

The Three-Dimensional Distribution of Galactic AGB Stars with ALLWISE

Nicholas M. Hunt-Walker, Željko Ivezić, Andrew C. Becker

University of Washington, Department of Astronomy, Seattle, WA 98195

`nmhw@uw.edu, ivezic@uw.edu, acbecker@uw.edu`

1. Introduction

The structure of the Milky Way holds clues to the processes of formation and evolution of galaxies. Historical models typically assumed three discrete components described by simple analytic expressions: the thin disk, thick disk, and halo (???). Recent surveys, such as the Sloan Digital Sky Survey (SDSS, ?) and the Two Micron All Sky Survey (2MASS, ?), have provided much more detail about these components. For example, SDSS data have constrained stellar distributions in the 7-dimensional space spanned by spatial coordinates (?), velocity components (?), and metallicity (?). The resulting maps revealed rich, complex substructure in the distribution of the Milky Way's stars (e.g. ?????????), deeply shaking the older view of a smooth Galaxy.

The large distance limit of about 100 kpc for some halo populations (e.g. RR Lyrae and BHB stars with SDSS, XXX Sesar 2010 ref and red giants with 2MASS, XXX Majewski ref), as well as about 10 kpc for very numerous main sequence stars, is not reached when studying disk component at low galactic latitudes. As shown, for example, by Berry et al., the extinction due to interstellar dust limits the SDSS sample of main sequence stars to heliocentric distances of only a few kpc. In order to avoid detrimental effect of dust extinction, stellar samples need to be probed at longer infrared wavelengths. Several infrared surveys covering the Galactic plane have recently become available (gives refs for WISE and GLIMPSE, VVV: arXiv:1111.5511).

Among populations suitable for studying Galactic structure using infrared surveys, Asymptotic Giant Branch (AGB) stars stand out. AGB stars represent the last stage of evolution for stars between 0.8 and 8 M_{\odot} (??). Stars from this mass range can reach the final stages of stellar evolution within the Galactic timescale (~ 10 Gyr, ?) and thus are bound to reside throughout the Galaxy wherever other stars are present. This phase of stellar evolution is marked by two distinct episodes with different observational characteristics: the early AGB phase (E-AGB) and the thermally-pulsing AGB phase (TP). During the thermally-pulsing phase, AGB stars produce substantial dust-driven stellar winds ($10^{-7} < \dot{M} < 10^{-4} M_{\odot} \text{ yr}^{-1}$, ?) rich in oxides (SiO, Al₂O₃, etc.) or carbon-rich molecules (SiC, AmC, etc.) The dominant chemical composition is highly dependent upon the metallicity of the host galaxy (?). High-metallicity galaxies like the Milky Way have a substantial population of oxygen-rich AGB stars (?), whereas low-metallicity galaxies such as the Magellanic clouds are dominated by carbon-rich AGB stars (?). In both cases, the other species of AGB star is rarely seen, as richness in one chemical type (e.g. oxides) necessitates the almost complete capture of the other chemical type (e.g. carbon) in CO (?).

The dust-rich winds in TP phase create vast circumstellar shells that are warmed by the stellar photosphere and emit copiously in the near- and mid-infrared (NIR & MIR, respectively). Together with high bolometric luminosity (10^3 – $10^4 M_\odot$), this redistribution of the output radiation to the infrared wavelength makes AGB stars excellent disk probe when infrared survey data are available. Indeed, they were detected all the way to the Galactic center even with the IRAS survey (?). Such disk studies can now be significantly improved thanks to the much more sensitive WISE survey.

The *Wide-field Infrared Survey Explorer* (WISE, ??), is a space observatory that has imaged essentially the entire sky in the MIR (3.4, 4.6, 12, and 22 μm). WISE catalog has been positionally matched to the 2MASS catalog, with the matched catalog listing NIR and MIR 7-band photometry for hundreds of millions of sources. Given the depths of the two surveys, this catalog should contain AGB stars to many kpc beyond the Galactic center. Here we develop selection methods for AGB stars using WISE and 2MASS data, and analyse the resulting samples.

In Section ??, we describe in detail the WISE, 2MASS and other auxiliary data used in our study and the data reduction process. In Section ??, we use samples of known Galactic and Magellanic AGB stars to derive WISE-2MASS color-based selection criteria, and calibrate color-absolute magnitude relations. In Section ??, we describe the spatial density distribution of selected AGB candidates from the Milky Way. Our conclusions are summarized in Section ??.

2. Input Catalogs and Data Preparation

Our principal data source is the merged WISE-2MASS catalog. In order to derive WISE-2MASS color-based selection criteria, we use several training samples described below: a heterogeneous sample of bright AGB stars listed in SIMBAD database and more homogeneous fainter samples selected using optical time-domain surveys OGLE-III and MACHO. The latter samples include AGB stars from the LMC and SMC and are also used to calibrate color-absolute magnitude relations that underlie our distance estimates. In order to assess what infrared populations could contribute to contamination of selected AGB candidates, we utilize a number of SDSS catalogs, also described below. We conclude this section by describing how these catalogs auxiliary catalogs were positionally merged with the WISE-2MASS catalog.

2.1. WISE-2MASS catalog

Write a few sentences about the WISE survey...

2.2. AGB stars from SIMBAD database

The SIMBAD database¹ is comprised of basic data, classifications, and measurements for all types of astronomical objects from a variety of surveys. Classifications are hierarchical, and are classified with increasing specificity based on the available data, though each star can have multiple classifications. AGB stars have subcategories of C-stars ($[C/O] > 1$) and S-stars ($[C/O] \sim 1$). Where subclassification is unclear, the main classification remains as AGB star. OH/IR stars are subsets of peculiar stars that specifically show strong OH maser emission and strong IR emission, identifying them clearly as AGB stars. Mira-type variables are a subset of variable stars, and are classified based upon their periodicity. Semi-regular variables are not included as they contain an unknown mixture of AGB stars and red supergiants. The sample of AGB stars from SIMBAD was obtained by querying all objects with types classified as C-stars (23,628), S-stars (1,514), AGB stars (4,459), OH/IR stars (1,265), and Mira-type variables (10,828). Without duplicate objects between the five types, there is a total of 39,348 AGB stars from SIMBAD.

2.3. AGB stars from optical time-domain surveys

Many of the identifications in the SIMBAD catalog are from spectroscopic studies, however many more AGB stars have been found through their characteristic optical periodicity. The *Optical Gravitational Lens Experiment* (OGLE) and the *MAssive Compacy Halo* (MACHO) survey have are two time-domain surveys that have captured and catalogued such variability.

2.3.1. MACHO

2.3.2. OGLE-III

2.4. Imaging and spectroscopic catalogs from SDSS

A few sentences of introduction...

¹<http://simbad.u-strasbg.fr/simbad/>

2.4.1. Galactic sources

2.4.2. Extragalactic populations

2.5. Merged sample

Each AGB sample as well as the SDSS sources were matched separately to the ALLWISE data products via the GATOR tool at the NASA/IPAC Infrared Science Archive². The accepted matching radius was 3". The matched sample was further reduced by enforcing the WISE 5 σ faint limits of $[W1/W2/W3/W4] < [16.83/15.6/11.32/8.0]$, saturation limits of $[W1/W2/W3/W4] > [2.0/1.5/-3.0/-4.0]$, signal-to-noise ratios greater than 3 in each WISE band. We constrain the possibility for a mismatched extragalactic IR source by asserting that the extended source flag ≤ 1 , corresponding either to a point source (flag = 0) or a source whose profile-fit photometry is goodness-of-fit is > 3 in any WISE band but is not coincident with any extended sources from the 2MASS survey. We also require only one match to a 2MASS source within 3".

How was it merged, which further cuts and why...

2.6. Completeness and contamination of the initial AGB sample

In the next section you are using the AGB sample from this section to estimate completeness and contamination of color-selected samples. Here you need to summarize what do you believe completeness and contamination are for your initial sample constructed using SIMBAD and optical surveys. In particular, it is important to describe expected distance limits.

2.7. XXX old stuff from here to the end of this section: Data Sources

In this study, we rely heavily on data from the ALLWISE extension of the WISE survey, combining data from the initial All-Sky Data Release, the 3-band cryogenic data release, and the NEOWISE post-cryogenic data release (?). The initial WISE All-Sky Data Release observed the sky between January and August 2010, observing the sky 1.2 times with four detectors, operating at 3.4, 4.6, 12, and 22 μ m. Hereon we refer to ALLWISE photometric bands at $[3.4\mu\text{m}/4.6\mu\text{m}/12\mu\text{m}/22\mu\text{m}]$ as $[W1/W2/W3/W4]$. The positions of objects in the WISE catalog were calibrated to the 2MASS point source catalog. The 3-band cryogenic data release contains data from $W1$, 2, and 3, and surveyed 30% of the sky between August and October 2010. During the 3-band cryogenic survey, $W1$ and $W2$ operated with nearly the same sensitivity as during the full survey. Warming of the telescope reduced sensitivity in $W3$ and fully saturated $W4$. The NEOWISE post-cryogenic data release contains $W1$ and $W2$ measurements, with sensitivities close

²<http://irsa.ipac.caltech.edu/cgi-bin/Gator/nph-scan?mission=irsa&submit=Select&projshort=WISE>

to those obtained during the full cryogenic phase. During this phase, WISE surveyed 70% of the sky. Data products from the post-cryogenic release included updated instrumental, astrometric, and photometric calibrations and reduction algorithms, resulting in much lower SNR. The overall number of sources compiled into ALLWISE totals over 747.6 million.

In order to generate a reliable, high-confidence catalog of Galactic candidate AGB stars, we must first define color-color criteria from known AGB star samples. We select AGB stars from three source catalogs: the *Optical Gravitational Lens Experiment-III Variable Star Catalog* (OGLE-III, ???), the *MAssive Compact Halo Objects* project (MACHO, ?), and the SIMBAD Astronomical Database (?).

OGLE-III photometry for Long-Period Variables (LPVs) in the Small and Large Magellanic Clouds (SMC and LMC respectively) was obtained between July 2001 and May 2009, with stars in the central 4.5-deg² of the LMC and SMC having an additional 5 observing seasons of photometry from OGLE-II. LPVs were classified into 3 categories: OGLE-III Small Amplitude Red Giants (OS-ARGs), Semi-Regular Variables (SRVs), and Miras. All AGB stars O-rich and C-rich AGB stars in OGLE-III were photometrically selected using reddening-free Wesenheit magnitudes, described in detail in ??). [Describe the selection bit in a little more detail, along with their sample completeness, selection biases, and contamination fractions] Data reduction techniques are described in ?). The resulting samples yield 46,467 AGB stars from the LMC (37,203 O-rich; 9,264 C-rich) and 6,509 stars from the SMC (3,727 O-rich; 2,782 C-rich).

From MACHO we obtain the sample of SMC, LMC, and Galactic Bulge AGB stars used in ?) (14,861 stars). Why were these objects selected? Howe were they selected? What is their completeness, selection bias, and contamination fraction? Following ?), the objects are divided into sequences (seq) 1-4. Sequence 1 primarily contains Mira variables pulsating in their fundamental modes, whereas Sequences 2-4 contain semi-regular variables in various pulsation modes.

The sample of AGB stars from SIMBAD was obtained by querying all objects classified as C-stars (18,656), S-stars (1,108), OH/IR stars (825), AGB stars (2,359), and Mira variables (9,608), for a total of 32,556 stars. Objects are classified spectroscopically, though by a variety of methods owing to the heterogeneous data housed within SIMBAD. Together with MACHO and OGLE-III, the total sample of AGB stars is 100,393. Because there is a high likelihood that samples between OGLE-III, MACHO, and SIMBAD overlap, we retain only unique objects after the initial data reduction in section ??.

We use SDSS spectroscopic catalogs to find and quantify regions in NIR-MIR color-color space populated by plausible contaminant sources. These include any Galactic stellar objects and planetary nebulae, as well as a host of extragalactic sources. Data for active galactic nuclei (AGN; 19,184 objects), quasi-stellar objects (QSOs; 122,550 objects), and star forming/burst galaxies (820,272 objects total) were drawn from SDSS DR7, specifically from the NYU Value Added Galaxy

Catalog³ (VAGC). Luminous Red Galaxies (LRGs) were selected from the SDSS Luminous Red Galaxy Survey (105,631 objects, VAGC). Data for stars in the SDSS stellar locus were drawn from the DR 9 SEGUE Stellar Parameters Pipeline (SSPP) (1,843,190 objects, SEGUE). **Include bit about YSOs and PNe from SIMBAD**

2.8. Data Reduction

We use NASA/IPAC IRSA’s **GATOR** tool⁴ to positionally match SDSS, OGLE-III, MACHO, and SIMBAD to ALLWISE. We select only matches within 3” between each sample and ALLWISE. All samples of AGB were required to be brighter than the published 5σ faint limits of [16.83/15.6/11.32/8.0], as well as fainter than the saturation limits of [2.0/1.5/-3.0/-4.0] extrapolated from the wings of the PSFs for point sources, for $[W1/W2/W3/W4]$ (VAGC), with no flags for confusion or contamination as a spurious source in any band. We also require only single associations with 2MASS objects within 3”, detections in $[J/K_S/W1/W2/W3/W4]$, and $\text{SNR} > 3$ in each ALLWISE band.

The population for each sample from initial matching as well as after the application of the ALLWISE faint limits, saturation limits, and 2MASS detection requirements are shown in Table ???. The WISE color-color distributions for the AGB and contaminant samples are shown in Figure ??.

Population	SIMBAD AGB*	C*	Mira	OH/IR	S*
3” match	1,689	14,209	9,027	406	1,081
Reduced	684	1,782	3,241	43	511
Population	MACHO seq1	seq2	seq3	seq4	
3” match	5,279	3,519	2,619	3,070	
Reduced	277	185	73	61	
Population	OGLE-III C-rich	O-rich			
3” match	11,542	38,848			
Reduced	249	730			
Population	DR12 SSPP	DR7 LRG	QSO	AGN	Galaxies
3” match	1,578,329	104,345	103,590	18,528	841,712
Reduced	67,508	84	3,977	1,069	44,314

Table 1: AGB and contaminant populations matched to WISE before and after sample reduction in section ???. MACHO sequences (seq1-seq4) are from (VAGC) and described briefly in section ??.

³<http://sdss.physics.nyu.edu/vagc/>

⁴<http://irsa.ipac.caltech.edu/cgi-bin/Gator/nph-scan?mission=irsa&submit=Select&projshort=WISE>

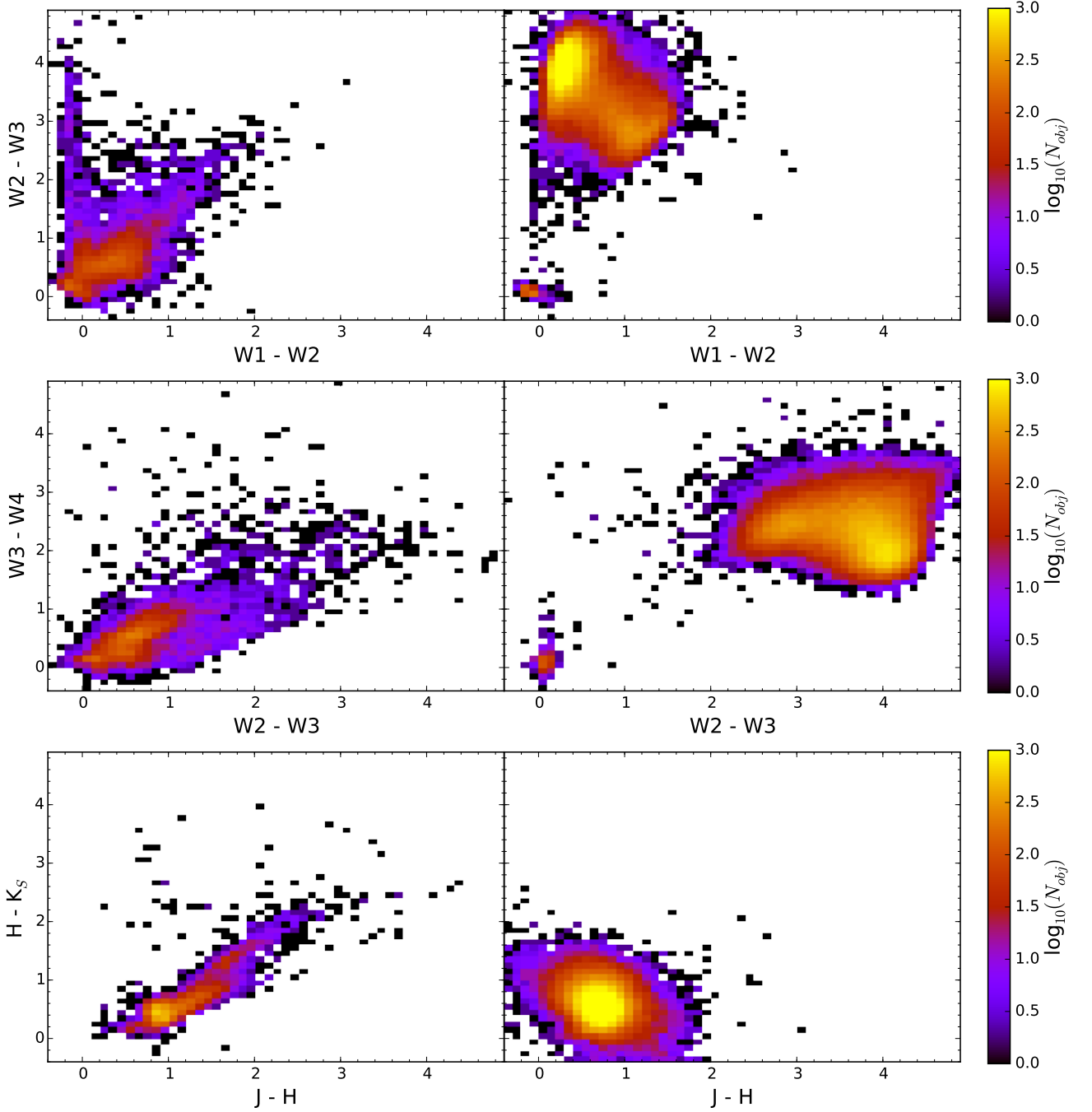


Fig. 1.— Logarithmic number densities for objects in WISE and 2MASS color-color space, binned in 0.1 dex on each axis. *Left:* The combined AGB sample matched to ALLWISE. *Right:* The combined contaminant sample.

3. Object Selection Criteria

In creating color-color criteria to generate a catalog of AGB candidates, we seek to maximize AGB completeness while minimizing contamination from non-AGB objects to beneath the 1% level.

The color-color criteria for AGB selection are as follows:

$$(J - K_s) > 1.1 \quad (1)$$

$$(W2 - W3) > 0.3 \quad (2)$$

$$(W3 - W4) < -0.83(W2 - W3) + 3.37 \quad (3)$$

The criteria in (??) and (??) are primarily concerned with rejecting objects from the stellar locus, and other objects whose NIR spectra are dominated by the Rayleigh-Jeans tail. (??) also rejects stars from the stellar locus, but primarily functions to remove IR-bright extragalactic sources.

We also experiment with other criteria to focus more closely on the high-reliability AGB population, instead of trying to capture the most AGB-like objects. These criteria are as follows:

$$(W1 - W2) < 1 \quad (4)$$

$$(W2 - W3) < 1 \quad (5)$$

Criteria (??) restricts NIR excesses. Criteria (??) primarily removes extragalactic sources, while only slightly cutting into the distribution of known AGB stars.

Sample completeness η is defined as

$$\eta = \frac{N - n_{\text{missed}}}{N}$$

where N is the total number of objects in the sample, and n_{missed} is the number of objects outside of the applied boundaries. Figure ?? shows the distribution of sample completeness amongst AGB sources after the application of the above criteria. The vast majority (79.07%) of Galactic AGB stars is recovered after criteria (??), (??), and (??) are applied. The largest losses occur at the edges of the Galactic disk ($|b| \approx 10^\circ$) and in regions of high stellar number density both in the Galaxy and the Magellanic Clouds. In the color-color diagrams, selection completeness degrades near the borders of the selection area, as objects that straddle these boundaries may exhibit enough emission in other color-color spaces to be removed by our criteria. Of the remaining sample of 5,709 objects, 52.94% lie within the disk ($|b| < 10^\circ$).

?) defines sample contamination as

$$\epsilon = \frac{n_{\text{spurious}}}{n_{\text{selected}}}.$$

where n_{spurious} is the number of spurious sources and $n_{\text{selected}} = N + n_{\text{spurious}}$. The contamination map is shown in Figure ??. Most bins in Figure ?? exhibit 0% contamination, with the overall

Table 2: Samples Recovered with Application of Criteria (%)

Object Type	Reduced	(1)	(2)	(3)	(1,2)	(1,2,3)	—	(4)	(5)	(1,2,4)	(1,2,5)	(1,2,4,5)
All AGBs	7220	95.15	84.89	95.21	82.62	79.07	—	94.22	88.82	77.05	72.71	71.14
O-rich AGB	3147	98.98	100.00	100.00	98.98	98.98	—	98.44	97.33	97.43	96.31	94.88
C-rich AGB	540	99.26	99.81	99.26	99.07	98.33	—	57.59	63.33	56.85	62.59	54.44
Unclassed AGB	3533	91.11	69.15	90.32	65.53	58.39	—	96.07	85.14	61.99	53.24	52.53
DR12 SSPP	67508	76.16	99.16	0.91	76.07	0.05	—	86.05	0.89	65.67	0.03	0.03
DR7 LRG	84	96.43	100.00	0.00	96.43	0.00	—	89.29	0.00	85.71	0.00	0.00
QSO	3783	73.54	99.97	0.13	73.54	0.11	—	34.50	0.11	27.54	0.08	0.08
Galaxy	42066	78.55	100.00	0.01	78.54	0.00	—	96.39	0.01	75.31	0.00	0.00
AGN	1011	79.33	100.00	0.00	79.33	0.00	—	96.34	0.00	75.87	0.00	0.00

contamination level at 0.38%. What contaminants do remain exist primarily at the very fringes of the AGB star distribution in $(J - K_S)$ vs $(W2 - W3)$ space and near the redder boundary in $(W3 - W4)$ vs $(W2 - W3)$, where bluer extragalactic sources creep into the selection region. The results of the applied criteria on both the collective AGB and contaminant samples are summarized in Table ??.

We note that while we do achieve a low rate of contamination, most of our contaminant sources are out of the plane of the Galaxy, while many of our AGB sources are within the Galactic disk. Considering how our extragalactic contaminant sources are from various iterations of SDSS, we know that their spatial distributions should be effectively uniform across the sky [cite me]. For contaminants unaffected by interstellar reddening, they would occupy the same regions of color-color space already marked by our existing contaminant sample, and would similarly be removed by (??), (??), and (??). This would produce the same zero-level contamination for all extragalactic sources aside from QSOs. We can adopt a QSO spatial density by dividing the total number of spectroscopically identified QSOs by the covered area on the sky. Using this, we recover a QSO spatial density of **some number per square degree**. Using that density, and assuming ubiquity of sources on the sky, complete detection of QSOs **down to some magnitude limit**, and a lack of interstellar reddening, we can estimate a maximum QSO contamination fraction of **some number**. We use this number to report our QSO contamination fraction in Table ??, as QSOs significantly affected by interstellar reddening would exist further outside of the limits of (??), (??), and (??). A more rigorous study of QSO spatial density is outside the scope of this paper.

The remaining contaminant sample is from the SDSS DR12 SSPP, with the number of remaining species found within shown in Figure ?. **the figures don't match up with each other. The number of red stars on the right side of the fig are significantly less than those in the histogram. Reconcile this.**

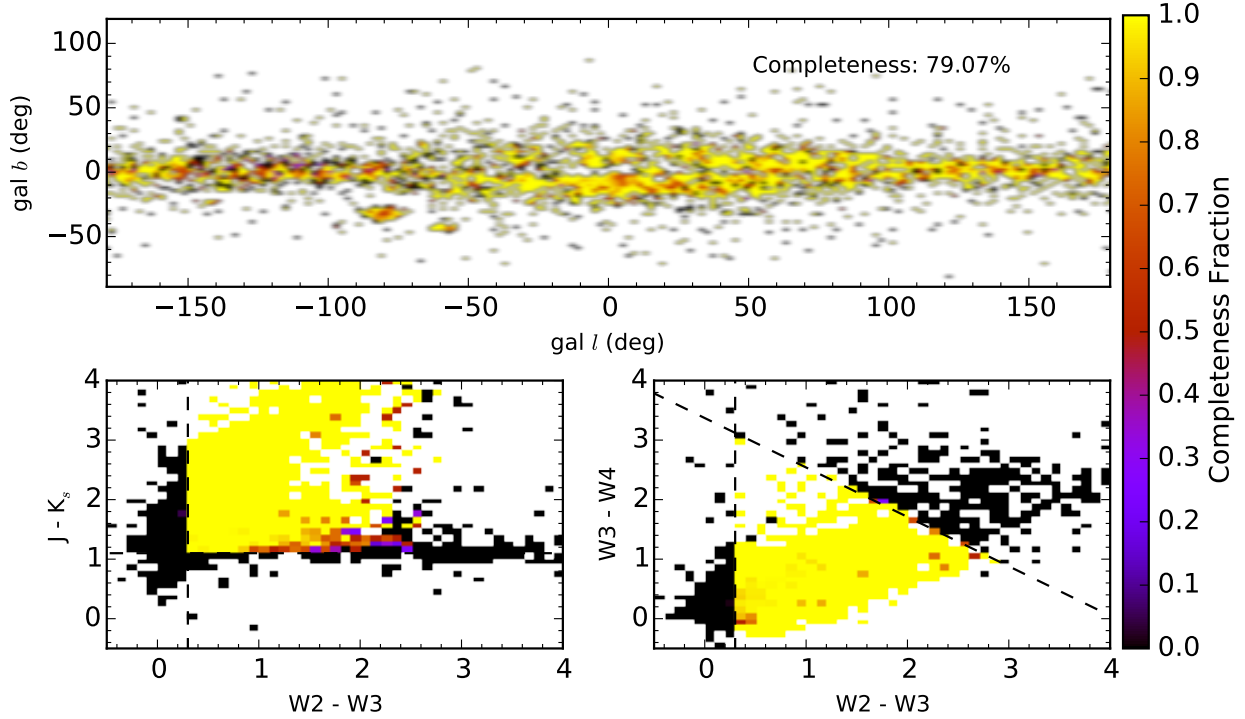


Fig. 2.— Sample selection completeness maps in (l, b) space (*top*) and the color-color space of our selection criteria (*bottom*). Color scale shows the completeness fraction per bin, with 4-deg^2 bins on the Galactic map and 0.1 dex bins on each axis of the color-color diagrams. Selection criteria are shown as dashed lines.

4. AGB Candidate Distribution

Put words here

5. Conclusions

Put words here

REFERENCES

- Ahn, C. P., et al. 2012, *Ap. J. Suppl.*, 203, 21
- Alcock, C., et al. 1997, *Ap. J.*, 482, 89
- Bahcall, J. N., & Soneira, R. M. 1980, *Ap. J. Suppl.*, 44, 73
- Belokurov, V., et al. 2006, *Ap. J. (Letters)*, 642, L137

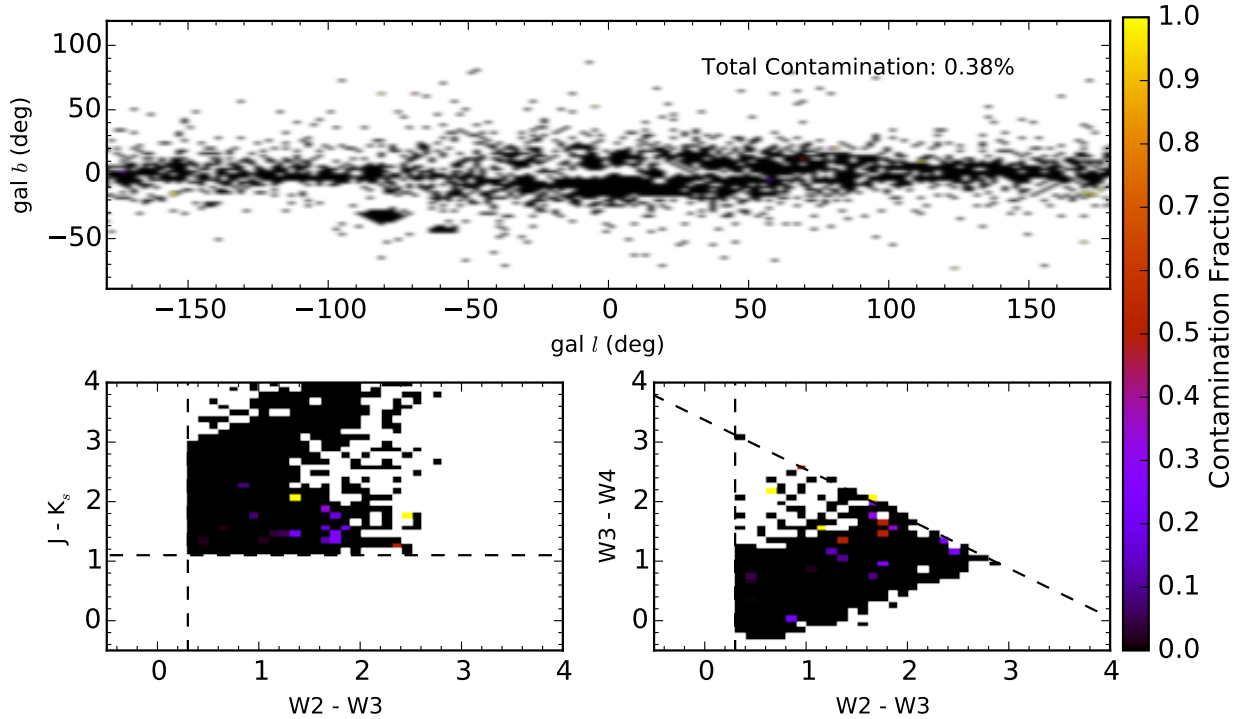


Fig. 3.— Sample contamination in Galactic (l, b) space (*top*) and color-color space (*bottom*). Boundaries and binning the same as Fig ??.

Blanton, M. R., et al. 2005, *A. J.*, 129, 2562

Bond, N. A., et al. 2010, *Ap. J.*, 716, 1

Boyer, M. L., et al. 2011, *A. J.*, 142, 103

Cutri, R. M., et al. 2012, Explanatory Supplement to the WISE All-Sky Data Release Products, Tech. rep.

—. 2013, Explanatory Supplement to the AllWISE Data Release Products, Tech. rep.

Fraser, O. J., Hawley, S. L., & Cook, K. H. 2008, *A. J.*, 136, 1242

Gilmore, G., Wyse, R. F. G., & Kuijken, K. 1989, *Ann. Rev. Astr. Ap.*, 27, 555

Grillmair, C. J. 2006, *Ap. J. (Letters)*, 651, L29

Habing, H. J., Olmon, F. M., Chester, T., Gillett, F., & Rowan-Robinson, M. 1985, *Astr. Ap.*, 152, L1

Herwig, F. 2005, *Ann. Rev. Astr. Ap.*, 43, 435

Iben, Jr., I., & Renzini, A. 1983, *Ann. Rev. Astr. Ap.*, 21, 271

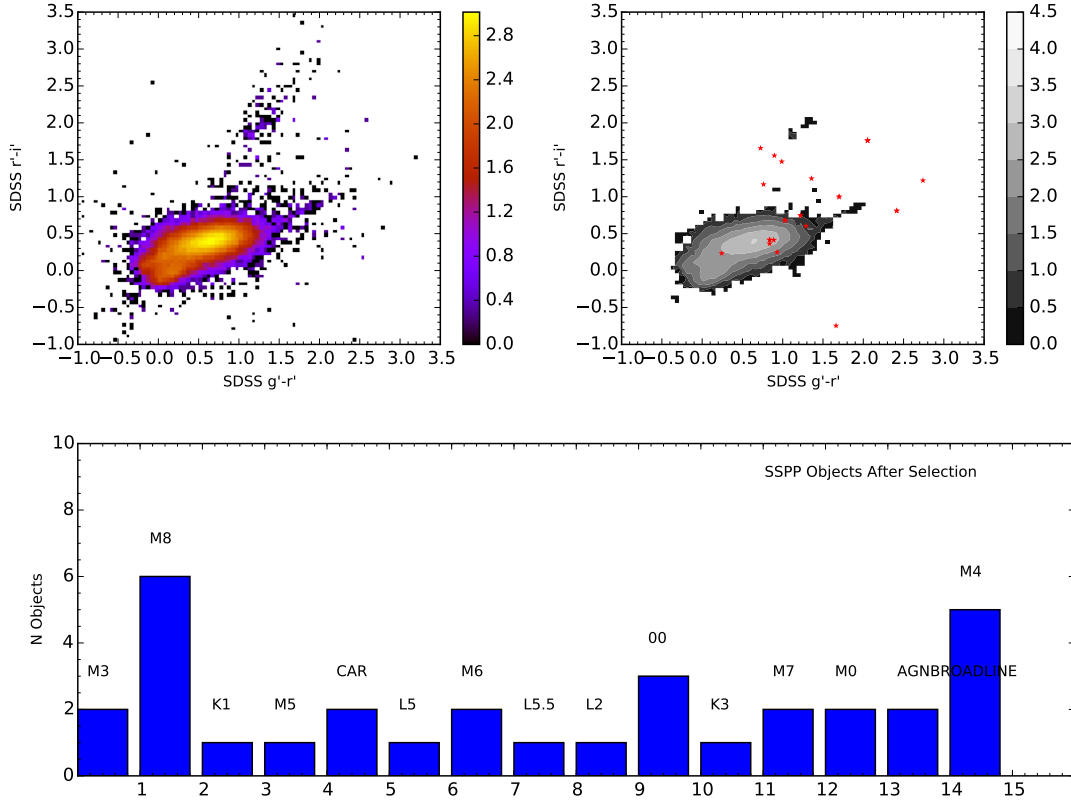


Fig. 4.— Some histogram

Ivezić, Ž., Connolly, A., VanderPlas, J., & Gray, A. 2013, *Statistics, Data Mining, and Machine Learning in Astronomy*

Ivezić, Ž., et al. 2000, *A. J.*, 120, 963

—. 2008, *Ap. J.*, 684, 287

Jackson, T., Ivezić, Ž., & Knapp, G. R. 2002, *M.N.R.A.S.*, 337, 749

Jurić, M., et al. 2008, *Ap. J.*, 673, 864

Kazin, E. A., et al. 2010, *Ap. J.*, 710, 1444

Majewski, S. R. 1993, *Ann. Rev. Astr. Ap.*, 31, 575

Majewski, S. R., Skrutskie, M. F., Weinberg, M. D., & Ostheimer, J. C. 2003, *Ap. J.*, 599, 1082

Matsuura, M., et al. 2005, *Astr. Ap.*, 434, 691

Newberg, H. J., et al. 2002, *Ap. J.*, 569, 245

Table 3: Sample Selection Completeness and Contamination

Population	SIMBAD AGB*	C*	Mira	OH/IR	S*
Completeness	89.62%	72.11%	95.62%	39.53%	22.31%
Population	MACHO seq1	seq2	seq3	seq4	
Completeness	88.45%	81.08%	28.77%	14.75%	
Population	OGLE-III C-rich	O-rich	All AGB Stars		
Completeness	73.09%	70.68%	79.07%		
Population	DR12 SSPP	DR7 LRG	Galaxies	QSO	AGN
Contamination	0.56%	0.00%	0.00%	0.07%	0.00%

Olofsson, H., González Delgado, D., Kerschbaum, F., & Schöier, F. L. 2002, *Astr. Ap.*, 391, 1053

Skrutskie, M. F., et al. 2006, *A. J.*, 131, 1163

Soszyński, I., et al. 2009, *Acta Astronomica*, 59, 239

—. 2011, *Acta Astronomica*, 61, 217

Udalski, A., Szymanski, M. K., Soszynski, I., & Poleski, R. 2008, *Acta Astronomica*, 58, 69

Vivas, A. K., & Zinn, R. 2006, *A. J.*, 132, 714

Vivas, A. K., et al. 2001, *Ap. J. (Letters)*, 554, L33

Wenger, M., et al. 2000, *Astr. Ap. Suppl.*, 143, 9

Wright, E. L., et al. 2010, *A. J.*, 140, 1868

Yanny, B., et al. 2000, *Ap. J.*, 540, 825

York, D. G., et al. 2000, *A. J.*, 120, 1579