











# Characterizing the Evolved Stellar Population in the Galactic Foreground. I. Bolometric Magnitudes, Spatial Distribution and Period–Luminosity Relations

Luis Henry Quiroga-Nuñez<sup>1,2,3,4,7</sup> , Huib Jan van Langevelde<sup>3,4</sup> , Loránt O. Sjouwerman<sup>1</sup> , Ylva M. Pihlström<sup>2,8</sup> , Anthony G. A. Brown<sup>3</sup> , R. Michael Rich<sup>5</sup> , Michael C. Stroh<sup>6</sup> , Megan O. Lewis<sup>1,2</sup> , and Harm J. Habing<sup>3</sup>

<sup>1</sup>National Radio Astronomy Observatory, Array Operations Center, Socorro, NM 87801, USA; [lquiroga@nrao.edu](mailto:lquiroga@nrao.edu)

<sup>2</sup>University of New Mexico, Department of Physics and Astronomy, Albuquerque, NM 87131, USA

<sup>3</sup>Leiden University, Leiden Observatory, Leiden, 2300RA, The Netherlands

<sup>4</sup>Joint Institute for VLBI ERIC (JIVE), Dwingeloo, 7990AA, The Netherlands

<sup>5</sup>University of California, Department of Physics and Astronomy, Los Angeles, CA 90095, USA

<sup>6</sup>Northwestern University, Center for Interdisciplinary Exploration and Research in Astrophysics, Evanston, IL 60201, USA

Received 2020 June 15; revised 2020 August 25; accepted 2020 September 3; published 2020 November 23

## 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 (IR) 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 very long baseline interferometry proper motions. In order to use the BAaDE sample for kinematic studies, the stellar population should be characterized. In this study, the BAaDE targets have been cross-matched with IR (2MASS) and optical (Gaia) samples. By exploring the synergies of this cross-match together with Gaia parallaxes and extinction maps, the local ( $d < 2$  kpc) asymptotic giant branch (AGB) stars can be characterized. 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 1812 evolved stars with accurate parallax measurements, confirmed evolved stellar evolution stage, and within 2 kpc distance around the Sun was selected, for which absolute (bolometric) magnitudes are estimated. The evolved stellar population with Gaia counterparts that are variable seems to be predominantly associated with AGB stars with moderate luminosity ( $1500_{-500}^{+3000} L_{\odot}$ ) and periods between 250 and 1250 days.

*Unified Astronomy Thesaurus concepts:* Galaxy stellar content (621); Solar neighborhood (1509); Asymptotic giant branch stars (2100); Periodic variable stars (1213); Astrometry (80); Surveys (1671)

## 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., Martínez-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. 2018b; Lindegren et al. 2018) delivers more accurate, reliable data in each data release (DR), 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 (Antoja et al. 2018; Helmi et al. 2018; Bland-Hawthorn et al. 2019; Belokurov et al. 2020).

Starting with the discovery of the Galactic 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 is seen in extragalactic edge-on boxy bulges. As these are the most prominent dynamic features in the inner Galaxy, research into the kinematics and stellar

populations that constitute the bar and the bulge is necessary to understand the morphology, structure, and evolution of the Milky Way (Bland-Hawthorn & Gerhard 2016). Evolved stars, which are prominent in the mid-infrared (mid-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; Rich et al. 2007) and 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 (SEDs) 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 (see, e.g., 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 microarcsecond accuracy (van Langevelde & Vlemmings 2003; Reid & Honma 2014).

Previous surveys focused first on OH masers (Sevenster et al. 2001; Fish et al. 2006) and later targeted SiO masers with

<sup>7</sup> Jansky Fellow of the National Radio Astronomy Observatory.

<sup>8</sup> Adjunct astronomer at the National Radio Astronomy Observatory.

single-dish telescopes (Messineo et al. 2018). When it was realized that the new capabilities at 7 mm of the NSF’s Karl G. Jansky Very Large Array (VLA) and at 3 mm of the Atacama Large Millimeter/submillimeter Array (ALMA) offer efficient ways to study SiO masers, the Bulge Asymmetries and Dynamical Evolution project (BAaDE<sup>9</sup>) was proposed. Using Midcourse Space Experiment (MSX) IR color selections, many thousands of SiO masers were 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.

Because 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 toward 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 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 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 region, recent studies have detected major Galactic structures (Reid et al. 2019; Alves et al. 2020) as well as several comoving groups and stellar structures (Kounkel et al. 2020).

In this paper, we cross-match the BAaDE sample with 2MASS and Gaia DR2. Because the BAaDE sample is based on MSX, it predominantly contains stars at low Galactic latitude. The cross section of the various surveys has IR as well as optical astrometric information. Through the Gaia DR2, we can evaluate other parameters such as the parallax and proper motion, but also information derived from the survey such as variability and stellar classification. 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), one can expect it to contain predominantly long-period variable (LPV) stars, likely Miras, with a modest circumstellar shell. But this sample may contain young stellar objects (YSOs), or main sequence (MS) or red giant branch (RGB) stars, which are very luminous (Lewis et al. 2020b), 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 Section 2. In the following sections, we present the main features of the different samples that resulted from the cross-matching. We start with the BAaDE–Gaia cross-match sample described in Section 3. Afterwards, we filter the sample to just the solar neighborhood defining the *local* sample in Section 4. In Section 5, we characterize the *local* sample in terms of luminosity, variability, Galactic distribution and period–luminosity (P-L) relation. These properties position us to comment on the nature of evolved stars in the foreground Galactic plane, for which we have Gaia counterparts with accurate distances. In a subsequent paper (L. H. Quiroga-Nuñez et al. 2020, in preparation), we will present an extended study of the features (e.g., kinematics, SiO maser emission and

<sup>9</sup> <http://www.phys.unm.edu/~baade/>

**Table 1**  
Sources Numbers Obtained for Different Samples and Cross-matches

Surveys	Sources
BAaDE (MSX)	28,062
$\cap$ 2MASS	25,809
$\cap$ Gaia DR2	20,111

**Note.** The intersection symbol ( $\cap$ ) indicates a cross-match between the surveys.

rates, carbon- and oxygen-rich stellar discernment, etc.) of the evolved stellar population in the Galactic foreground by using the current results of the BAaDE survey.

## 2. Cross-match at Different Wavelengths

We have cross-matched the MSX-based BAaDE target sample with 2MASS and Gaia DR2 (see Table 1), using the Gaia data archive interface.<sup>10</sup> This sample was defined as the BAaDE–Gaia sample. In the following subsections, we describe how this process was implemented, starting from the BAaDE target selection, followed by the cross-match criteria.

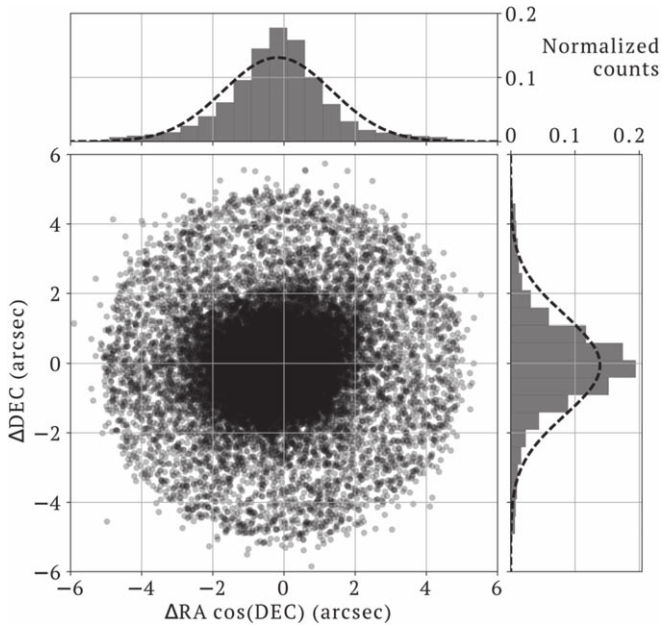
### 2.1. BAaDE Target Sample Selection

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 two color–color diagram (2CD), van der Veen & Habing (1988) 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 2CD, suggesting an evolutionary track with increasing mass-loss rate. In the IRAS 2CD, SiO maser stars are expected to be found in a specific color region, facilitating a selection based on the IRAS colors. However, the angular resolution of IRAS varied between about  $0.5'$  and  $2'$ , limiting a large-scale survey, particularly in the Galactic plane. Later, Sjouwerman et al. (2009) were able to transform the IRAS 2CD sequence to colors in the mid-IR, using MSX data. By doing this, the positional accuracy of identified IR sources was improved to  $2''$  (Price 1995), and a new sample of AGB stellar 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  $|b| < 5^\circ$ . One-third of the BAaDE target sample is expected to lie in the Galactic bulge (Sjouwerman et al. 2017). The 28,062 targets are being followed up in order to detect SiO maser emission at 43 GHz with the VLA or 86 GHz with ALMA. So far, 20,600 candidates have been observed, of which 16,335 have already been analyzed (14,548 with the VLA and 1787 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.

### 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 a  $3''$  radius around the BAaDE targets. The motivation for this separation was based on the assumption that the distribution of deviations from the actual positions is

<sup>10</sup> <http://gea.esac.esa.int/archive/>



**Figure 1.** 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 (see Section 2.2).

dominated by the MSX data (as confirmed by Pihlström et al. 2018) and has Gaussian distributions in both components ( $\Delta\alpha \cos(\delta)$ ,  $\Delta\delta$ ) with absolute mean values  $<0''.2$  and positional accuracy of around  $2''$ , as seen in Figure 1. Therefore, a  $3''$  radius was chosen as a conservative match radius between the BAaDE targets and Gaia. Note that the criterion we used 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. 2018), 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. 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 (see Table 1), where all of them were found to be one-to-one correspondences. This last sample of 20,111 is called the BAaDE–Gaia sample throughout the paper and thus includes 2MASS information. Notably, for 7951 BAaDE targets (33%), there were no Gaia counterparts, probably due to the fact that these targets lie behind considerable dust extinction at optical wavelengths. Figure 2 shows how the

distribution of these “missing” sources indeed correlates with the dust-obscured regions that Gaia could not penetrate.

### 2.2.1. Statistics of the Cross-matches

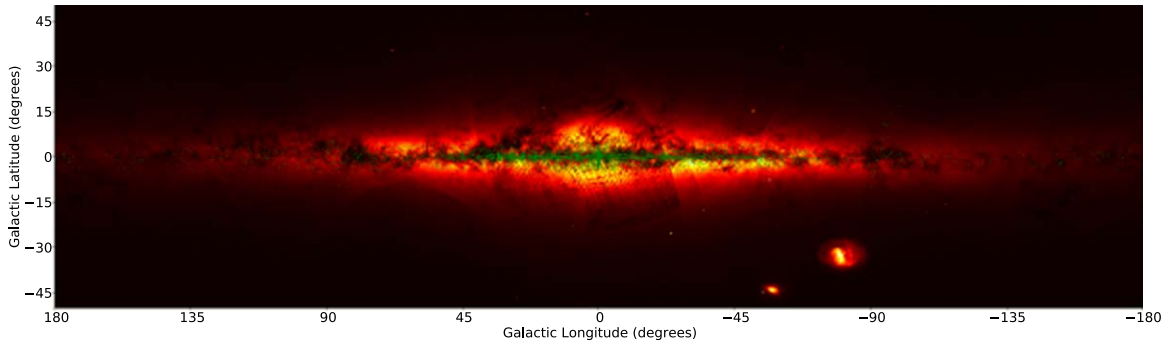
Assuming a uniform distribution of sources in the bulge for the Gaia detections as well as for BAaDE targets, one can calculate the number of sources that will give random matches at the given resolution of each survey. We estimated that the number of random matches should be less than 1200; this is 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 1200, indicating that our sample has at most a modest contamination of sources with unrelated counterparts.

## 3. Features of the BAaDE–Gaia Sample

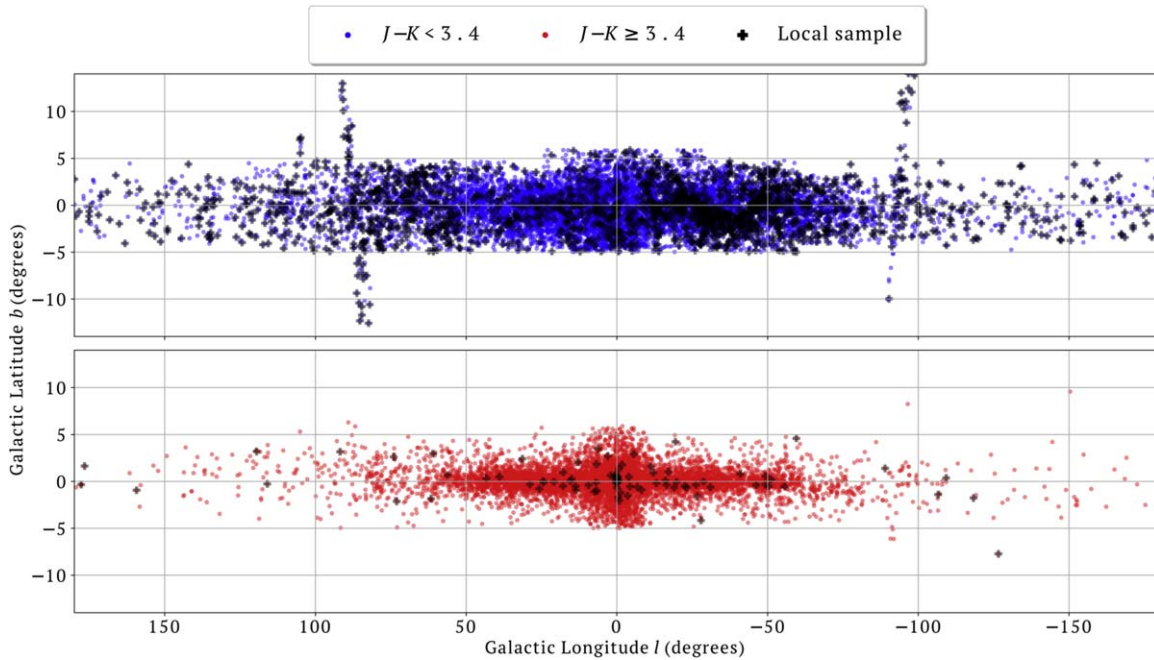
Because the BAaDE–Gaia sample was obtained through 2MASS, the mean value of the near-IR color ( $J-K$ ) can be used to split the sample in two equal-size subsamples: i.e.,  $(J-K) < 3.4$  for the bluer stars and  $(J-K) \geq 3.4$  for the redder stars. More extreme AGB stars (more luminous and with thicker shells) are expected to have steeper slopes in their SEDs at near-IR wavelengths, resulting in increasingly redder IR colors. Figure 3 shows the subsamples of red and blue stars in a Galactic latitude–longitude diagram. Red stars seem to better trace the inner part of the Galaxy (Galactic bulge and plane) while bluer stars seem to dominate the foreground population. Indeed, as we will detail in Sections 4 and 5, Figure 3 also shows that sources in the solar neighborhood ( $<2$  kpc) are mainly stars that are bluer (in the context of the BAaDE selection), in particular those with Gaia counterparts. We confirm that by splitting the sample using IR photometry, two samples can be traced. This has already been observed by Trapp et al. (2018), who made the split using  $K$  magnitudes, and labeled the two kinematic populations “cold” (the bluer, brighter stars in the Galactic disk) and “hot” (the redder stars in the bulge/bar). However, although it is indeed expected that more extreme stars are redder, we must highlight that the increased extinction with distance (toward the bulge) also makes them redder. Therefore, this partly explains why these stars show up nicely as bulge sources (Figure 3) and seem to be a better tracer of the inner galaxy.

Another property that can be investigated for the Gaia–BAaDE sample is variability. Although Gaia DR2 has variability information for a considerable number of stars (Mowlavi et al. 2018), Belokurov et al. (2017) have shown that—already with Gaia DR1—flux uncertainties quoted in the Gaia catalog 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 to as amplitude index throughout this paper, using the mean flux ( $\bar{I}_g$ ) and its error ( $\sigma_{\bar{I}_g}$ ) in the optical  $G$  band as  $\log_{10}\left(\sqrt{N_{\text{obs}}}\frac{\sigma_{\bar{I}_g}}{\bar{I}_g}\right)$ , where  $N_{\text{obs}}$  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), finding that





**Figure 2.** Galactic distribution of the BAaDE stellar targets without a Gaia counterpart (green points) overplotted on the sky map from Gaia DR1. This sample accurately correlates with highly obscured regions in the optical regime. Credit: ESA/Gaia/DPAC.



**Figure 3.** Galactic longitude–latitude diagram for the cross-matches obtained between BAaDE, 2MASS, and Gaia defined as the BAaDE–Gaia sample. This sample was split in two populations (upper and lower panels) 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 using a subset of the BAaDE survey. Black crosses represent the defined *local* sample, a subsample of evolved stars in the BAaDE–Gaia sample with accurate parallax measurements at  $< 2$  kpc distance (see Sections 4 and 5 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 the MSX target list (and are also BAaDE targets) caused by the target selection made by MSX in order to fill the strips missed by the all-sky survey by IRAS (Egan et al. 2003).

Mira variables have an amplitude index  $> -1.0$ . Figure 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. However, although the amplitude index seem to be an useful tool to estimate variability, it might have issues when it is applied to a sample with different properties from the Magellanic system or when more observations become available ( $N_{\text{obs}}$ ). Therefore, we could expect that the cutoff changes depending on the stellar population.

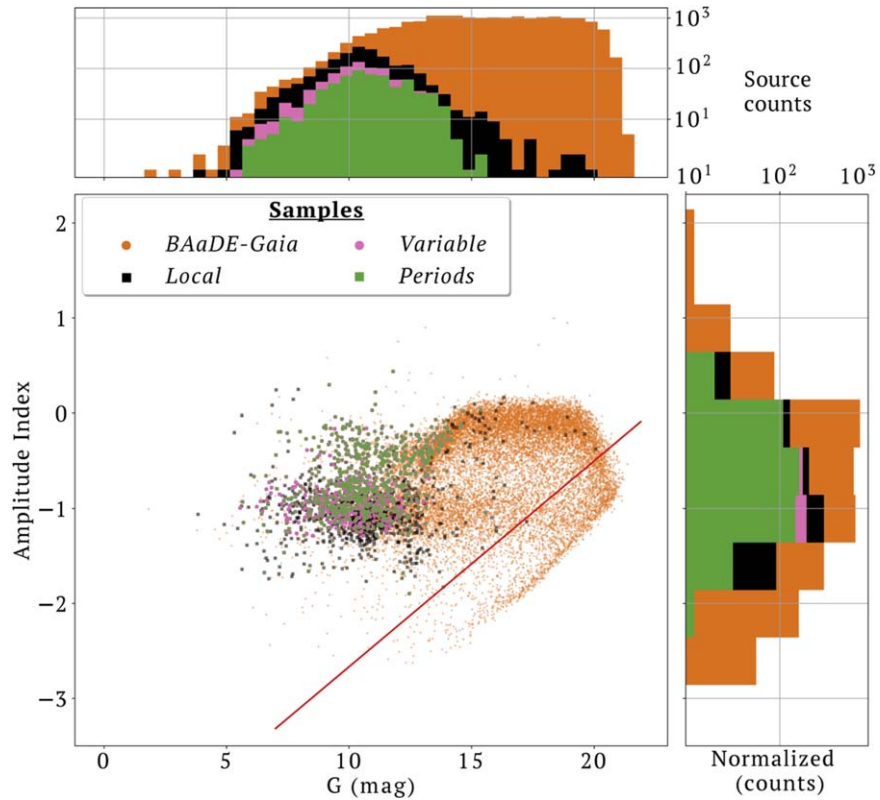
Figure 4 also shows that the variable stars defined by Mowlavi et al. (2018) within Gaia DR2 (see Section 5) coincide with larger-amplitude values as expected, confirming that indeed the IR classification made by the BAaDE project correlates with variable stars. However, this qualification is restricted to stars that are bright in the  $G$  band.

#### 4. Filtering the BAaDE–Gaia Sample for Galactic Foreground Sources

As the objective of this study is to characterize the evolved stars in the BAaDE target list, we apply 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, finally arriving at the resulting sample of evolved stars in the foreground Galactic plane, which we define as the *local* sample. Figure 5 summarizes the resulting subsamples.

##### 4.1. Parallax Measurements

Obtaining distance estimates from noisy parallax measurements can be a complex issue (see e.g., Bailer-Jones 2015).



**Figure 4.** Amplitude–magnitude diagram first suggested by Belokurov et al. (2017), used to distinguish variable stars where larger amplitudes are likely associated with pulsating AGB stars. The orange points represent the BAaDE–Gaia sample ( $\text{BAaDE} \cap 2\text{MASS} \cap \text{Gaia DR2}$ ), whereas black sources represent the stars within the *local* sample (i.e., accurate distance estimates for evolved stars at  $<2$  kpc from the Sun). The pink and green points show two subsamples (*variable* and *periods*) derived from the *local* sample (see Figure 5 and Section 5). The solid red line represents a threshold above which sources are predicted to show variability larger than expected for a constant star at the given  $G$  magnitude (Belokurov et al. 2017). Possible features present in the plot (e.g., the apparent bimodality of the amplitude index) will be further explored in a subsequent paper.

Sample name	Description or filters used	Sources
<i>BAaDE-Gaia</i>	$\text{BAaDE (MSX)} \cap 2\text{MASS} \cap \text{Gaia DR2}$	20,111
	└─ $\sigma_{\varpi}/\varpi < 0.2$	2,277
<i>Local</i>	└─ $r < 2$ kpc & evolutionary stage	1,812
<i>Variable</i>	└─ Variable	898
<i>Periods</i>	└─ Periods	649

**Figure 5.** Number of sources obtained for each subsample of the BAaDE–Gaia sample. Each row represents a filter used. See Section 5 for a detailed description of each filter. The definition of the BAaDE–Gaia and local samples is given in Sections 2.2 and 5, respectively. The arrow symbols indicate the subsample, whereas the intersection symbol ( $\cap$ ) indicates cross-match between the surveys.

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 of 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, if for a specific star  $\sigma_{\varpi}/\varpi < 0.2$ , one could obtain an accurate estimate of the distance without further considerations (Bailer-Jones 2015). In this sense, we find that most (91%) of the stars of the BAaDE–Gaia sample that have  $\sigma_{\varpi}/\varpi < 0.2$  are limited up to 2 kpc. Moreover, because (1) the aim of this research is to study the foreground population of evolved stars and (2) accurate extinction maps are limited to 2 kpc (see the following

subsection), we focus on the solar neighborhood ( $<2$  kpc). Finally, an analysis of the foreground sample can be considered an initial step for doing a full statistical analysis.

We have also investigated the effect of the Gaia parallax zero point for our targets. In principle, the Gaia parallax zero point can be up to  $100 \mu\text{as}$  depending on the method and sample used (e.g., Zinn et al. 2019; Chan & Bovy 2020, and references therein). However, several studies concur that the Gaia parallax zero point for red clump and variable stars oscillates around  $-50 \mu\text{as}$  (Groenewegen 2018; Riess et al. 2018; Van Langevelde et al. 2018; Zinn et al. 2019; Chan & Bovy 2020). Such an offset may cause a shift of less than 5% of the average values of the physical quantities reported in this work.

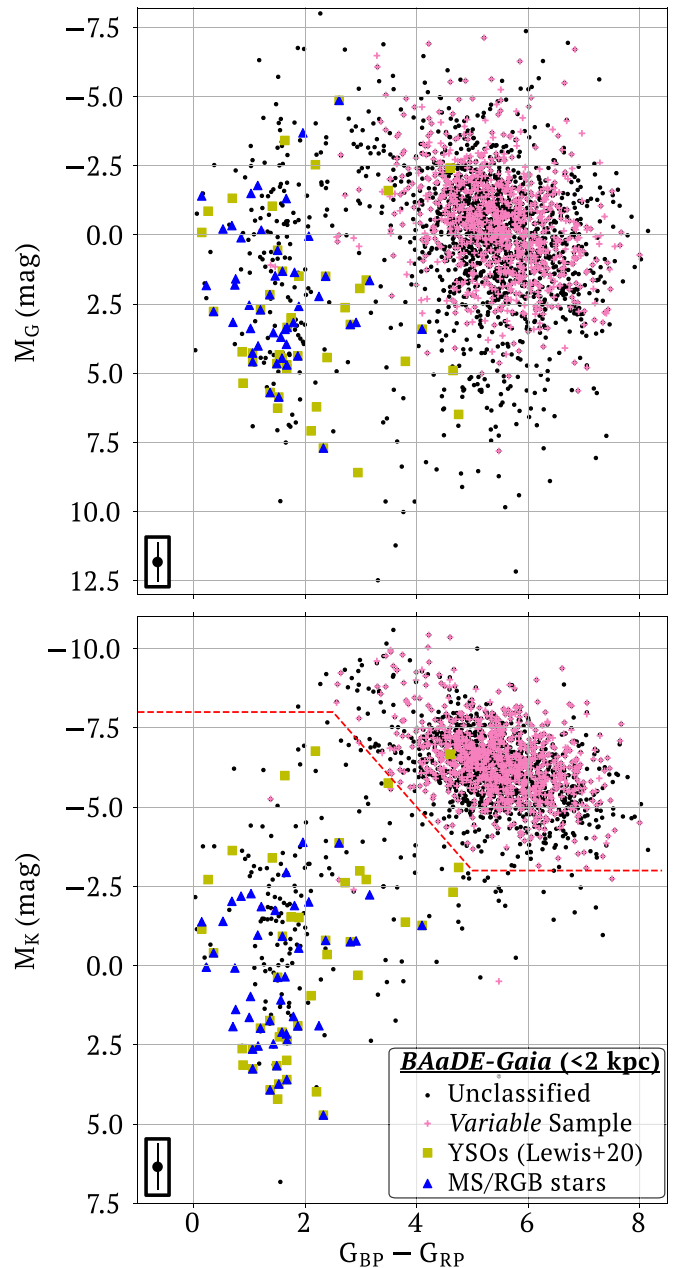
#### 4.2. Extinction Maps up to 2 kpc

Capitani 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 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.

#### 4.3. Younger Stars Detected in Hertzsprung–Russell Diagrams

Using the Gaia DR2 results, Gaia Collaboration et al. (2018a) have generated several observational Hertzsprung–Russell diagrams (HRDs). Particularly, for a sample of 29,288 low-extinction nearby giants (i.e.,  $>2$  mas parallax,  $E(B - V) < 0.015$ , and  $M_G < 2.5$ ), they were able to locate the AGB bump (at  $M_G \sim -0.5$  and  $G_{BP} - G_{RP} \sim 1.5$ ), which corresponds to the starting point of the AGB where stars are burning their helium shell (see Figure 10 in Gaia Collaboration et al. 2018a). In order to compare these results with the BAaDE targets in the foreground Galactic plane, we generate the HRDs shown in Figure 6. These diagrams use Gaia colors,  $K$ -apparent magnitudes from 2MASS, accurate distance estimates, and extinction maps for the resulting BAaDE–Gaia cross-match around the Sun ( $<2$  kpc). Moreover, the *variable* sample was also overplotted (pink crosses in Figure 6) to support the statement that sources already classified as Mira by Gaia fall in a defined location within the diagram (see Section 5 for the definition of the *variable* sample). This location is close to where the expected AGB bump is happening ( $M_G \sim 0.5$ ; Gaia Collaboration et al. 2018a), but with expected redder colors and fainter sources due to dust—a combination of circumstellar and ISM reddening—than those reported in the Gaia HRD, as their sample was filtered for brighter sources ( $M_G < 2.5$ ; Gaia Collaboration et al. 2018a). Therefore, we conclude that sources concentrated around  $M_G \sim -0.5$  with redder colors are certainly AGB stars with significant circumstellar shells.

On the other hand, we have also confirmed that the IR color selection, by which the BAaDE targets were selected, does not prevent a fraction of YSOs and MS/RGB stars from entering into the target sample. In fact, Lewis et al. (2020b) have shown that by using MSX colors, specifically the MSX color  $[D] - [E]$ , a differentiation between YSOs and AGB stars can be invoked within the BAaDE sources. Moreover, Y. M. Pihlström et al.



**Figure 6.** Absolute magnitude in  $G$  (upper panel) and  $K$  (lower panel) bands as a function of Gaia color for the BAaDE–Gaia sample with a distance of  $<2$  kpc. The typical errors in magnitude are shown in the left corner of each diagram. The sources in the sample that are also classified as MIRA (*variable* sample; see Section 5) by Gaia are marked as pink crosses. As yellow squares, we marked the sources that Lewis et al. (2020b) have confirmed as YSOs, where blue triangles are those likely associated with MS/RGB stars (Y. M. Pihlström et al. 2020, in preparation). The dashed red line was established to distinguish evolved stars in the upper region (which defines the *local* sample) from other BAaDE targets in the lower region as YSOs and MS/RGB stars (see Section 4.3). Absolute magnitudes in the  $K$  band for typical low-metallicity Mira variables reported in the LMC by Whitelock et al. (2008) cover a range between  $-8.0$  and  $-6.0$  mag, whereas the AGB bump for Gaia data seems to be located at  $M_G \sim -0.5$  (Gaia Collaboration et al. 2018a).

(2020, in preparation) identified BAaDE sources that, based on their IR photometry, are likely associated with either reddened, massive MS stars (B or A stellar type) or RGB stars. These samples are also shown in Figure 6 to highlight their position in the HRDs, confirming that they are likely not AGBs.

Because the distribution in  $M_K$  is less dispersed for the preselected AGBs sources, we made a cut in this diagram (lower panel in Figure 6) to filter non-AGB sources. We define the *local* sample as the sources with the following Gaia colors and absolute  $K$ -magnitude conditions:

$$G_{\text{BP}} - G_{\text{RP}} \begin{cases} < 3 & \text{and } M_k < -8 \\ > 5 & \text{and } M_k < -3 \\ \text{rest} & \text{and } M_k \leq 2(G_{\text{BP}} - G_{\text{RP}}) - 13. \end{cases}$$

The non-AGB stellar sources were filtered out, leaving 1812 local stellar sources in the AGB regime. This means that we were able to confirm the evolved stellar evolutionary stage for most (88%) of the BAaDE targets in the foreground Galactic plane. From those, there was only one source that Lewis et al. (2020b) classified as a YSO. We have confirmed that this source falls very close to the empirical MSX color frontier defined by Lewis et al. (2020b) for YSOs and, therefore, is likely a genuine AGB star after all.

## 5. The Foreground Population of Evolved Stars: the 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 in Section 4, leaving a sample of 1812 stellar sources that we have defined as the *local* sample. This sample contains BAaDE targets associated with AGB stars within a 2 kpc distance around the Sun with accurate distance estimates, IR and optical photometry, and proper motions.

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 extracted those objects that were flagged as variables of any kind, which we define as the *variable* sample (898 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 the *periods* sample (649 sources). Note that all sources within the *local* sample contained in the table `vari_classifier_result` were classified by Gaia as Mira/semiregular (SR) stars (MIRA\_SR). The characteristics (variable and periods) described allowed us to generate subsamples of the *local* sample, as shown in Figure 5.

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 color variations of the stellar photosphere (see e.g., Lindegren et al. 2018; Van Langevelde et al. 2018), 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.

In the following subsections, we research the Galactic foreground sample of BAaDE targets (*local* sample) and its different subsamples (*variable* and *periods*) in terms of IR photometry, absolute and bolometric magnitudes, variability, Galactic distribution, and the P-L relation.

### 5.1. Infrared Photometry

The SED of AGB stars usually peaks at IR wavelengths; therefore, these stars have been usually identified by their IR colors (see, e.g., van der Veen & Habing 1988). In particular, after the 2MASS data release (Skrutskie et al. 2006),  $K$

measurements have been widely used to characterize these populations (Whitlock et al. 2008; Messineo et al. 2018). Figure 7 shows the distribution of the apparent  $K$  magnitude obtained from 2MASS for the entire cross-matched sample, with the different subsamples in Figure 5. We note that by filtering the foreground sample with Gaia counterparts (the *local* sample), we are selecting brighter stars in the  $K$  band.

By using the optical extinction maps described in Section 4.2, we obtain the extinction and reddening estimates at the  $K$  band by assuming  $A_\lambda/A_k = (\lambda/2.12 \mu\text{m})^{-1.9}$  (Messineo 2004). As one could expect for the region around the Sun, the IR extinction estimates at the IR  $K$  band for the filtered sample are usually lower than 0.5 mag (see lower panel of Figure 8). Finally, as we have accurate distance estimates for the *local* sample (see the upper panel of Figure 8), we are able to estimate the absolute  $K$ -magnitude distribution ( $M_k$  in Figure 9) as described in Section 5.2.

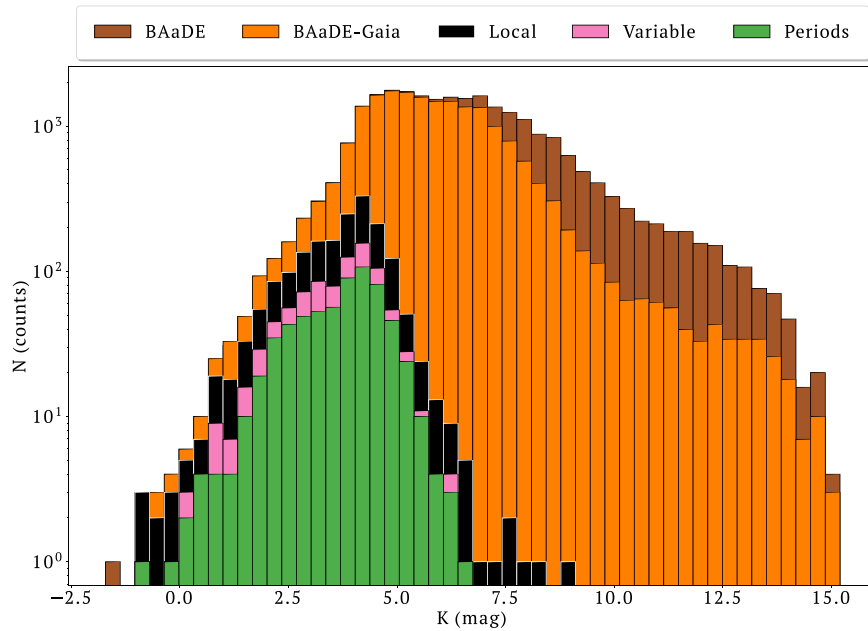
### 5.2. Absolute Magnitudes for the Foreground Mira Population

Several studies have been carried out to estimate IR absolute magnitudes of the AGB populations in the LMC, where the distance to the stellar system is known, and therefore, the distance modulus (and presumably also the IR extinction) can be assumed to be the same for each object (see e.g., Whitlock et al. 2008). From Figure 6, we can determine an average absolute  $K$  magnitude of  $M_k = -6.3 \pm 1.2$  mag for the *local* sample. Although the magnitude values found roughly correspond to those found in the LMC (i.e., between  $-8.0$  and  $-6.0$  mag; see, e.g., Whitlock et al. 2008), one should keep in mind that we established a fainter limit of  $M_k = -2.5$ . Nevertheless, with the aim of analyzing the absolute magnitude distribution using the variability classification defined by Gaia (Mowlavi et al. 2018) and the amplitude index (Belokurov et al. 2017), we made Figure 9, where the distribution of absolute magnitude for the *local* sample without considering the filtering of YSOs and MS/RGB (made in Section 4.3) is shown. Figure 9 shows that by including these younger sources, the distributions are much broader in terms of absolute magnitude range. This can be partly explained by the current sample being mostly based on a single 2MASS observation and including the effects of large amplitude variability. However, when we filter the sample by a variability qualification (as shown in both panels of Figure 9), the low-luminosity tail (in  $M_k$ ) is cut out. In this sense, the Gaia classification as Mira or SR variable seems to narrow the distribution more. Following a similar discussion in Mowlavi et al. (2018), we argue that the low-luminosity tail in Figure 9 and also Figure 7 is due to contamination with YSOs and MS/RGB stars that can also peak in the IR but do not show the same variability (Lewis et al. 2020b; Y. M. Pihlström et al. 2020, in preparation).

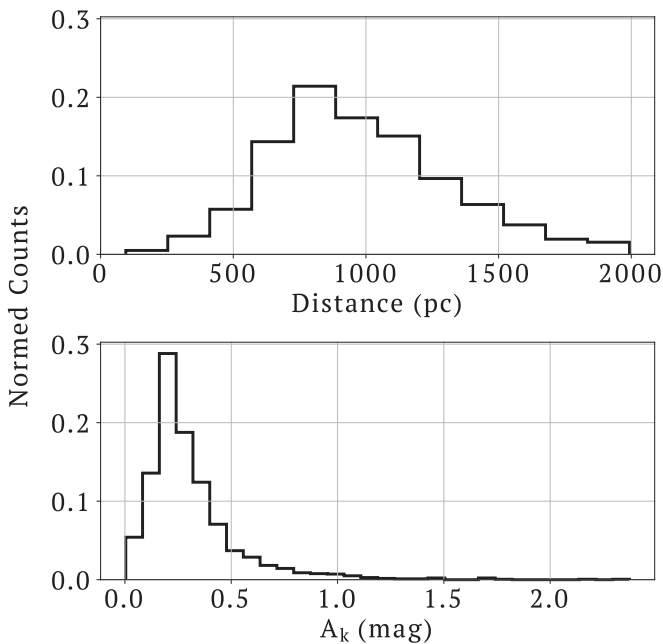
### 5.3. Bolometric Magnitudes for the Foreground Mira Population

The bolometric luminosity is a fundamental property useful for classifying stellar populations and evolutionary stages (Srinivasan et al. 2009), as it measures the intrinsic stellar power. Although its definition is straightforwardly formulated as the total integrated power over all frequencies, in practice, complete photometric measurements that allow a direct bolometric luminosity estimate are hardly ever available. Therefore, under various assumptions, a limited set of





**Figure 7.** Histogram comparison for the distribution of the  $K$ -magnitude observed by 2MASS for different samples as shown in Table 1 and Figure 5. The figure also shows that when more filters are applied, we are targeting the nearby stars, making the distribution narrower toward the apparently brighter stars.



**Figure 8.** Upper panel: distribution of the 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 the IR  $K$  band following Messineo (2004).

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 BC, which is usually parameterized using IR colors (see, e.g., Whitelock et al. 2008; Messineo et al. 2018; Lebzelter et al. 2019).

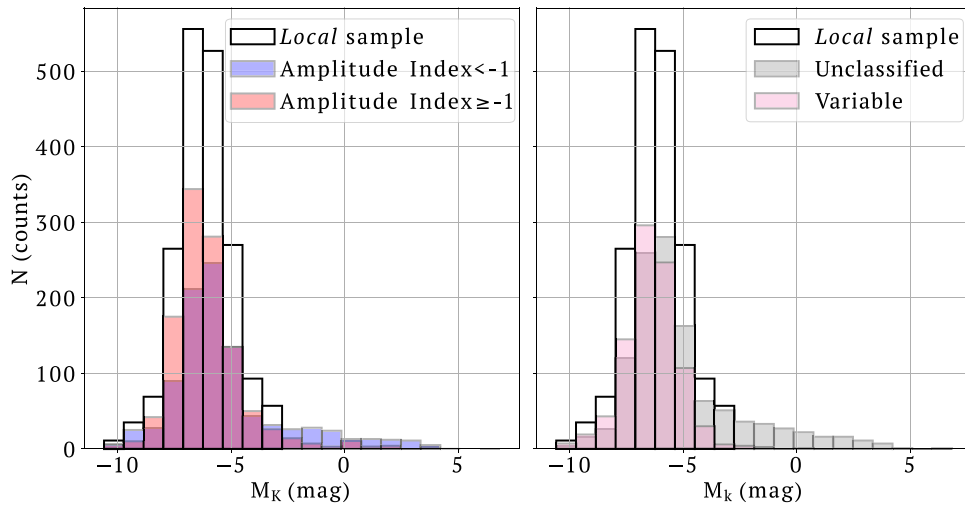
Trapp et al. (2018) have estimated the bolometric magnitude for a subset of the BAaDE sample. They considered a kinematically “cold” population of Galactic disk stars, which is similar to what is defined here as the foreground population

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 10 shows the bolometric distributions obtained. The offset between the bolometric distributions (the ones obtained for our samples and those obtained by Trapp et al. 2018) is likely caused by the distance assumption made by Trapp et al. (2018), which is equivalent to a range between 2.2 and 3.8 mag when taking the average distance of the *local* sample ( $1000 \pm 350$  pc; upper panel of Figure 8).

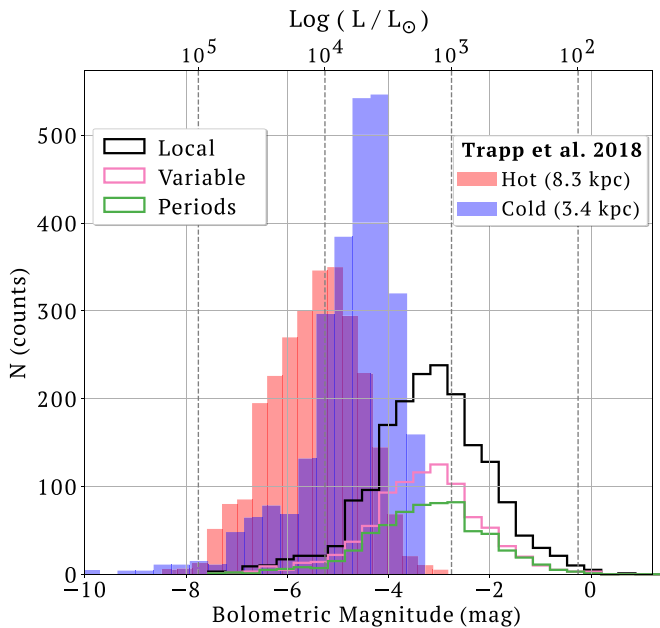
So far, we have based our bolometric magnitude estimates on  $M_K$  (Section 5.2), which in turn was estimated by using 2MASS IR photometry, Gaia parallaxes, and extinction maps following the approach of Messineo (2004). We define it as  $M_K$  2MASS. However, we can also obtain absolute  $K$  magnitudes directly from Gaia data, and we define it as  $M_K$  Gaia-Messineo. This one was calculated by using Gaia photometry in the  $G$  band, Gaia parallaxes, and dust maps to estimate first the absolute  $G$  magnitude for each star. Then, we use the  $BC_G$  provided for Gaia data (Andrae et al. 2018) to estimate the bolometric magnitude, and finally, we estimate  $M_K$  using the  $BC_k$  provided by Messineo (2004).

If we assume that both BCs produce similar results, then the  $M_K$  obtained for each star should be the same. In other words, a plot between  $M_K$  2MASS and  $M_K$  Gaia-Messineo should produce a linear relation with a slope of 1 and an intercept of 0, which is not the case. Figure 11 shows this plot, where we have done a linear fitting by forcing a slope of one—and therefore, we are assuming that both  $M_K$  estimates must be equal—finding an offset of  $-1.4 \pm 1.0$  mag ( $M_K$  2MASS value at  $M_K$  Gaia-Messineo = 0), indicating that the BCs in the Gaia DR2 seem to be overestimating the total luminosity calculated by Andrae et al. (2018) of our very red, variable objects. Further research comparing different stellar populations is needed to





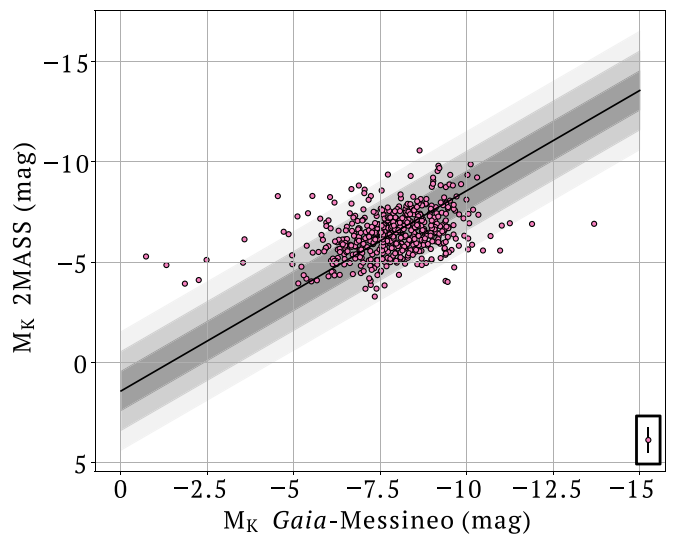
**Figure 9.** Each panel shows the absolute  $K$ -magnitude distribution as filled bars for the *local* sample before removing by YSOs and MS/RGB stars, and split by the variability criteria. These criteria are the amplitude index (Belokurov et al. 2017) and the Gaia DR2 variability classification (Mowlavi et al. 2018), in the left and right panels, respectively. The *local* sample (without YSOs and MS/RGB stars) is also shown for comparison as a black unfilled histograms. We found that the amplitude index method seems to find variability in objects that were not classified as variables in the Gaia DR2.



**Figure 10.** Luminosity and bolometric magnitude distributions for the *local* sample and the subsamples resulting from different filters (Figure 5). 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 the 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” populations using an approximate kinematic model are shown as filled histograms. Note that the kinematic cold population proposed by Trapp et al. (2018) is made up of stars in the Galactic disk and not in the bulge; therefore, similarities with respect to the *local* sample defined in this work are expected.

refine the Gaia BC at least for these red stars. We continue to use only the  $M_K$  based on 2MASS IR photometry, Gaia parallaxes, and extinction maps, i.e.,  $M_K$  2MASS.

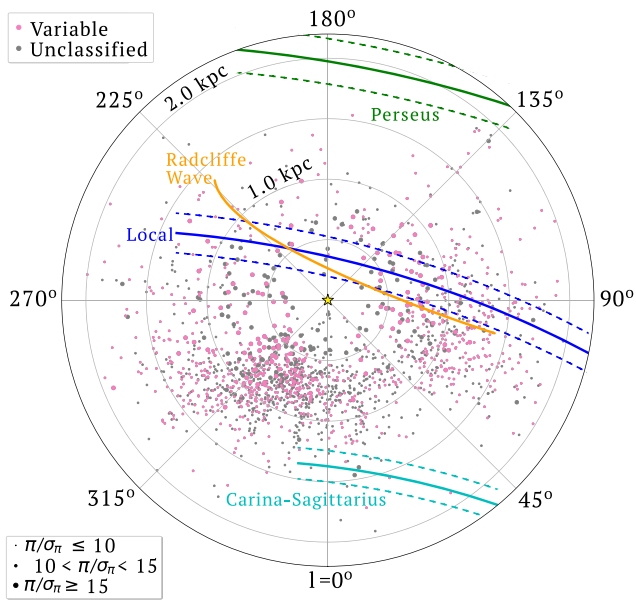
In Figure 10, we also present the luminosity distribution for the *local* sample. It shows that our sample is made up of giant stars with a luminosity range that is consistent with AGB stars, mostly Mira variables (Srinivasan et al. 2009). Compared to previous studies of Mira variables in the LMC or Galactic bulge (where fixed distances have been assumed), we have



**Figure 11.** Comparison between the absolute  $K$  magnitudes estimated from 2MASS data, Gaia parallaxes, and extinction maps ( $M_K$  2MASS, typical errors are shown in the bottom-right corner) with respect to the absolute bolometric estimates reported for variable stars in Gaia DR2 that were transformed to the absolute  $K$  magnitude using the BC $_k$  in Messineo et al. (2018;  $M_K$  Gaia-Messineo). The black line describes the linear fitting that was forced to have a slope of 1. The gray layers contain the  $1\sigma$ ,  $2\sigma$ , and  $3\sigma$  deviation from the linear fitting (see Section 5.3).

found less luminous objects. This, of course, is expected in our selection that was based on a combination of IR detections, optical Gaia counterparts, extinction maps, and distance selection. We found that the typical luminosity for the *local* sample is estimated as  $1500_{-500}^{+3000} L_\odot$ , suggesting that evolved stars in the solar neighborhood are found to be moderately luminous stars, likely associated with low-mass stars.

Another plausible explanation of the moderately luminous stars found in the solar neighborhood could come from the BC used in this research. In principle, we can affirm that we are obtaining an accurate absolute  $K$ - and  $G$ -magnitude distribution for the *local* sample sources given that (1) the magnitude ranges found are expected for AGB stars, and (2) their evolutionary stages were confirmed by comparing with Gaia HRDs (see Figure 6 and Section 4.3). Therefore, if there is a



**Figure 12.** Foreground Galactic distribution of the *local* sample. The sample was split into *Variable* and *Unclassified* according to Gaia DR2 (Mowlavi et al. 2018). The size of the marker is a measure of the relative parallax uncertainty, and, therefore, the distance uncertainty for each source. The positions and widths of the spiral arms are based on Reid et al. (2014), whereas the Radcliffe wave is based on Alves et al. (2020). These structures were found not to be correlated with the occurrence of the evolved stellar sources (see Section 5.5). The stellar distribution found relies on (1) the BAADE target selection made and (2) the dust distribution (Capitanio et al. 2017; Lallement et al. 2019).

miscalculation in the bolometric estimates, it could come from the BC proposed by Messineo (2004), as this BC was determined for AGB stars located in the bulge. The metallicity difference with respect to nearby AGB stars could cause a significant change in the luminosity estimate made. A planned research using BAADE targets in the bulge would clarify this point.

#### 5.4. Variability

At the beginning of Section 5, we have described the tables from Gaia DR2 that yield the variability classification that can be used for the *local* sample. Moreover, we noted that the ratio between the flux error and mean in *G* magnitude can be used to identify pulsating stars when the amplitude index  $> -1$ . We have considered both methods, in particular in relation to the *K*-band apparent and absolute magnitude distributions. Figures 4 and 9 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.

#### 5.5. Spatial Distribution in the Solar Neighborhood

The Galactic distribution of AGB stars has been studied extensively using IRAS, Wide-field Infrared Survey Explorer, 2MASS, and MSX data (Habing 1996; Jackson et al. 2002; Sjouwerman et al. 2009; Lian et al. 2014; Messineo et al. 2018).

Generally, it has been found that AGB stars are tracing the dynamically relaxed stellar population of the Galactic thick disk. This is expected, as these old stars have already migrated from their birth place, which is now detached from the spiral structure usually traced by young massive stellar objects (Quiroga-Nuñez et al. 2017; Reid et al. 2019). 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 Figure 12, which shows the projected spatial distribution of the *local* sample.

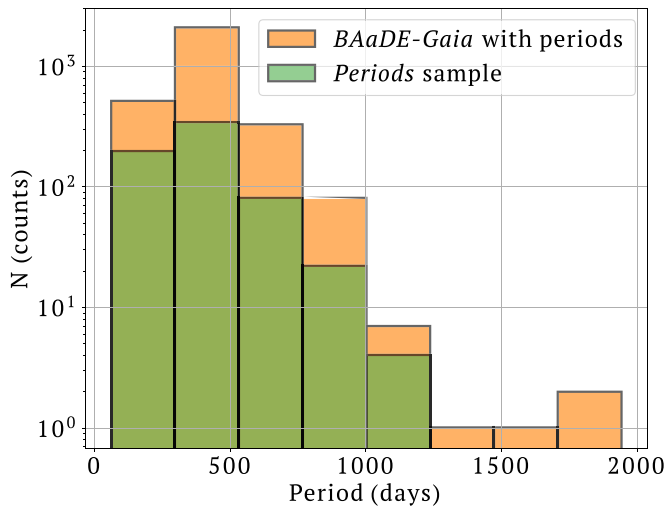
We have found that the number density of sources in the solar vicinity ( $< 0.5$  kpc) is considerably lower than farther out. The depletion of targets around the Sun originates from the fact that the MSX catalog, on which our sample is based, is mostly limited to  $|b| < 5^\circ$ . Therefore, the volume sampled increases with distance. From this, we estimate a corresponding scale height of  $\sim 50$  pc, equally for most of the objects in the *local* sample ( $> 0.5$  kpc). This seems to suggest that the scale height for our BAADE targets is lower than the scale height of the Galactic disk, i.e.,  $\sim 300$  pc (Jackson et al. 2002).

Recent studies using hundreds of maser-bearing stars have suggested a correlation between the position of evolved stars and the spiral arm structure at larger Galactic scales (up to 6 kpc; Gorski & Barmby 2020; Urago et al. 2020). In our study, however, we have found that this is not the case at least for AGBs in the solar neighborhood ( $< 2$  kpc). This can be seen in Figure 12, where there is not a clear correlation with any of the two major Galactic structures in the region: the local spiral arm (Reid et al. 2014) and the recently discovered Radcliffe wave (Alves et al. 2020). We must add that the *local* sample is affected by the interstellar extinction in the Gaia bands, meaning that we might be biased to miss some sources at the highly extinguished regions (i.e., the large Galactic structures as they are defined in terms of star-forming regions usually not reachable by Gaia; Quiroga-Nuñez et al. 2019). Moreover, there is still a radial gradient detectable with more targets toward the center than observed in the anticenter direction. This arises due to the MSX criteria, defined by Sjouwerman et al. (2009), optimized (1) to look for higher stellar density toward the inner Galaxy and (2) to detect SiO masers that are hosted by 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; Lewis et al. 2020b), which will be verified using the SiO maser detections currently done by BAADE in this region.

Finally, we have also noted two features in the source distribution presented in Figure 12. First, the outer Galactic region contains a lower number of confirmed cross-matches than is expected from the BAADE target selection, which, in turn, comes from the MSX selection. Second, there is no notable difference in the spatial distribution of variable stars compared to the unclassified (nonvariable) sources according to Gaia DR2.

#### 5.6. P-L Relations

Accurately determined periods have been the means to distinguish variable stars, particularly pulsating variables (LPV, Cepheids, RR Lyrae, RV Tauri, etc.) within Gaia (Gaia Collaboration et al. 2019). Figure 13 displays the distribution



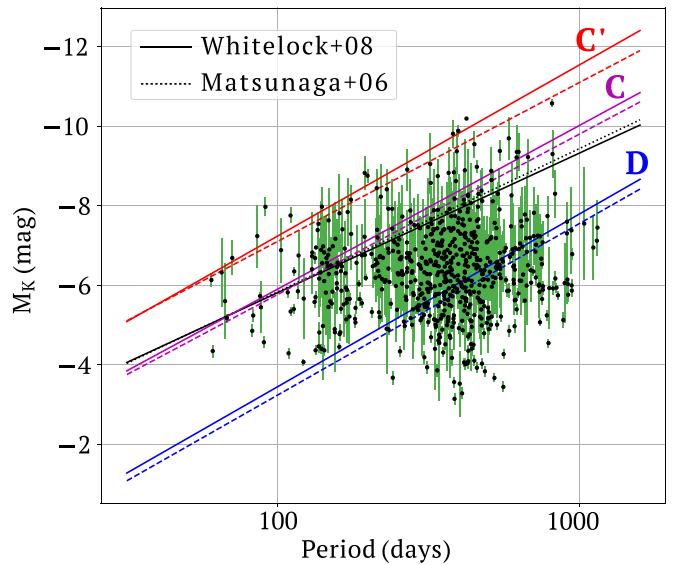
**Figure 13.** Period distributions obtained for variable stars in Gaia DR2 for the Gaia–BAaDE and *local* samples (see Figure 5 and Section 5.6).

of periods available from Gaia for sources in the *local* sample. It can be noted that the whole sample contains LPV stars with a wide range of periods, but that those within the solar neighborhood are restricted to stars with periods of  $\lesssim 1250$  days, presumably Mira or SR variables, as classified by Gaia (Molnár et al. 2018; Mowlavi et al. 2018).

It has been established that P–L relations are a very powerful tool to distinguish AGB stars of different natures (Wood et al. 1999; Ita et al. 2004; Lebzelter et al. 2019). By recognizing 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 (Whitelock et al. 2008). With 2MASS  $K$  magnitudes, Gaia DR2 parallaxes, extinction maps, and periods for a subsample of the *local* sample (*period* sample), we are able to make a comparison of the BAaDE targets with previously studied variable stars.

In Figure 14, 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 the apparent  $K$  magnitude, extinction, and distance (indicated by the error bars) and IR variability (not indicated). A comparison is made of the P–L distributions with known variability sequences, associated with distinct pulsation modes, that have been derived from Gaia DR2 data for LMC 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) and Matsunaga & Team (2006) in the LMC are added.

The *local* sample appears to be much flatter than the empirical P–L relations for the LMC, and it is clear that most of the stars in the *local* sample fall below the LMC P–L relations (Matsunaga & Team 2006; Whitelock et al. 2008). This is likely related to differences between the stellar samples used in Figure 14. We preferentially select the closer, less luminous AGB stars when we make our *local* sample, while in the LMC, the sample is (1) biased toward the most luminous stars and (2) has a different stellar metallicity. Therefore, it is possible that the LMC P–L relations previously related are missing a big, low-luminous clump that we are reporting. In fact, including



**Figure 14.** Period–luminosity relations found for the variable stars within the *local* sample. The sequences marked as C, C', and D represent different known variability sequences, associated with distinct pulsation modes derived for the 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 (Matsunaga & Team 2006; Whitelock et al. 2008).

more initial high-mass sources would add more sources to the upper right of Figure 14 (see, e.g., Vassiliadis & Wood 1993).

Using the analysis by Lebzelter et al. (2019), it is possible to further interpret Figure 14. 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. Both these sequences are supposedly 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. In this sense, we have already confirmed carbon stars in the BAaDE sample based on IR color cuts and detection rates (Lewis et al. 2020a, 2020b), but we plan on analyzing the observational results—including the implications for the P–L diagram—in a subsequent paper. Moreover, at longer periods and lying on sequence D, we typically find stars with mass slightly higher than the solar mass and ages below 1 Gyr (Grady et al. 2019).

## 6. Conclusions

We have cross-matched the BAaDE target list, which consists of 28,062 IR sources mainly preselected from the MSX colors at latitudes  $|b| < 5^\circ$  to match evolved stars in the inner Galaxy (van der Veen & Habing 1988; Sjouwerman et al. 2009), with the Gaia DR2 catalog (Gaia Collaboration et al. 2018b), 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). One-third of the BAaDE target list was not detected in Gaia DR2, these sources correlate with lines of sight of high optical extinction in the Galactic plane. From the 20,111 cross-matched sources, stars with accurate parallax estimates and within a 2 kpc radius around the Sun (where we can obtain accurate extinction maps) were selected after removing YSOs and MS/RGB stars. The remaining 1812 stars constitute



our *local* sample, representing a foreground population of evolved stars in the Galactic plane. Among the *local* sample, the Gaia DR2 shows large amplitude variability for 898 stars that have been classified as Mira variables (Mowlavi et al. 2018), of which another 649 have period estimates.

Using IR and optical data for this sample, we have characterized the evolved stellar population around the Sun in terms of spatial, variability, bolometric, and P-L distributions. The population of evolved stars close to the Sun displays the following features:









1. The absolute magnitude distribution at the *K* band peaks at  $-6.3 \pm 1.2$  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, mainly contains moderate-luminosity variables.
2. Using extinction and BCs from the literature, we are able to estimate bolometric magnitudes for the *local* foreground Galactic sample. The distribution peaks at  $-3.2$  with a width of 1.2 mag ( $1500_{-500}^{+3000} L_{\odot}$ ). This peak is at fainter magnitudes than those obtained for Miras in the LMC (Whitelock et al. 2008) and also at a lower value than inferred for the inner Galaxy (Trapp et al. 2018). Although variability and uncertainties in the extinction and BCs 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. By applying variability filters such as the amplitude index, it is possible to restrict the sample to LPV stars in the solar neighborhood. This filtering has shown to reproduce a narrower *K*-magnitude distribution similar to what we obtained after removing YSOs and MS/RGB stars.
4. Given that our samples are severely affected by interstellar extinction at Gaia wavelengths, we have found that the distribution of moderately luminous evolved stars in the solar neighborhood disk ( $r < 2$  kpc) seem to be not correlated with respect to the location of major Galactic structures in the region. This confirms that the BAaDE target selection traces an old, dynamically relaxed stellar population.
5. 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. Among the targets in the sample, carbon-rich LPV stars also seem to be abundant.

Overall, we conclude that the BAaDE targets are—as discussed—predominantly made up of LPVs, optically detectable Miras, and carbon stars. The IR selection also picks up lower-luminosity objects within 2 kpc from the Sun. The sample of evolved stars at these distances is mostly 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 combined with metallicity information would be required. In a subsequent paper, we will analyze the kinematics of the *local* sample by using a preliminary catalog of  $\sim 17,000$  sources, which corresponds to  $\sim 70\%$  of the VLA and  $\sim 20\%$  of

the ALMA targets that have been already observed and analyzed as part of the BAaDE collaboration.

The authors sincerely thank the anonymous referee for making valuable suggestions that genuinely improved the paper. The BAaDE project is funded by National Science Foundation grant 1517970/1518271. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. M.C.S. is funded by the Heising-Simons Foundation under grant #2018-0911. M.O.L. received support for this work by the NSF through the Grote Reber Fellowship Program administered by Associated Universities, Inc./National Radio Astronomy Observatory. This research has 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.

### ORCID iDs

Luis Henry Quiroga-Nuñez  <https://orcid.org/0000-0002-9390-955X>  
 Huib Jan van Langevelde  <https://orcid.org/0000-0002-0230-5946>  
 Loránt O. Sjouwerman  <https://orcid.org/0000-0003-3096-3062>  
 Ylva M. Pihlström  <https://orcid.org/0000-0003-0615-1785>  
 Anthony G. A. Brown  <https://orcid.org/0000-0002-7419-9679>  
 R. Michael Rich  <https://orcid.org/0000-0003-0427-8387>  
 Michael C. Stroh  <https://orcid.org/0000-0002-3019-4577>  
 Megan O. Lewis  <https://orcid.org/0000-0002-8069-8060>

### References

- Alves, J., Zucker, C., Goodman, A. A., et al. 2020, *Natur*, 578, 237  
 Andrae, R., Fouesneau, M., Creevey, O., et al. 2018, *A&A*, 616, A8  
 Antoja, T., Helmi, A., Romero-Gómez, M., et al. 2018, *Natur*, 561, 360  
 Babusiaux, C., & Gilmore, G. 2005, *MNRAS*, 358, 1309  
 Bailer-Jones, C. A. L. 2015, *PASP*, 127, 994  
 Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, *AJ*, 156, 58  
 Belokurov, V., Erkal, D., Deason, A. J., et al. 2017, *MNRAS*, 466, 4711  
 Belokurov, V., Sanders, J. L., Fattahi, A., et al. 2020, *MNRAS*, 494, 3880  
 Bland-Hawthorn, J., & Gerhard, O. 2016, *ARA&A*, 54, 529  
 Bland-Hawthorn, J., Sharma, S., Tepper-García, T., et al. 2019, *MNRAS*, 486, 1167  
 Blitz, L., & Spergel, D. N. 1991, *ApJ*, 379, 631  
 Capitanio, L., Lallement, R., Vergely, J. L., Elyajouri, M., & Monreal-Ibero, A. 2017, *A&A*, 606, A65  
 Chan, V. C., & Bovy, J. 2020, *MNRAS*, 493, 4367  
 Chiavassa, A., Freytag, B., & Schultheis, M. 2018, *A&A*, 617, L1  
 Colom, P., Lekht, E. E., Pashchenko, M. I., & Rudnitskij, G. M. 2015, *A&A*, 575, A49  
 Dwek, E., Arendt, R. G., Hauser, M. G., et al. 1995, *ApJ*, 445, 716

- Egan, M., Price, S., Kraemer, K., et al. 2003, The Midcourse Space Experiment Point Source Catalog Version 2.3, Tech. Rep. AFRL-VS-TR-2003-1589, Air Force Research Laboratory, [https://irsa.ipac.caltech.edu/data/MSX/docs/msxpsc2.3\\_explguide.pdf](https://irsa.ipac.caltech.edu/data/MSX/docs/msxpsc2.3_explguide.pdf)
- Fish, V. L., Briske, W. F., & Sjouwerman, L. O. 2006, *ApJ*, **647**, 418
- Gaia Collaboration, Babusiaux, C., Van Leeuwen, F., et al. 2018a, *A&A*, **616**, A10
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018b, *A&A*, **616**, A1
- Gaia Collaboration, Eyer, L., Rimoldini, L., et al. 2019, *A&A*, **623**, A110
- Gómez, F. A., Minchev, I., O’Shea, B. W., et al. 2012, *MNRAS*, **423**, 3727
- Gorski, M. D., & Barmby, P. 2020, *MNRAS*, **495**, 726
- Grady, J., Belokurov, V., & Evans, N. W. 2019, *MNRAS*, **483**, 3022
- Groenewegen, M. A. T. 2018, *A&A*, **619**, A8
- Groenewegen, M. A. T., & Sloan, G. C. 2018, *A&A*, **609**, A114
- Habing, H. J. 1996, *A&A*, **7**, 97
- Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, *Natur*, **563**, 85
- Höfner, S., & Olofsson, H. 2018, *A&ARv*, **26**, 1
- Ibata, R. A., McConnachie, A., Cuillandre, J.-C., et al. 2017, *ApJ*, **848**, 129
- Ita, Y., Tanabé, T., Matsunaga, N., et al. 2004, *MNRAS*, **347**, 720
- Jackson, T., Ivezić, Ž., & Knapp, G. R. 2002, *MNRAS*, **337**, 749
- Kounkel, M., Covey, K., & Stassun, K. G. 2020, arXiv:2004.07261
- Kunder, A., Koch, A., Michael Rich, R., et al. 2012, *AJ*, **143**, 57
- Lallement, R., Babusiaux, C., Vergely, J. L., et al. 2019, *A&A*, **625**, A135
- Lebzelter, T., Trabucchi, M., Mowlavi, N., et al. 2019, *A&A*, **631**, A24
- Lewis, M. O., Pihlström, Y. M., Sjouwerman, L. O., et al. 2020b, *ApJ*, **892**, 52
- Lewis, M. O., Pihlström, Y. M., Sjouwerman, L. O., & Quiroga-Núñez, L. H. 2020a, *ApJ*, **901**, 98
- Lian, J., Zhu, Q., Kong, X., & He, J. 2014, *A&A*, **564**, A84
- Lindgren, L., Hernández, J., Bombrun, A., et al. 2018, *A&A*, **616**, A2
- Luri, X., Brown, A. G., Sarro, L. M., et al. 2018, *A&A*, **616**, A9
- Mackereth, J. T., Bovy, J., & Schiavon, R. P. 2017, *MNRAS*, **471**, 3057
- Marrese, P. M., Marinoni, S., Fabrizio, M., & Altavilla, G. 2019, *A&A*, **621**, 144
- Martínez-Medina, L. A., Pichardo, B., Peimbert, A., & Carigi, L. 2017, *MNRAS*, **468**, 3615
- Matsunaga, N., & Team, T. I. 2006, in ASP Conf. Ser. 378, Why Galaxies Care About AGB Stars: Their Importance as Actors and Probes, ed. F. Kerschbaum et al. (San Francisco, CA: ASP), 86
- Messineo, M. 2004, PhD thesis, Leiden Univ, <https://local.strw.leidenuniv.nl/events/phdtheses/messineo/>
- Messineo, M., Habing, H. J., Sjouwerman, L. O., Omont, A., & Menten, K. M. 2018, *A&A*, **619**, 35
- Molnár, L., Plachy, E., Juhász, Á L., & Rimoldini, L. 2018, *A&A*, **620**, A127
- Mowlavi, N., Lecoœur-Taïbi, I., Lebzelter, T., et al. 2018, *A&A*, **618**, A58
- Oort, J. H., Kerr, F., & Westerhout, G. 1958, *MNRAS*, **118**, 379
- Pihlström, Y. M., Sjouwerman, L. O., Claussen, M. J., et al. 2018, *ApJ*, **868**, 72
- Price, S. D. 1995, *SSRv*, **74**, 81
- Quiroga-Núñez, L. H., Immer, K., Van Langevelde, H. J., Reid, M. J., & Burns, R. A. 2019, *A&A*, **625**, A70
- Quiroga-Núñez, L. H., van Langevelde, H. J., Reid, M. J., & Green, J. A. 2017, *A&A*, **604**, A72
- Reid, M., & Honma, M. 2014, *ARA&A*, **52**, 339
- Reid, M., Menten, K., Brunthaler, A., et al. 2014, *ApJ*, **783**, 130
- Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2019, *ApJ*, **885**, 131
- Rich, R. M., Origlia, L., & Valenti, E. 2007, *ApJL*, **665**, L119
- Riess, A. G., Casertano, S., Yuan, W., et al. 2018, *ApJ*, **861**, 126
- Salaris, M., Weiss, A., Cassarà, L. P., Piovani, L., & Chiosi, C. 2014, *A&A*, **565**, 9
- Sande, M. V. D., Decin, L., Lombaert, R., et al. 2018, *A&A*, **609**, 63
- Sevenster, M. N., Van Langevelde, H. J., Moody, R. A., et al. 2001, *A&A*, **366**, 481
- Sjouwerman, L. O., Capen, S. M., & Claussen, M. J. 2009, *ApJ*, **705**, 1554
- Sjouwerman, L. O., Pihlström, Y. M., Rich, R. M., Claussen, M. J., & Morris, M. R. 2017, in IAU Symp. 336, Astrophysical Masers: Unlocking the Mysteries of the Universe (Cambridge: Cambridge Univ. Press), 180
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *ApJ*, **131**, 1163
- Srinivasan, S., Meixner, M., Leitherer, C., et al. 2009, *ApJ*, **137**, 4810
- Stroh, M. C., Pihlström, Y. M., Sjouwerman, L. O., et al. 2019, *ApJS*, **244**, 25
- Trapp, A. C., Rich, R. M., Morris, M. R., et al. 2018, *ApJ*, **861**, 75
- Urago, R., Omodaka, T., Nagayama, T., et al. 2020, *ApJ*, **891**, 50
- van der Veen, W., & Habing, H. 1988, *A&A*, **194**, 125
- Van Langevelde, H., Quiroga-Núñez, L. H., Vlemmings, W. H. T., et al. 2018, in 14th European VLBI Network Symp. Users Meeting (EVN 2018), 43, (doi:/10.22323/1.344.0043)
- van Langevelde, H., & Vlemmings, W. H. T. 2003, in Mass-Losing Pulsating Stars and their Circumstellar Matter. Astrophysics and Space Science Library, Vol. 283, ed. Y. Nakada, M. Honma, & M. Seki (Springer: Dordrecht), 381
- Vassiliadis, E., & Wood, P. R. 1993, *ApJ*, **413**, 641
- Wegg, C., & Gerhard, O. 2013, *MNRAS*, **435**, 1874
- Whitelock, P. 2003, in Mass-losing Pulsating Stars and their Circumstellar Matter Workshop, Astrophysics and Space Science Library, Vol. 283, ed. Y. Nakada, M. Honma, & M. Seki (Dordrecht: Kluwer), 19
- Whitelock, P. A., Feast, M. W., & Van Leeuwen, F. 2008, *MNRAS*, **386**, 313
- Wood, P. R., Alcock, C., Allsman, R. A., et al. 1999, in IAU Symp. 191, Asymptotic Giant Branch Stars, ed. T. Le Bertre, A. Lebre, & C. Waelkens (Cambridge: Cambridge Univ. Press), 632
- Zinn, J. C., Pinsonneault, M. H., Huber, D., & Stello, D. 2019, *ApJ*, **878**, 136