This sample may thus facilitate a detailed study the kinematics of the bulge, bar and inner Galaxy.

Since only very few SiO masers are known from young stars (Colom et al. 2015), those stars that show emission at 43 and/or 86 GHz are almost exclusively Asymptotic Giant Branch (AGB) stars. But stars of a very wide mass range are expected to spend time in this phase, as they become unstable towards the end of their lives. As a consequence, the ages of these star can vary considerably, ranging from 100 Myr to a fraction of the age of the universe (e.g. Salaris et al. 2014, and the references therein). Metallicity effects also affect the observables of the AGB population, as stars for which the envelope becomes low in oxygen may not easily produce sufficient SiO (e.g., Sande et al. 2018). Although the *Gaia* mission cannot

¹<http://www.phys.unm.edu/~baade/>

provide information on all of the BAaDE targets — and certainly not the majority of targets that sample the inner Galaxy — it can be used to characterize the stars in the BAaDE sample, particularly those in the local region ($d < 2$ kpc).

In this paper we cross-match the BAaDE sample with 2MASS and *Gaia* DR2. Because the BAaDE sample is based on *MSX*, it contains predominantly stars at low Galactic latitude. The cross section of the various surveys has infrared photometric information, as well as optical. Through the *Gaia* DR2, we can also evaluate other parameters as the parallax and proper motion, but also information derived from the survey such as variability and stellar classification. Finally, for a subsample there are already SiO detections. The objective of this work is to understand the nature of stars that enter the BAaDE survey. As we selected objects from their *IR* colors in *MSX* (with SiO maser emission detected for $\sim 70\%$, Trapp et al. 2018, Stroh et al. 2019 subm.), one can expect it to contain predominantly Long Period Variable (LPV) stars, likely Miras, with a modest circumstellar shell. But this sample can contain young stars, that are very luminous, or older, less massive stars that progress on the AGB track with lower luminosity. In order to address these issues, we present the cross-matches in Sect. 4.2, results for the various sub-samples are presented in Sect. 4.3. We discuss the characterization and distribution of the samples in Sect. 4.4, where we are in particular well positioned to comment on the nature of evolved stars in the foreground Galactic plane, for which we have *Gaia* counterparts with accurate distances.

Table 4.1: Number of sources obtained for different samples and cross-matches used in this research. See Sect. 4.2.1 and 4.2.2 for a detailed description of how the cross-matches were done and the filtering criteria.

			Sources / Cross-matches
BAaDE (MSX)			28,062
\cap	2MASS		25,809
	\cap	Gaia DR2	20,111
			SiO Observed
			4,154

Notes. The intersection symbol (\cap) indicates cross-match where the arrow indicates subsample. The “SiO Observed” refers to the subsample for which SiO maser flux densities and radial velocities from spectral lines are already available. Note that this observed SiO maser sample is still in preliminary stage.

4.2 Cross-match between radio, IR and optical observations

We have cross-matched the *MSX*-based BAaDE target sample with 2MASS and *Gaia* DR2 (see Table 4.1), using the *Gaia* data archive interface². This sample was defined as the *BAaDE-Gaia* sample (see Sect. 4.2.2) In the following subsections, we describe how this process was implemented, starting from the BAaDE target selection, followed by the cross-match criteria, the filters we applied to reduce false positives (including the extinction maps used), and finally arriving at the resulting sample of evolved stars in the foreground Galactic plane, which we define as the *local* sample (see Sect. 4.2.4 and Table 4.2).

²<http://gea.esac.esa.int/archive/>

Sample name	Description / Filters used	Sources/Cross-matches
<i>BAaDe-Gaia</i>	BAaDE (MSX) \cap 2MASS \cap Gaia DR2	20,111
	Variable \hookrightarrow	3,017
	SiO Observed \hookrightarrow	395
	$\sigma_\pi / \pi < 0.2$ \hookrightarrow	2,277
<i>Local</i>	< 2 kpc \hookrightarrow	2,061
	SiO Observed \hookrightarrow	147
<i>Variable</i>	Variable \hookrightarrow	908
<i>Periods</i>	Periods \hookrightarrow	653
<i>SiO maser</i>	SiO Observed \hookrightarrow	42

Table 4.2: Number of sources obtained for each subsample of the BAaDE-Gaia sample. Each row represents a filter used. See Sect. 4.2.3 and 4.2.4 for a detailed description of each filter. The definition of the BAaDE-Gaia and local samples is given in Sects. 4.2.2 and 4.2.4, respectively.

Notes. The arrow symbols indicate subsample. When two arrows are aligned in the same column, it means that the corresponding sample was split by two different criteria.

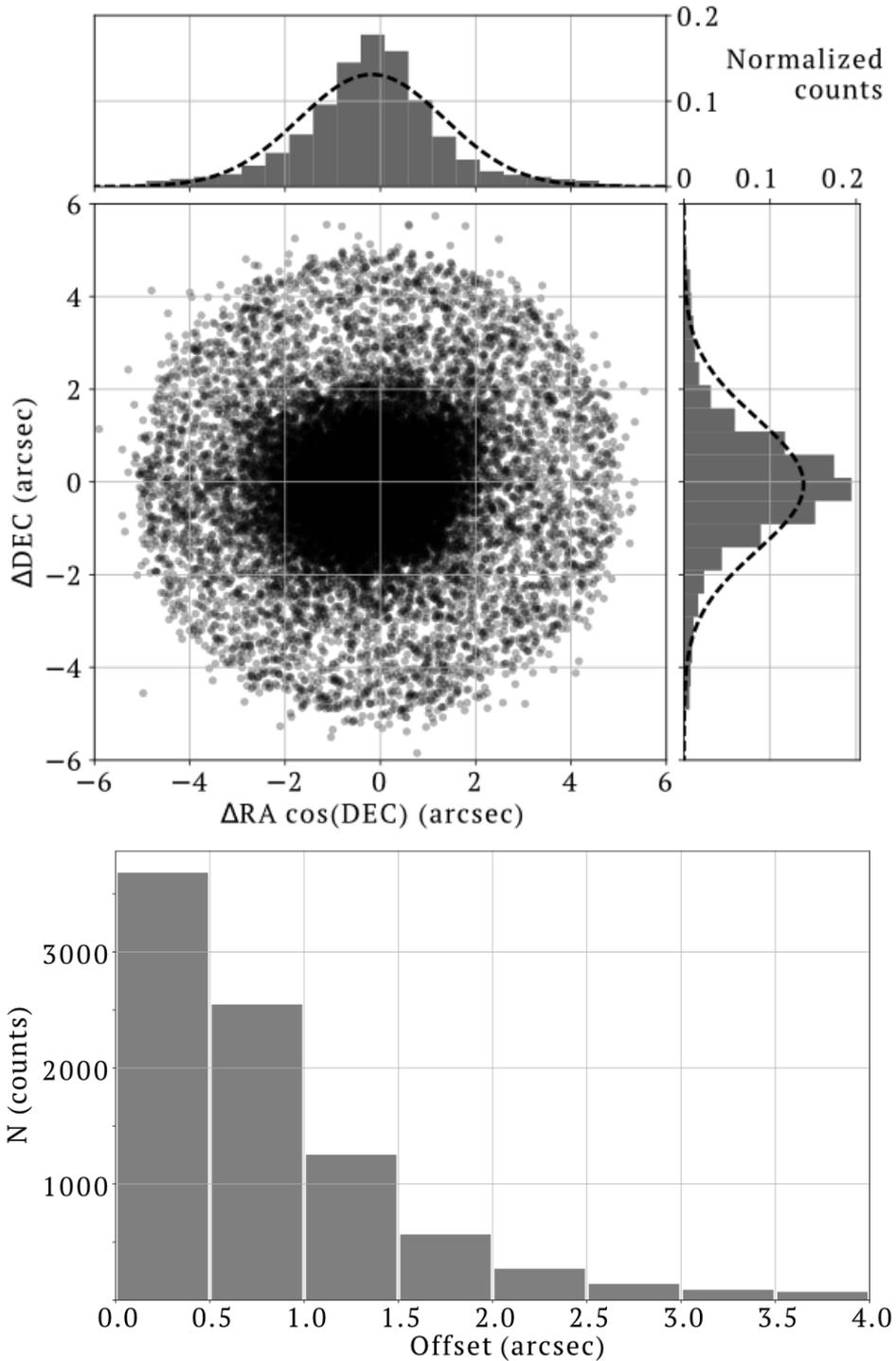


Figure 4.1: Upper panel: Distribution of the equatorial coordinate offsets between BAaDE targets and *Gaia* DR2 counterparts. Each offset component can be well-represented by a 1D Gaussian distribution. **Lower panel:** Histogram distribution of the total offset between BAaDE targets and *Gaia* DR2 counterparts. The histogram follows a Rayleigh distribution were sources with offsets of $>3''$ are likely false positives (see Sect. 4.2.2).

4.2.1 BAaDE target sample selection

The BAaDE target selection was based on IR photometry and designed to identify red giant stars with envelopes likely to harbor SiO maser emission. Starting from the IRAS color-color diagram (CCD), van der Veen & Habing (1998) studied dust and gas envelopes of AGB stars. They pointed out that circumstellar shell properties of AGB stars appear in a sequence in the IRAS CCD, suggesting an evolutionary track with increasing mass-loss rate. In the IRAS CCD, SiO maser stars are expected to be found in a specific color region, facilitating a selection based on the IRAS colors. However, the astrometric accuracy of IRAS is limited to one arc-minute for sources in the Galactic plane, prohibiting a large-scale survey. Later, Sjouwerman et al. (2009) were able to transform the IRAS CCD sequence to colors in the mid-IR, using *MSX* data. By doing this, the angular resolution was improved to $2''$ (Price 1995), and a new sample of AGB star candidates with mid-IR information was obtained. This way, 28,062 stellar targets were selected with the objective to sample the evolved stellar population in the Galactic plane, bar and bulge, mostly limited to $|b| < 5^\circ$. It is expected that one third of the BAaDE target sample lies in the Galactic bulge (Sjouwerman et al. 2017). These 28,062 targets are being followed up in order to detect SiO maser emission at 43.1 GHz ($J = 1 \rightarrow 0$ [$\nu = 1$]), 42.8 GHz ($J = 1 \rightarrow 0$ [$\nu = 2$]) with the VLA or 86 GHz ($J = 2 \rightarrow 1$ [$\nu = 1$]) with ALMA. So far, 20,600 candidates have been observed, of which 4,996 have already been analyzed (3,209 with the VLA and 1,787 with ALMA) and the scientific products are planned to be released publicly soon. The remaining sources are expected to be observed with ALMA in future cycles. Preliminary BAaDE results for 80 targets, comparing quasi-simultaneous observations at 43 and 86 GHz SiO maser lines have been already published by Stroh et al. (2018).

4.2.2 Cross-match description

In order to match the BAaDE targets with other surveys in position, we considered a conservative sky-projected circular area with $3''$ radius around the BAaDE targets. The motivation for a $3''$ maximum separation was based on the assumption that the distribution of deviations from the actual positions is dominated by the *MSX* data (as confirmed by Pihlström et al. 2018b) and has Gaussian distributions in both components ($\Delta\text{ff} \cos(\text{ffi})$, Δffi) with absolute mean values $< 0.2''$ and positional accuracy of $2''$ (see upper panel of Fig. 4.1). This 2D Gaussian distribution is represented as a the Rayleigh distribution. The lower panel of Fig. 4.1 shows the offset distribution, where sources at offsets above $\sim 3''$ are likely false positives. Note that the criterion we use here is more restrictive than the first cross-match done for a pilot of BAaDE sources and 2MASS (i.e., $5''$ in Trapp et al. 2018).

After defining a sky-projected circular area, we proceed with cross-matching the BAaDE target sample with 2MASS and *Gaia* DR2. Although the cross-match in principle can be done directly with *Gaia*, as it has typically lower positional offsets with respect to SiO masers positions (Pihlström et al. 2018b), we started instead by cross-matching BAaDE and 2MASS, motivated by two different arguments. First, BAaDE targets that display both mid-IR emission (*MSX*) and optical emission (*Gaia*) are not expected to be extinct at NIR (2MASS). Hence, by initially cross-matching with 2MASS, we are already avoiding some false positives (see Sect. 4.2.3). Second, the cross-match between 2MASS and *Gaia* was already established by Marrese et al. (2019), using a robust best neighbor algorithm, which found more than 90% overlap between both surveys.

By using the 2MASS survey, we have found 25,809 counterparts for the BAaDE target list. Next, after cross-matching with *Gaia* DR2, the sample was reduced to 20,111 cross-matches

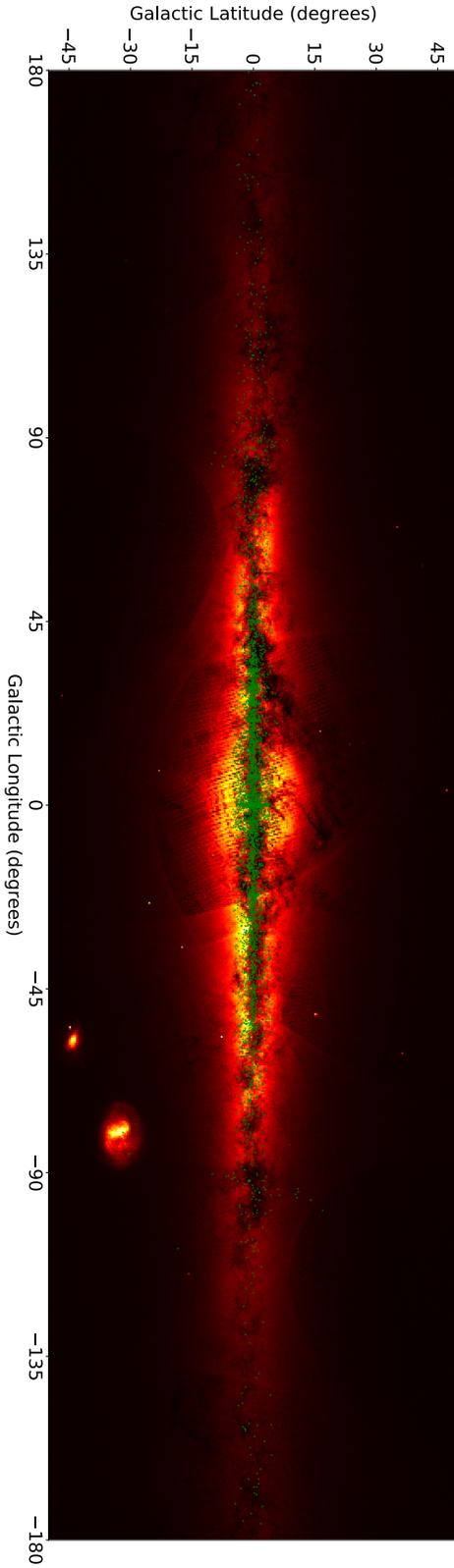


Figure 4.2: Galactic distribution of the BADE stellar targets without a Gaia counterpart (green points) overlaid on the first Gaia sky map. This sample accurately correlates with highly obscured regions in the optical regime. Credit: ESA/Gaia/DPAC.

(see Table 4.1), where all of them were found to be one-one correspondences. This last sample of 20,111 is called *BAaDE-Gaia* sample throughout the paper and thus includes 2MASS information. Notably, for 8,491 BAaDE targets, there were no *Gaia* counterparts, probably due to the fact that these targets lie behind considerable dust extinction at optical wavelengths. Figure 4.2 shows how the distribution of these “missing” sources indeed correlates with the dust obscured regions that *Gaia* could not penetrate.

Statistics of the cross-matches

Assuming a uniform distribution of sources in the bulge for the *Gaia* detections, as well as for BAaDE targets, one can calculate the number of sources that will give random matches at the given resolution of each survey. We estimated that the number of random matches should be less than 1,200; this is a small fraction of the 20,111 cross-matches that we have. Moreover, in this statistical estimate we have assumed that there is no optical extinction limiting the number of *Gaia* sources. Therefore, the actual number of chance matches will be much lower than 1,200 indicating that our sample has a modest contamination of sources with unrelated counterparts.

4.2.3 Refining the *BAaDE-Gaia* sample

As the objective is to characterize the BAaDE targets, we consider additional refinements of the cross-matches in order to identify contaminating sources. Several filters have been considered, which in turn have generated several subsamples from the *BAaDE-Gaia* sample of 20,111 sources. Below, we outline the criteria that have been applied. Table 4.2 summarizes the resulting subsamples, also indicating the number of sources that have already been observed by the BAaDE project with the *VLA* or *ALMA* and subsequently analyzed for SiO maser emission.

Parallax measurements

Obtaining distance estimates from noisy parallax measurements can be a complex issue (see e.g., Bailer-Jones 2015). Several tools are available to extract statistically robust distances from parallax measurements with limited accuracy — even from negative parallaxes — (see e.g. Bailer-Jones et al. 2018; Luri et al. 2018). However, such distance estimates strongly rely on robust expectations on the stellar properties for a target sample. In our case, the best approach would be to compute the parameters of a probability distribution specifically for AGB stars by maximizing a likelihood function, so that under an assumed statistical model the distance distribution for the observed evolved stellar data is the most probable. However, since (1) the aim of this research is to study foreground population of evolved stars and (2) precise extinction maps are limited to 2 kpc (see following subsection), we have found that 91% of the stars in the relevant sample have a $\sigma_\pi/\pi < 0.2$, which allows a direct and precise estimate of the distance without further considerations (Bailer-Jones 2015). Moreover, an analysis of the foreground sample can be considered an initial step for doing a full statistical analysis.

Extinction maps up to 2 kpc

Capitaniao et al. (2017) and Lallement et al. (2019) have produced local dust maps, based mainly on a regularized Bayesian inversion of individual color excess measurements using *Gaia* data. Additionally, the authors combined several tracers to confirm accurate extinction

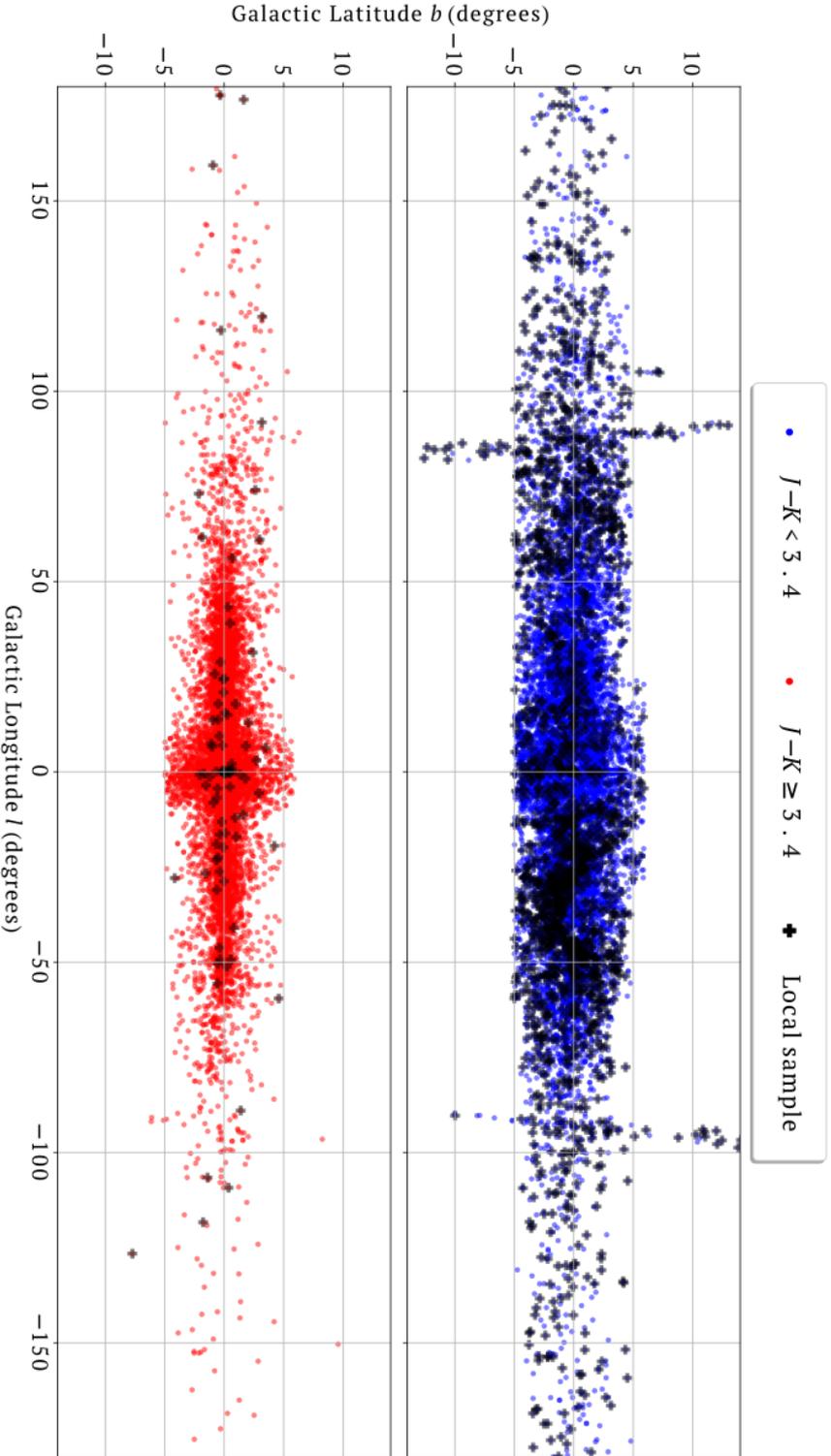


Figure 4.3: Galactic longitude-latitude diagram for the cross-matches obtained between BADE, 2MASS and Gaia defined as the BADE-Gaia sample. This sample was split in two populations (upper and lower panel) based on the mean 2MASS color ($J - K$) obtained, similar to what Trapp et al. (2018) have done to identify “cold” and “hot” kinematic populations within the BADE database. Black crosses represent the defined local sample, a subsample of the BADE-Gaia sample with precise parallax measurements at < 2 kpc distance (see Sect 4.2.4 for further details). The local sample is mainly made of foreground Galactic stars. The linear features observed at $l \sim 85^\circ$ and $l \sim -85^\circ$ for sources with $|b| > 8^\circ$ are part of MSX target list (and are also BADE targets) caused by the rotation of the spacecraft to map the entire Galactic plane (Price 1995).

maps and reddening estimates up to 2 kpc. This tool is extremely useful to estimate intrinsic luminosities for the stars in our sample, which is an important physical property that can be used to characterize the stellar population. Although for local AGB stars, which emit mostly in the (mid-) IR, the effects will be small, we do adopt these maps, and thus a distance limit of 2.0 kpc.

4.2.4 Resulting cross-match: *local* sample

Using the 20,111 cross-matched sources that we have found between BAaDE, 2MASS and Gaia DR2 (*BAaDE-Gaia* sample), we have applied the additional filters (previously described), leaving a sample of 2,061 stellar sources that we have defined as *local* sample. This sample contains BAaDE targets within 2 kpc distance around the Sun with precise distance estimates, IR and optical photometry and proper motions. From those, for 147 sources we have already obtained radio data, i.e., flux densities and radial velocities from SiO maser spectral lines at 43.1, 42.8, and/or 86 GHz. These samples uniquely enable us to study the late-type stars in the Galactic plane.

In addition, the *local* sample can be filtered by variability. For this, we have used the *Gaia* DR2 variability information contained in the *Gaia* table `vari_classifier_result`, and extract those objects that were flagged as variables of any kind, and named it as *variable* sample (908 sources). Next, we have refined the sample by extracting the sources with period estimates from the *Gaia* table `vari_long_period_variable`, and named it as *periods* sample (653 sources). Note that all the sources within the *local* sample contained in the table `vari_classifier_result` were classified by *Gaia* as Mira or semi-regular stars (`MIRA_SR`). Then, from the sources in the *periods* sample where we already have SiO maser emission confirmed, we defined the *SiO maser* sample (42 sources). All of the characteristics previously described helped us to generate subsamples of the *local* sample, as shown in Table 4.2. Finally, it should be noted that there are two effects that play a role when distances to individual AGB stars are estimated. First, the strong colour variations on the stellar photosphere (see e.g., Lindegren et al. 2018; van Langevelde et al. 2019), and second, the photocenter movements caused by large atmospheres with convective motions (Chiavassa et al. 2018). We have checked and added the *Gaia* `astrometric_excess_noise` uncertainty when discussing individual objects.

4.3 Results

In this section we present the main features of the different samples that resulted from the cross-matching. We start with the *BAaDE-Gaia* sample of 20,111 targets described in Sect. 4.3.1. Next, in Sect. 4.3.2, we characterise the *local* sample in terms of IR photometry, variability and Galactic distribution.

4.3.1 Features of the *BAaDE-Gaia* sample

Since the *BAaDE-Gaia* sample was obtained through 2MASS, the mean value of near-IR color ($J - K$) can be used to split the sample in two equal sized subsamples: i.e., $(J - K) < 3.4$ for the bluer stars and $(J - K) \geq 3.4$ for the redder stars. More extreme AGB stars (more luminous and with thicker shells) are to expected to have steeper slope in their SEDs at IR wavelengths, resulting in increasingly redder IR colors. Figure 4.3 shows the subsamples of red and blue stars in a Galactic latitude-longitude diagram. Red stars seem to better trace

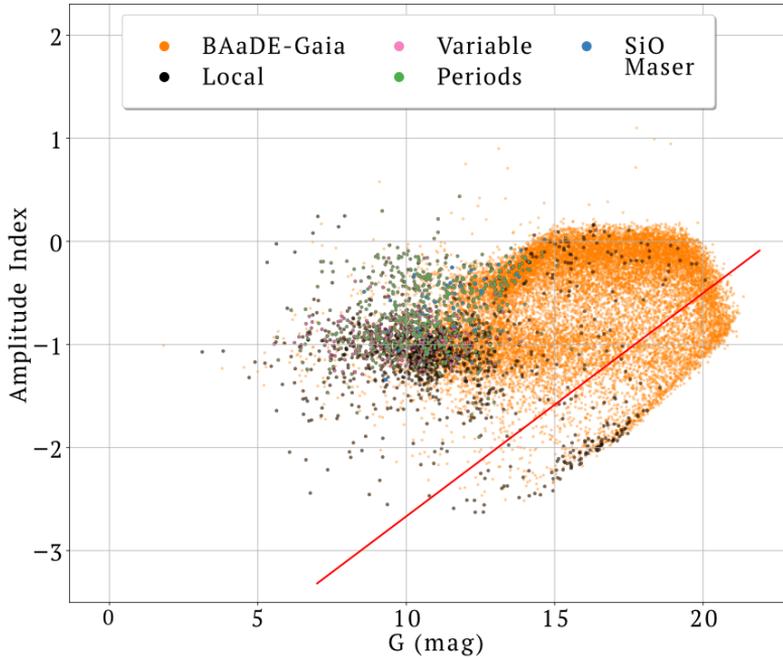


Figure 4.4: Amplitude-magnitude diagram suggested by Belokurov et al. (2017) to distinguish variable stars where larger amplitudes are likely associated with pulsating AGB stars. The green points represent the BAaDE-Gaia sample ($BAaDE \cap 2MASS \cap Gaia\ DR2$), whereas black sources represent the variable stars within the local sample (i.e., precise distance estimates at <2 kpc). The various subsamples were derived from the local sample following Table 4.2. The solid red line represents a threshold above which stars are considered variable Belokurov et al. (2017).

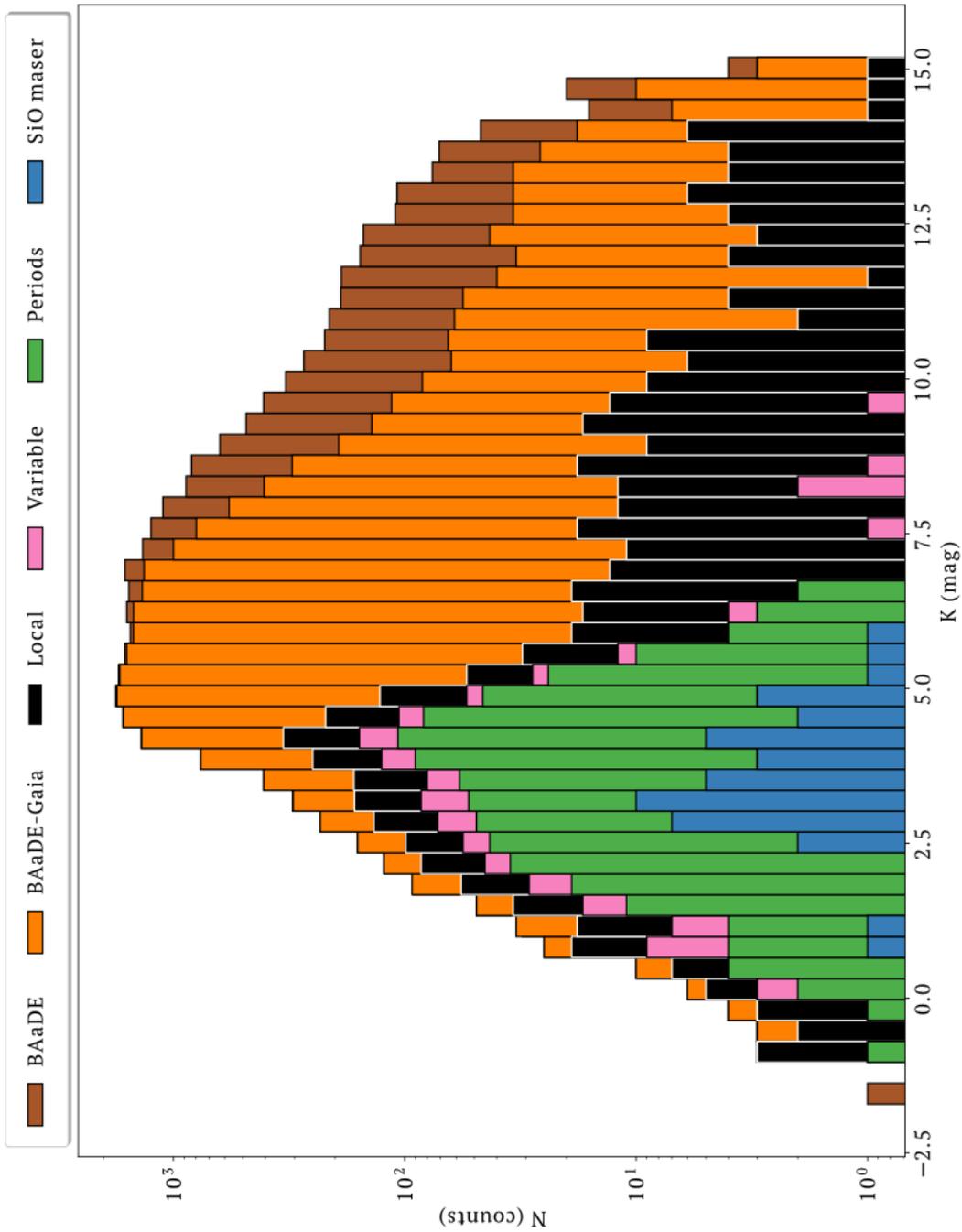


Figure 4.5: Cumulative histogram comparison for the distribution of K-magnitude observed by 2MASS for different samples. Each distribution results from a drawn bigger sample as shown in Tables 4.1 and 4.2. The figure shows that as more filters are applied, the apparent K-magnitude range is narrower.

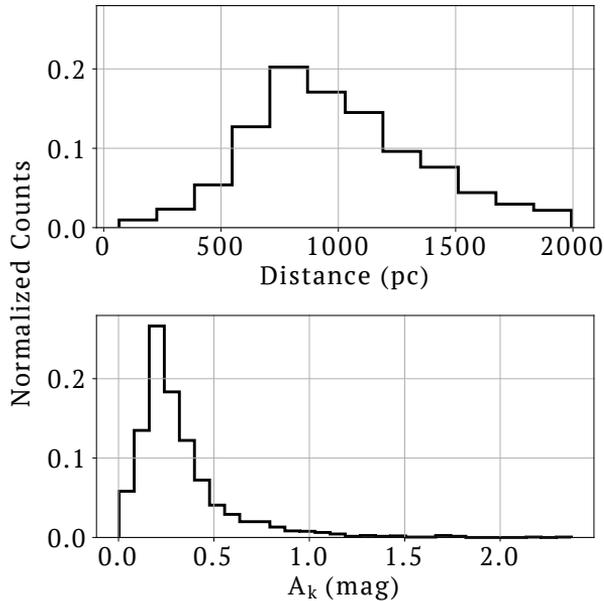


Figure 4.6: *Upper panel:* Distribution of distance to the Sun for the *local* sample. *Lower panel:* Extinction in the *K*-band obtained from the optical extinction maps developed by Capitanio et al. (2017) and Lallement et al. (2019), and converted to IR *K*-band following Messineo (2004).

the inner part of the Galaxy (Galactic bulge and plane) while bluer stars seem to dominate the foreground population. Indeed, the figure also shows that the sources of the *local* sample are mainly stars that are bluer (in the context of the BAaDE selection). We confirm that by splitting sample using IR photometry, two populations can be traced. This was also observed by Trapp et al. (2018), who made the split using *K* magnitudes, and labeled the two a kinematic populations “cold” (the bluer, less red, stars in the Galactic disk) and “hot” (the redder stars in the bulge/bar).

Another property that can be investigated for the *Gaia*-BAaDE sample is variability. Although the *Gaia* DR2 has variability information for a considerable number of stars (Mowlavi et al. 2018), Belokurov et al. (2017) has shown that —already in *Gaia* DR1— flux uncertainties quoted in the *Gaia* catalogue reflect the dispersion of the *G*-band flux measurements, which will thus lead to apparently larger uncertainties for variable stars. They have defined an amplitude variation over error, which we refer as amplitude index, using the flux in the optical *G*-band as $\log_{10}(\sqrt{N_{obs}} \frac{\alpha_g}{g})$, where N_{obs} is the number of observations. Using this quantity, Belokurov et al. (2017) calculated the amplitude for different stellar populations in the Large Magellanic Cloud (LMC) and Small Magellanic Cloud (SMC), finding that Mira variables are located in the upper region of Fig. 4.4, where the amplitude index is > -1.0 . Figure 4.4 shows an amplitude-magnitude plot for the *Gaia*-BAaDE sample, where stars with amplitudes larger than -1 in this diagram are likely pulsating stars. At the same time, we also have described in Sect. 4.2.4 that in *Gaia* DR2, sources were classified as “Variable”, following the criteria described in Mowlavi et al. (2018), and defined the *variable* sample. Figure 4.4 shows that these variable stars coincide with larger amplitude values as expected, confirming that indeed the IR classification made by the BAaDE project correlates with variable stars. However, this

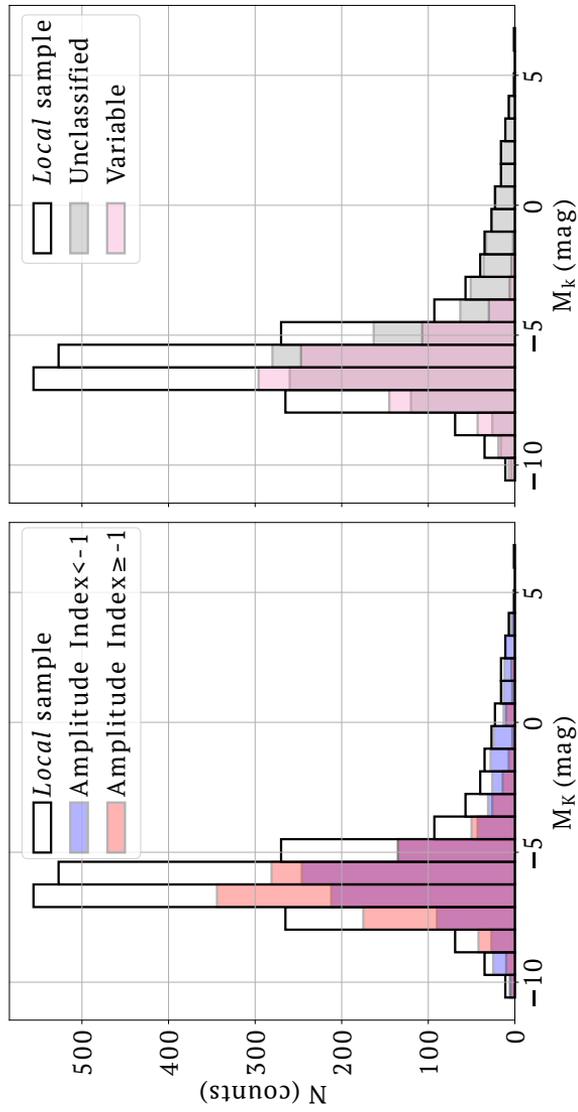


Figure 4.7: Absolute K -magnitude distribution for the local sample. In each panel, the distribution was split by variability criteria, i.e., using their Amplitude (Belokurov et al. 2017) and the *Gaia* variability classification, in left and right panel, respectively.

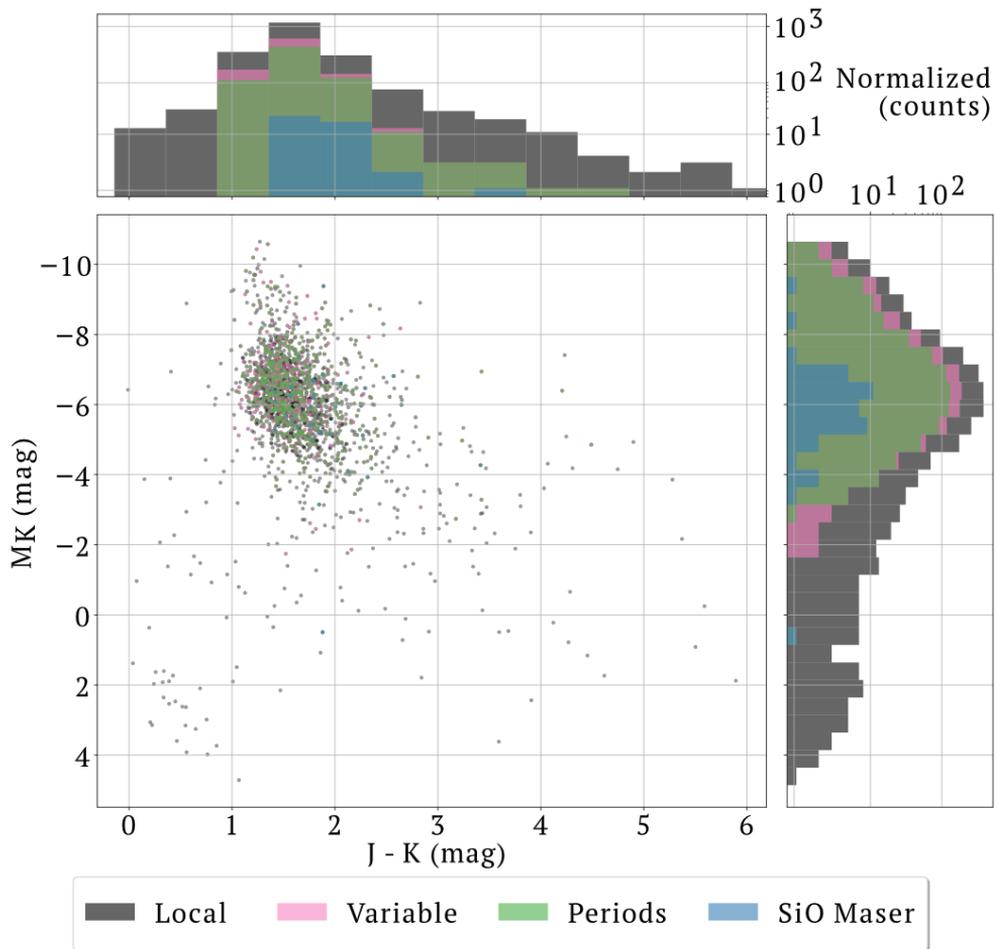


Figure 4.8: Magnitude (K -band) color diagram for the local sample (black) and the different subsamples obtained (Table 4.2). The distribution for each component is also shown.

qualification is restricted to stars that are bright in the G band.

4.3.2 The *local* sample

We study the Galactic foreground sample of BAaDE targets in terms of IR photometry, variability and Galactic distribution.

Infrared photometry

AGB stars can be identified by their IR colors, given that their SEDs peak at these wavelengths (see e.g., van der Veen & Habing 1998). In particular, after the 2MASS data release (Skrutskie et al. 2006), K measurements have been widely used to characterize these populations (Whitelock et al. 2008; Messineo et al. 2018). Figure 4.5 shows the distribution of the apparent K magnitude obtained from 2MASS for the entire cross-matched sample, with the different sub-samples related in Table 4.2. We note that by filtering the foreground sample with *Gaia* counterparts (the *local* sample), we are selecting brighter stars in the K -band.

Next, using the optical extinction maps described in Sect. 4.2.3, we can obtain the extinction and reddening estimates at K -band by assuming $A_V/A_K = (A_V/2.12\ \mu\text{m})^{-1.9}$ (Messineo 2004). As one could expect for the region around the Sun, the IR extinction estimates at the infrared K -band for the filtered sample are usually lower than 0.5 mag (see lower panel of Fig. 4.6). Finally, as we have precise distance estimates for the *local* sample (see upper panel of Fig. 4.6), we are able to estimate the absolute K magnitude distribution (M_K). Figure 4.7 shows this distribution where we have split the sample in terms of variability, as will be explained in the following subsection. Moreover, using the absolute K magnitudes together with *MSX* and 2MASS colors, we have constructed a Magnitude-color diagram (Fig. 4.8) for the *local* sample. This diagram confirms that the population of variable stars are restricted to well-defined ranges of mid-IR colors, as it was pointed out by van der Veen & Habing (1998); Sjouwerman et al. (2009); Lewis et al. (2018).

Variability

In Sect. 4.2.4, we have described the tables from *Gaia* DR2 that yield the variability classification that can be used for the *local* sample. And we noted that also the ratio between the flux error and mean in G magnitude can be used to identify pulsating stars when the amplitude index > -1 (Sect. 4.3.1). We have considered both methods, in particular in relation to the K -band apparent and absolute magnitude distributions. Figures 4.4 and 4.7 show these distributions split according to both variability criteria. Although both methods seem to produce similar results, the variability criterion from the *Gaia* DR2 tables, achieves narrower ranges of absolute magnitudes (particularly for less luminous objects). In other words, the amplitude estimator based on the G variance can presumably also pick up variability from objects that are not classified as variables in the *Gaia* DR2.

Galactic spatial distribution

Under the assumption that distances can be directly derived from *Gaia* parallax measurements (Sect. 4.2.3), the projected spatial distribution of the *local* sample can be generated (Fig. 4.9). The Galactic distribution of the *local* sample displays a number of distinct features:

1. The number of sources in the vicinity of the Sun (<0.5 kpc) is considerably lower than further out. This is to be expected as the BAaDE targets only sample latitudes of $|b| < 5^\circ$,

therefore the volume sampled increases with distance away from the Sun, until it covers the height of the galactic disk (scale height of ~ 300 pc, Jackson et al. 2002).

2. The inter-arm region between the Local arm and the Carina-Sagittarius arm seem to be the most populated region. The location of the main spiral arm is anti-correlated with the occurrence of stellar sources, which we speculate is related to regions with higher extinction.
3. At a Galactic longitude of $\sim 45^\circ$, there is a region depleted, which is possibly caused by the clouds of high optical depth that block this entire line of sight (Capitani et al. 2017).
4. The Outer Galactic region contains a lower number of sources from the BAaDE target selection. This is expected as the stellar density for AGB stars decreases as SiO maser detection.
5. In Fig. 4.9, there is no notable difference in the spatial distribution of variable stars compared to the unclassified (non-variable) sources according with *Gaia* DR2.

4.4 Discussion

4.4.1 Absolute magnitudes for the foreground Mira population

AGB stellar populations have been characterized by IR absolute magnitudes, because they radiate predominantly at these wavelengths (Vassiliadis & Wood 1993). Such IR observations allows one to study large populations, given the low extinction at these wavelengths, but unbiased distance estimates can be hard to obtain.

Several studies have been carried out to estimate IR absolute magnitudes of the AGB populations in the LMC and SMC, where the distance to the stellar system is known, and therefore, the distance modulus (and presumably also the IR extinction) can be assumed the same for each object. Figure 4.7 shows the distribution of absolute magnitude for the *local* sample with various indicators for variability. Although the magnitude values found correspond to those found at LMC and SMC, our distributions are much broader in terms of absolute magnitude range if we consider the entire *local* sample. This can be partly explained, as the current sample is mostly based on a single 2MASS observation and include the effects of large amplitude variability. However, when we filter the sample by a variability qualification (as shown in both panels of Fig. 4.7), the low luminosity tail (in M_K) is cut out. In this sense, the *Gaia* classification as Mira or Semi-Regular (SR) variable seems to narrow the distribution more, with a corresponding a peak at $M_K = -6.4 \pm 1.2$ mag. Preliminary results of the SiO masers detected by the BAaDE project also suggest that the variability classification (together with the IR color selection that generated the BAaDE target list) is selecting stellar objects in a specific M_K range between -7.2 and -3.6 with a peak at -6.5 . Following a similar discussion in Mowlavi et al. (2018), we argue that the low luminosity tail in Fig. 4.7 and also Fig. 4.5 is due to contamination with Young Stellar Objects (YSOs) that can also peak in the infrared, but do not show the same variability (Lewis, et al., in prep.).

4.4.2 Bolometric magnitudes for the foreground Mira population

The bolometric luminosity is a fundamental property useful for classifying stellar populations and evolutionary stages (Srinivasan et al. 2009), since they measure the intrinsic stellar power.

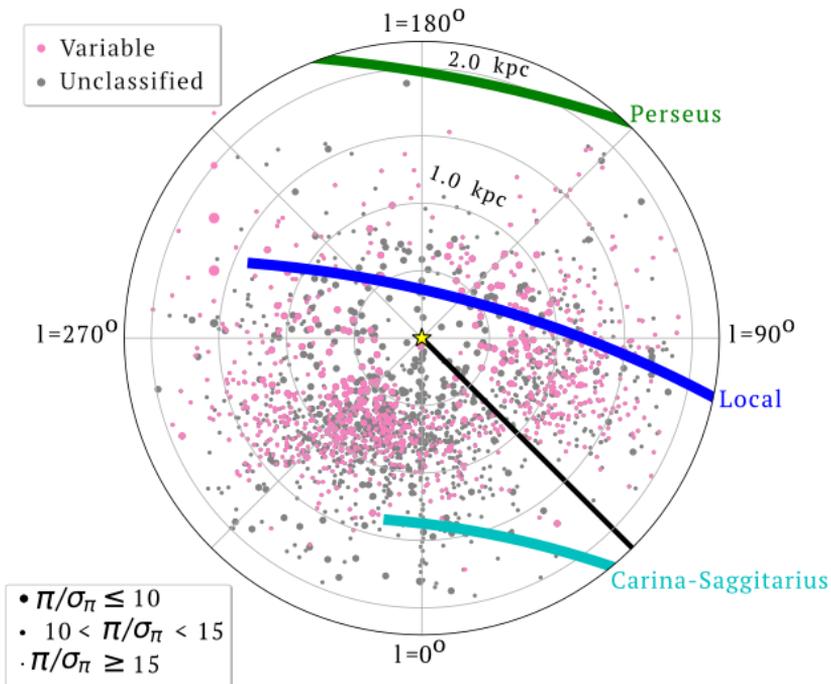


Figure 4.9: Foreground Galactic distribution of the local sample. The sample was split “Variable” and Unclassified according with *Gaia* DR2 (Mowlavi et al. 2018). The size of the marker is a measure of the relative parallax uncertainty, and therefore, distance uncertainty for each source. The spiral arms location and width (with of the lines) are based on Reid et al. (2014). The black line represent Galactic longitude of 45° , where a high extinction line of sight is expected (see 4.3.2).

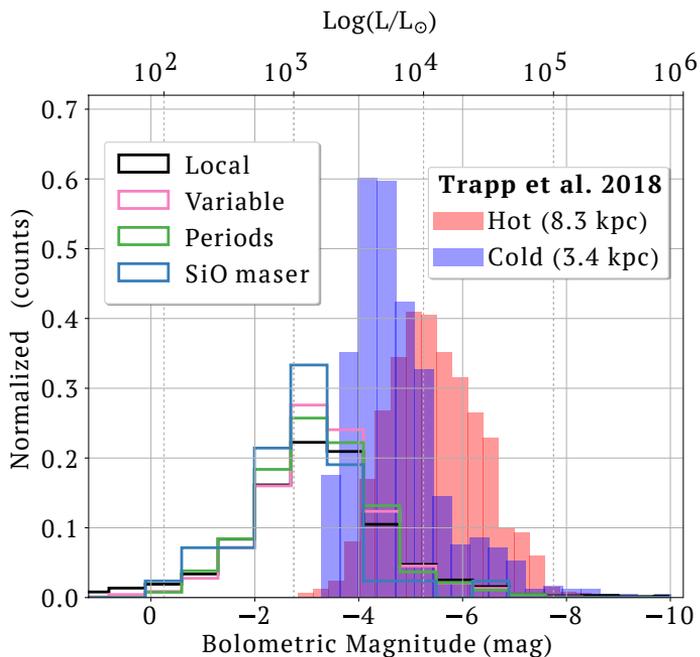


Figure 4.10: Luminosity and bolometric magnitude distributions for the *local* sample and the subsamples resulting from different filters (Table 4.2). These distributions were obtained by applying the BC proposed by Messineo et al. (2018) to the absolute magnitudes in the *K*-band. The absolute *K*-magnitudes were estimated from 2MASS *K*-band, Gaia parallaxes and extinction maps from Capitanio et al. (2017) and Lallement et al. (2019). The bolometric distribution estimated by Trapp et al. (2018) for the “hot” and “cold” kinematic populations are shown as filled histograms. Note that the kinematic cold population proposed by Trapp et al. (2018) is made up of stars in the disks and not in the bulge, therefore, similarities with respect to the *local* sample defined in this work are expected.

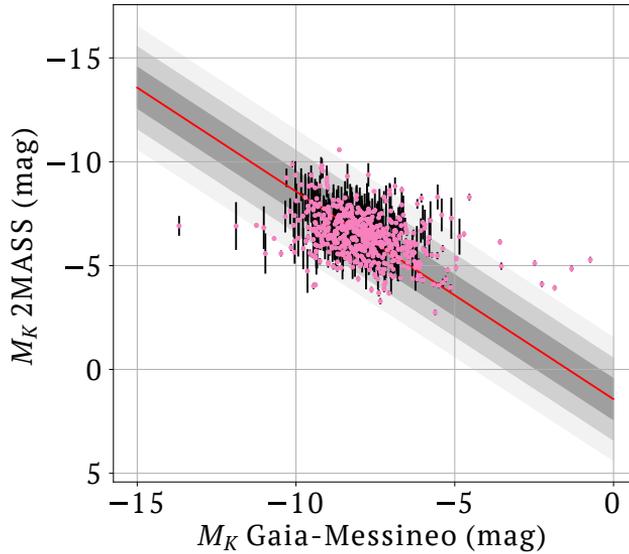


Figure 4.11: Comparison between the absolute K -magnitudes estimated from 2MASS data, *Gaia* parallaxes and extinction maps with respect to the absolute bolometric estimates reported for variable stars in *Gaia* DR2 that were transformed to absolute K -magnitude using the BC_K in Messineo et al. (2018). The red line describes the linear fitting that was forced to have slope of 1. The grey layers contains 1, 2, and 3 σ deviation from the linear fitting.

Although its definition is straightforward formulated as the total integrated power overall frequencies, in practice complete photometric measurements that allow a direct bolometric luminosity estimates are hardly ever available. Therefore, under various assumptions a limited set of photometric measurements, preferably near the peak of the SED, can be used to apply a bolometric correction (BC) in order to determine the integrated stellar luminosity. In particular, for AGB stars IR absolute magnitudes are converted to bolometric luminosities using a bolometric correction, which is usually parameterized using IR colors (see e.g., Messineo et al. 2018; Whitelock et al. 2008; Lebzelter et al. 2019).

Trapp et al. (2018) have estimated the bolometric magnitude for a sub-set of the BAaDE sample. They considered a kinematically “cold” population disk stars, which is similar to what is defined here as the foreground population or *local* sample. In their analysis, they have assumed a common distance of 3.8 kpc for this population and have applied a BC_K based on Messineo (2004). In order to compare the *local* sample with their kinematically “cold” population, we have applied the same BC_K , but not before confirming that other proposed BC_K s for AGB samples produced similar results (Whitelock 2003; Srinivasan et al. 2009). Figure 4.10 shows these bolometric distributions. The offset between the two distributions is likely caused by the distance assumption made by Trapp et al. (2018), which is equivalent to ~ 3.5 mag when taking the average distance of the *local* sample (767 pc, upper panel of Fig. 4.6).

We did our bolometric magnitude estimates from M_K , using 2MASS IR photometry, *Gaia* parallaxes and extinction maps. The *Gaia* data also provide bolometric corrections based on G magnitudes, allowing us to derive estimates of the bolometric magnitudes in particular for the stars classified as variables. But, one would assume that the 2MASS bolometric corrections are more robust for the IR selected stars in our sample. To compare our results with the

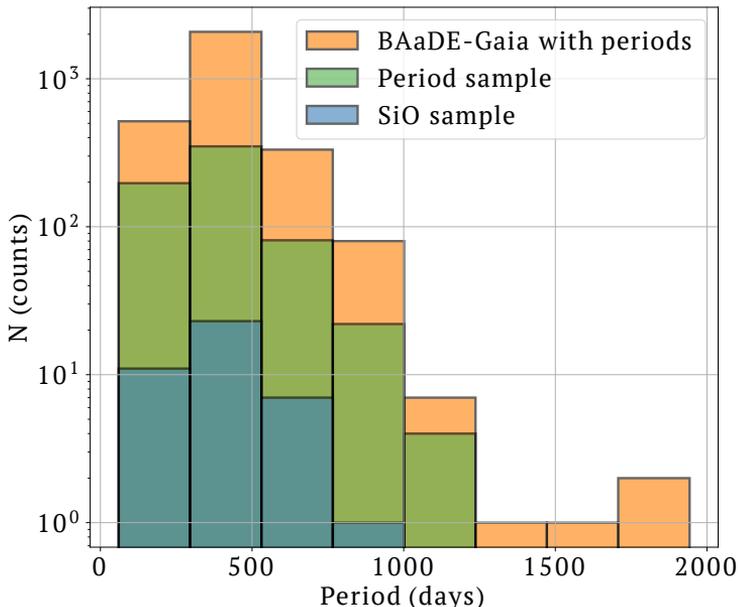


Figure 4.12: Histogram of periods obtained for variable stars in *Gaia* DR2 for the different samples used in this research (Table 4.2).

estimated that *Gaia* provides, we have compared the M_K that we have obtained in 4.3.2 with estimates of M_K obtained by working back from the *Gaia* bolometric magnitudes using the BC_K from Messineo (2004). The result in Fig. 4.11 shows a substantial difference of -1.6 mag, indicating that the bolometric corrections in the *Gaia* DR2 are overestimating the total luminosity estimated by *Gaia* DR2 (Andrae et al. 2018) of our very red objects. We continue to use only the M_K based using 2MASS IR photometry, *Gaia* parallaxes and extinction maps.

In Fig. 4.10, we also present the luminosity distribution for the *local* sample. It shows that our sample is made up of giant stars with a luminosity range that is consistent with AGB stars, mostly Mira variables (Srinivasan et al. 2009). In particular the sub-set of stars that have SiO masers has a similar luminosity distribution. Compared to previous studies of stars in the LMC, SMC or Galactic bulge (where fixed distances have been assumed), we have found less luminous objects. This is of course expected in our selection that was based on a combination of IR detections, optical *Gaia* counterparts, extinction maps and distance selection. The typical luminosity for the *local* sample was estimated in $1,250 L_{\odot}$.

4.4.3 Mira distribution around the Sun

The Galactic distribution of AGB stars has been studied extensively using IRAS, WISE, 2MASS and MSX data (Jackson et al. 2002; Habing 1996; Lian et al. 2014; Messineo et al. 2018; Sjouwerman et al. 2009). Generally, it has been found that AGB stars are tracing the relaxed stellar population of the Galactic (thick) disk and bulge. In particular, Jackson et al. (2002) found a density distribution based on revised IR photometric data from IRAS that they called universal, implying that there are no statistically significant differences in the spatial distribution of AGB stars based on IR colors. Adopting their radial scale length of 1.6 kpc (outside of $R > 5$ kpc) and scale height of 300 pc, we consider Fig. 4.9. We attribute the

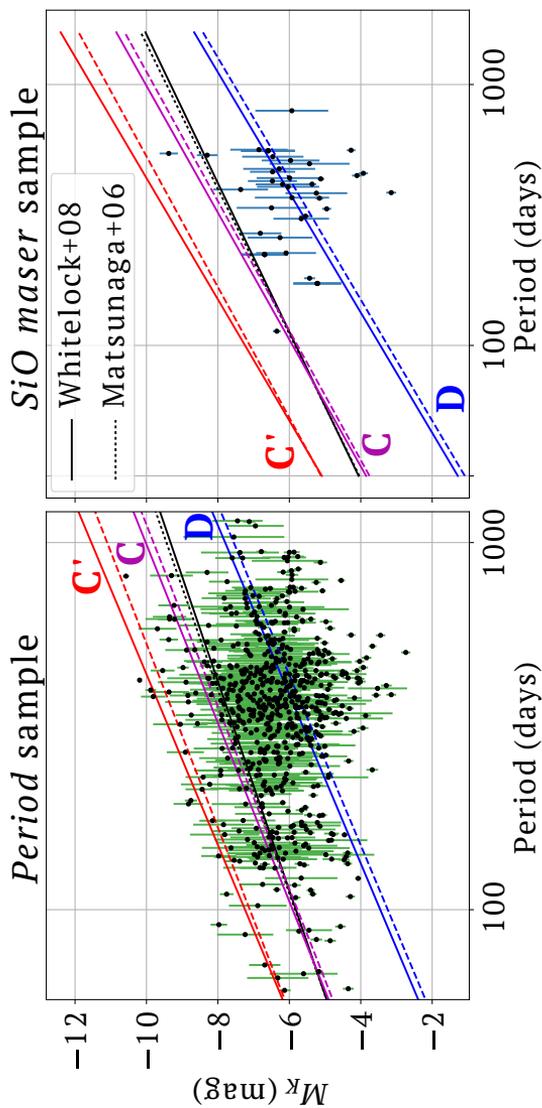


Figure 4.13: Period-Luminosity relations found for the variable stars with in the local sample (**left panel**) and those with SiO maser emission detected (**right panel**). The sequences marked as C, C' and D represent different known variability sequences, associated with distinct pulsation modes that have been derived for LMC based on *Gaia* data (Lebzelter et al. 2019). Note that the sequences were corrected for a distance modulus of 18.49 mag. Period-Luminosity relations reported for LMC using other surveys are also shown (Whitelock et al. 2008; Matsunaga & Team 2006).

depletion of targets < 500 pc around the Sun to the fact that the *MSX* catalogue, on which our sample is based, is mostly limited to $|b| < 5^\circ$. The corresponding scale height is estimated to be 100 pc, equally for all objects in the *local* sample, as well as for identified Mira variables. This seems to suggest that the scale height for our BAaDE targets is a bit lower than that estimated for previous AGB stars, but it requires detailed simulations to confirm that.

From Fig. 4.9, we note that contrary to a homogeneous distribution, there is an anti-correlation with the location of spiral arms as defined by Reid & Honma (2014). It appears that, the *local* sample is severely affected by interstellar extinction in the *Gaia* band. But, there is still a radial gradient detectable with more targets towards the center than observed in the anti-center direction. This can possibly arise, because the *MSX* criteria defined by Sjouwerman et al. (2009) were optimised for detecting SiO masers, O-rich AGB stars. It has been established that outside of the solar circle the AGB population contains a higher fraction of carbon rich stars (Lian et al. 2014; Groenewegen & Sloan 2018). This will become clearer when the SiO detections are discussed in Sect. 4.4.5.

Finally, we point out that in our sample there is one variable star standing out for its proximity. The source *IRAS* 17375-3434 is part of the *local* sample, it has been classified as a Mira by *Gaia* with a period of 373.5 ± 47.5 days and a distance of 65.5 ± 3.9 pc. Although we have confirmed SiO maser emission at 42 GHz coming from that region ($V_{\text{LSR}} = -15.5 \text{ km s}^{-1}$), its luminosity was estimated in $2.6 L_\odot$, which is unrealistically low for an evolved star. Based on DSS images of the region, we confirm that this star is located in a very crowded region in the Galactic plane, suggesting that it is likely a false positive.

4.4.4 Period-Luminosity relations

Figure 4.12 displays the distribution of periods available from *Gaia* for sources in the *BAaDE-Gaia* sample, the *local* sub-sample and for those stars bearing SiO masers (*SiO maser* sample). It can be noted that the whole sample contains LPV stars with a wide range of periods, but that SiO masers are mostly restricted to stars with periods of $\lesssim 750$ days, presumably Mira variables.

It has been established that Period-Luminosity (P-L) relations are a very powerful tool to distinguish AGB stars of different nature (Wood et al. 1999; Ita et al. 2004; Lebzelter et al. 2019). By recognising that Mira variables pulsate dominantly in the fundamental mode, they can be promising candidates for distance determinations of remote galaxies, using empirical relations based on the LMC and SMC (Whitelock et al. 2008). With 2MASS *K*-magnitudes, *Gaia* DR2 parallaxes, extinction maps and periods for a sub-sample of the *local* sample (*period* sample), we are able to make a comparison of the BAaDE targets with previously studied variable stars.

In Fig. 4.13, we present the P-L relation for those BAaDE stars in the *local* sample with measured *Gaia* periods (*period* sample), where there is a spread in the magnitude, resulting from uncertainties in apparent *K* magnitude, extinction and distance (indicated by the error bars) and infrared variability (not indicated). We present a separate panel for those stars that have confirmed SiO maser emission (*SiO maser* sample). A comparison is made of the P-L distribution with known variability sequences, associated with distinct pulsation modes, that have been derived from *Gaia* DR2 data for LMC (and SMC) populations as discussed by Lebzelter et al. (2019). These sequences have been transformed to M_K , using the LMC distance modulus in that work (18.49 mag). Moreover, the established P-L relations for Miras from Whitelock et al. (2008); Matsunaga & Team (2006) in the LMC and SMC are added.

In comparison with the analysis by Lebzelter et al. (2019), it is possible to interpret Fig. 4.13, where it should be noted that these results were obtained for the stellar popula-

tions in the Milky Way satellites, i.e. a different metallicity. At short periods one can identify stars associated with sequence *C*, while at the most extreme long periods most stars lie closer to sequence *D*. Supposedly both these sequences are being traced by low mass, oxygen rich Miras. At the intermediate periods, where there is the highest density of objects, there is no clear distinction between the two sequences. In Lebzelter et al. (2019), the corresponding objects are mostly (extreme) carbon rich Miras. It is quite remarkable that the stars that have confirmed SiO masers, have dominantly longer periods and lie on sequence *D*, as expected for oxygen-rich Miras. Typically, stars in that period range have a mass slightly over the solar mass and ages below 1 Gyr (Grady et al. 2019).

Finally, it is clear that most of the stars in the *local* sample are below the empirical P-L relation for the LMC (and SMC) Whitelock et al. (2008); Matsunaga & Team (2006). Possibly this is again related with the fact that we preferably select the closer, less luminous AGB stars, when we make our *BAaDE-Gaia* sample, while in the LMC (and SMC) the sample is biased towards the most luminous stars.

4.4.5 SiO maser emission from Mira variables

Although the subsample from the *local* sample for which SiO maser emission lines are detectable, is not complete and currently limited to 42 stars, preliminary statistics of this subsample can be presented briefly. Figure 4.14 shows the distribution of these stars and their motions from *Gaia* DR2 projected on the Galactic plane. Although this result is preliminary, it is clear that there is a preference to find SiO masers in the direction of the Galactic center. As discussed before, this can be due to an asymmetry between carbon rich stars and oxygen rich stars with respect to the solar circle. Otherwise, the stars appear to be found in the region between the high extinction regions associated with the local and Carina-Sagittarius arm. The apparently random motions with an rms of 30 km s^{-1} (mainly in the direction of the Galactic rotation component) are consistent with the expected kinematics for a relaxed, old stellar population (e.g., Schubert et al. 2010).

VLBI astrometry of selected bright BAaDE targets is currently being considered, in order to provide parallaxes and 3D orbits with a target accuracy of $\sim 50 \mu\text{as}$. This can provide orbits of stars beyond the *local* sample, allowing us to study the kinematics in the inner Galaxy. Astrometric VLBI would also provide a direct comparison of the parallax technique between optical (*Gaia* DR2) and radio, verifying alignment between the stellar photosphere and the origin of the circumstellar masers. Notably, synergies between VLBI and *Gaia* data are currently being used to refine astrometric measurements for *Gaia* objects (Lindgren 2019).

4.5 Conclusions

We have cross-matched the BAaDE target list, which consists of 28,062 IR sources preselected from the MSX colors at latitudes $|b| < 5^\circ$ to match evolved stars in the inner Galaxy (van der Veen & Habing 1998; Sjouwerman et al. 2009), with the *Gaia* DR2 catalogue (Gaia-Collaboration et al. 2018), finding 20,111 cross-matches. The cross-match was made using a conservative radius of $3''$ around the MSX position which has a positional accuracy of $2''$ (Price 1995). The 7,951 sources from the BAaDE target list that were not detected in *Gaia* DR2 correlate with lines of sight of high optical extinction in the Galactic plane. From the 20,111 cross-matched sources, stars with precise parallax estimates and within a 2 kpc radius around the Sun (where we can obtain accurate extinction maps) were selected. The remaining 2,060 stars constitute our *local* sample, representing a foreground population of evolved stars in the Galactic plane.

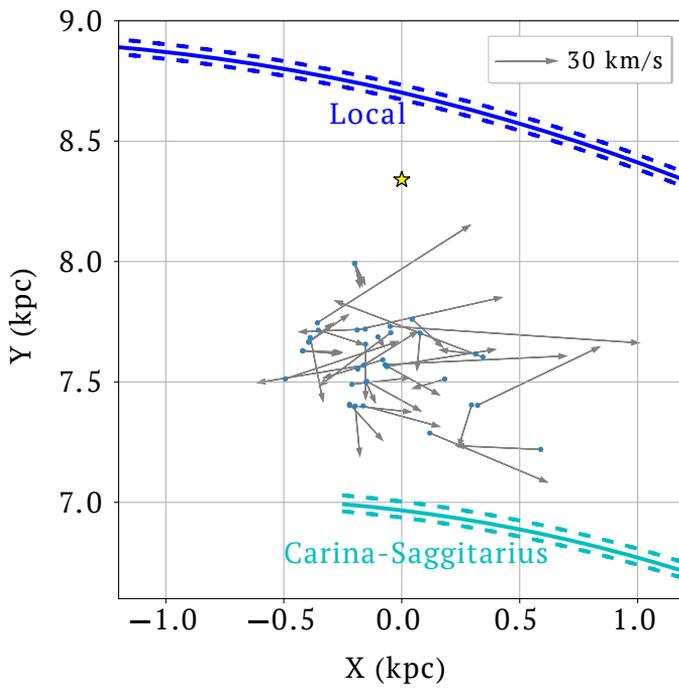


Figure 4.14: Galactic projection of the motion of BAaDE targets within the local sample with confirmed SiO maser emission. In this view, the Sun is located at $[0, 8.34]$ kpc (Reid et al. 2014), and the Galactic center at $[0.0]$. The location of the spiral arms and their widths follow (Reid et al. 2014).

Among the *local* sample, the Gaia DR2 shows large amplitude variability for 908 stars that have been classified as Mira variables (Mowlavi et al. 2018), of which another 653 have period estimates. Moreover, the BAaDE survey has already observed 395 of those 653 sources with the VLA and ALMA detecting SiO maser emission lines at 42 and 86 GHz coming from 45 sources.

Using radio, infrared and optical data for this sample, we have characterized the evolved stellar population around the Sun in terms of spatial, variability, bolometric luminosity distribution. The population of evolved stars close to the Sun displays the following features:

1. The absolute magnitude distribution at K-band peaks at -6.4 ± 1.2 mag with a spread of approximately 4 mag around the peak for the stars classified by *Gaia* as variables. While the brightest sources are consistent with the expected luminosities for optically identified Mira variables, it is clear that our sample, at distances <2 kpc, is contaminated by red, low luminosity objects, most likely YSOs. By applying additional filters based on variability we can restrict the sample to LPV stars, which have a narrower K magnitude distribution.
2. Using extinction and bolometric corrections from the literature, we are able to estimate bolometric magnitudes for the *local* foreground Galactic sample. The distribution peaks at -3.0 with a width of 1.8 mag. This peak is at fainter magnitudes than that obtained for Miras in the LMC (Whitelock et al. 2008) and also at a lower value than inferred for BAaDE sources in the inner Galaxy (Trapp et al. 2018). Although variability and uncertainties in the extinction and bolometric corrections are important, we argue that the main reason is the selection of faint, but nearby, sources that can be identified in the optical regime.
3. For those BAaDE objects that have *Gaia* periods, we are able to associate these with fundamental mode and first overtone pulsation sequences. The BAaDE foreground population contains moderate mass Mira variables, which are the sources that have a high probability to show SiO masers. Among the targets in the sample, carbon rich LPV stars also seem to be abundant.
4. The SiO masers detected so far are associated with fundamental mode pulsators inside Solar circle. Moreover, the kinematic distribution is consistent with that expected for a relaxed, old stellar population.

Overall we conclude that the BAaDE targets are predominantly made up of LPVs, optically detectable Miras and carbon stars. The infrared selection also picks up lower luminosity objects within 2 kpc from the sun. The sample of evolved stars at these distances is made up of AGB stars of moderate luminosity. To understand the nature of stars that make up the BAaDE sample in the inner Galaxy, advanced statistical methods that can use more uncertain *Gaia* data in a robust way will be required.

Acknowledgements. The BAaDE project is funded by National Science Foundation Grant 1517970/1518271. This paper uses data products obtained with instruments run by the National Radio Astronomy Observatory (NRAO): NSF's Karl G. Jansky Very Large Array (VLA; program ID: VLA/13A-331) and the Atacama Large Millimeter/submillimeter Array (ALMA). This paper makes use of the following ALMA data: ALMA/2013.1.01180.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated

by ESO, AUI/NRAO and NAOJ. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. This research made use of data products from the Midcourse Space Experiment. Processing of the data was funded by the Ballistic Missile Defense Organization with additional support from NASA Office of Space Science. This research has also made use of the NASA/IPAC Infrared Science Archive, which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

This publication makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

This work also has made use of data from the European Space Agency mission *Gaia*, processed by the *Gaia* Data Processing and Analysis Consortium (DPAC). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. This material is based upon work supported by the National Science Foundation under Grant Number 1517970.