QUASAR SELECTION TECHNIQUES GOING INTO THE GAIA ERA

N. SECREST[1], J. FROUARD[2], G. HENNESSY[3]

U.S. Naval Observatory - USA
[1]nathan.j.secrest.civ@mail.mil, [2]julien.h.frouard.civ@navy.mil, [3]gregory.s.hennessy.civ@mail.mil

ABSTRACT. In the last five years, there has been an explosion in the number of known active galactic nuclei (AGN) and quasars, thanks largely to the *Wide-field Infrared Survey Explorer* (WISE). This has been a major boon for celestial reference frame (CRF) work, which requires a quasi-inertial reference frame by which positions and proper motions can be defined and self-consistently measured. In this work, we describe new developments in the pursuit of a maximally dense, maximally uniform all-sky sample of AGN and quasars, as well as how the improved CRF may in turn benefit quasar work.

1. WHY QUASARS?

How is motion defined? When discussing the positions of astronomical objects, we must be precise about exactly with respect to *what* we are defining these positions, which determines how independently we can measure any associated motion that we observe. When we observe an astronomical object, our measurement of its apparent motion is affected by the Earth's rotation about its axis, the Earth's orbit around the Sun, the Sun's motion through the Galaxy (the secular aberration), the Galaxy's motion with respect to the Local Group, the Local Group's motion induced by the inhomogeneous distribution of local matter in the universe (the "clustering dipole" in cosmology parlance), and finally any inhomogeneity that may exist in the distribution and motion of matter at the largest observable scales. Fortunately, the success of ΛCDM at describing the universe on the largest ($> 100h^{-1}$ Mpc) scales suggests that this latter term is negligible (although see Colin et al. 2017; Rameez et al. 2018; and references therein), and so distant extragalactic objects are the best, and likely final, choice to realize a quasi-inertial reference system, i.e. the International Celestial Reference System (ICRS; e.g., Arias et al. 1995). The most distant (and therefore most ideal) extragalactic objects, quasars, are fortunately also powerful sources of compact radio emission, allowing us to take advantage of VLBI's exquisite astrometric precision both to produce a realization of the ICRS, the International Celestial Reference Frame (Ma et al. 1998; Fey et al. 2015; Charlot et al. 2020) and to tie it to the International Terrestrial Reference Frame (ITRF; e.g., Altamimi et al. 2016).

2. FINDING QUASARS

While efforts to realize the ICRS have been successful, several issues remain, such as a relative under-density of sources below declinations of $\sim -30°$, despite efforts to increase the uniformity of the *defining* sources (Figure 1), caused by relatively few southern baselines in the global VLBI network. A related issue, which is the focus of this work, is the effect of defining a fundamental CRF at one wavelength (i.e., in the radio S and X bands) on the tie to CRFs at other wavelengths. This issue has become especially pressing with the advent of *Gaia*, which for the first time has produces a CRF at visual (optical) wavelengths comparable in precision to VLBI measurements. Alignment of the *Gaia* reference frame to the ICRS was done with a subset of 2843 sources identified as optical counterparts to the ICRF3 prototype (Lindegren et al. 2018), but to ensure that this aligned CRF is non-rotating a much larger, all-sky, and uniform catalog of quasars is preferred. As faint stars and quasars are often photometrically indistinguishable in the optical, spectroscopic confirmation of the presence of redshifted broad Balmer and high-ionization narrow emission lines has generally been required in the absence of a strong radio counterpart, which has so far been too observationally expensive to produce a proper all-sky sample as needed. The launch of the mid-infrared *Wide-field Infrared Survey Explorer* (WISE; Wright et al. 2010) in 2009 provided a serendipitous solution to this problem. In the mid-infrared, AGN and quasars occupy a distinct region of photometric color-color space (i.e., [3.4 μm]$-$[4.6 μm], [4.6 μm]$-$[12 μm] for WISE) because of their power-law spectral energy distributions, and these colors are nearly insensitive to heavy obscuration (e.g., Donley et al. 2012). Using the WISE color-color selection wedge of Mateos et al. (2012), Secrest et al. (2015) produced an all-sky sample of 1.4 million AGN, which they showed to contain few to no star contaminants that would be detectable by *Gaia*. This sample of AGN, which has half a million *Gaia* counterparts, after applying a few astrometric quality cuts, was subsequently used for the *Gaia*-CRF (Lindegren et al. 2018).

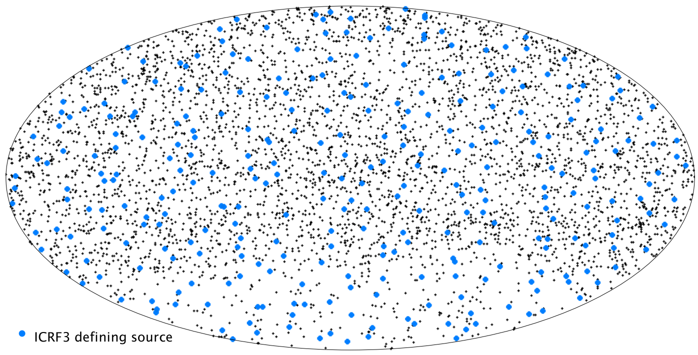

Figure 1: Aitoff projection of ICRF3, in equatorial coordinates.

While WISE data has been hugely beneficial for CRF work, there are two problems that we discuss here. The first is non-uniformity in the distribution of WISE sources across the sky (see Figure 1 in Secrest et al. 2015; Figure 2 in this work). This is caused by an over-density of sources at the ecliptic poles due to the WISE scanning pattern, which is easily remedied with a simple magnitude cut, and an under-density of sources along the Galactic plane due to source confusion. This latter issue cannot be remedied using WISE data alone, and so we have begun exploring other methods of filling in the decrement of sources along the Galactic plane.

It is not possible, even in principle, to find AGN with *Gaia* counterparts lying behind much of the Galactic plane, as no optical counterpart exists for sight lines with $E(B-V) \gtrsim 2$. However, a significant portion of the Galactic plane in the zone of WISE confusion has sight lines with $E(B-V) < 2$, suggesting that it is worthwhile to look at alternate methods to identify AGN with *Gaia* counterparts (Figure 2). While mid-infrared color-based diagnostics are the ideal method to photometrically identify AGN and quasars, the near-infrared, which carries the advantage of much higher angular resolution ($< 1''$ for facilities such as UKIRT, versus $\sim 6''$ to $8''$ for WISE), shows promise. Moreover, the proposed *GaiaNIR* mission (e.g., Hobbs et al. 2019) will operate at roughly the near-infrared H-band, where the extinction coefficient $A/E(B-V)$ is a factor of ~ 4 times smaller than in the *Gaia* G band (e.g., Yuan et al. 2013), allowing background AGN counterpart identification to $E(B-V) \lesssim 8$, freeing up significantly more of the Galactic plane (Figure 2).

We have developed a machine learning-based method to identify likely AGN candidates along the Galactic plane that folds in prior information, specifically the Galactic extinction, with optical photometry from *Gaia* and near-infrared photometry from the UKIDSS Galactic Plane Survey (GPS;

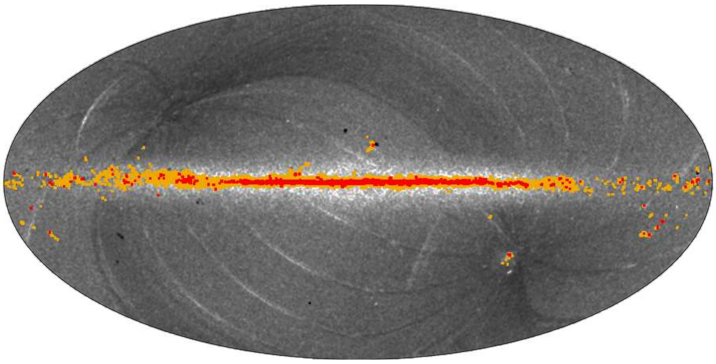

Figure 2: Aitoff projection of AGN from Secrest et al. (2015), shaded by density. Gold regions denote $E(B-V) > 2$, where *Gaia* is insensitive to background AGN, while red regions denote $E(B-V) > 8$, where *GaiaNIR* will be insensitive to background AGN.

Lucas et al. 2008). Using `scikit-learn`,[1] we train a k-NN classifier with AGN selected from the AllWISE catalog at high Galactic longitude, where source confusion is less of an issue, in order to sample a similar parameter space in $E(B-V)$ and representative stellar populations. We use photometry from *Gaia* DR2 and the UKIDSS GPS, and $E(B-V)$ from Schlafly & Finkbeiner (2011). By validating against our training sample, we found reliability and completeness values of $\sim 70\%$ and $\sim 30\%$. Given that the percentage of AGN in the training set is only 0.45%, this means that our classifier is performing ~ 160 times better than random chance. In order to validate our classifier, we have an accepted program to spectroscopically confirm our AGN candidates in the near-IR, where optical strong emission lines should exist given the typical redshifts of mid-IR AGN, using the UIST instrument on UKIRT, although logistical problems have prevented this program from being executed thus far. We show a sample of our Galactic plane AGN candidates in Figure 3.

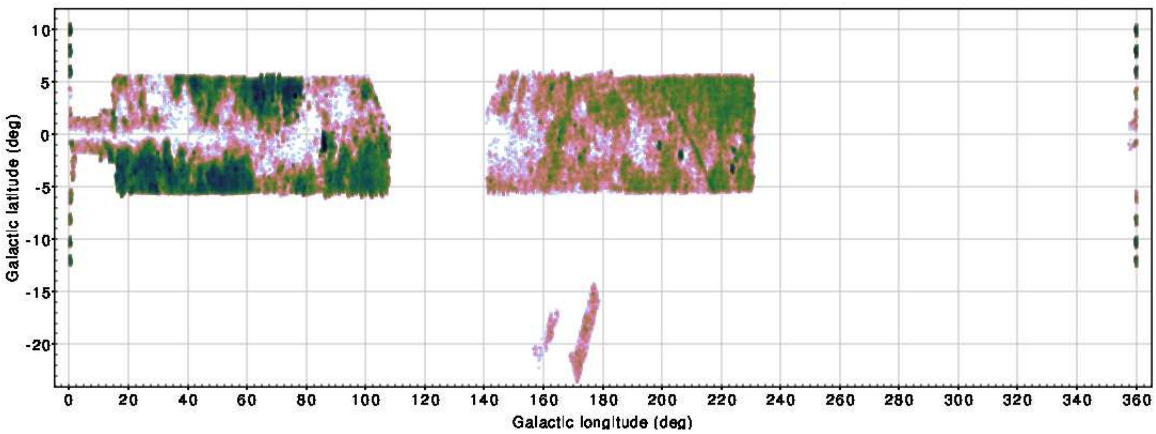

Figure 3: Cutout of the Galactic plane, with darker shading indicating a higher density of AGN candidates (arbitrary scaling).

3. RECIPROCITY

We have shown how massive photometric catalogs have been and may continue to be used to assist with CRF work in the *Gaia* era. This benefit is not one-way, however, and CRF work provides

[1] `https://scikit-learn.org/stable`.

reciprocal benefit to photometric catalogs in the form of refined astrometric reference catalogs. In the case of WISE and *Gaia*, we have begun re-deriving the astrometry of the AllWISE catalog, as well as the deeper unWISE catalog (Schlafly et al. 2019), using astrometry from *Gaia*. The WISE catalog astrometry is based on 2MASS (Skrutskie et al. 2006) with proper motion corrections from UCAC4 (Zacharias et al. 2006) [2] that, while of excellent astrometric quality, do nonetheless exhibit residual position-dependent astrometric offsets (zonal errors), which could potentially be alleviated by using *Gaia* data (e.g., Spoto et al. 2017). In Figure 4, we show the signed ($\pm$) astrometric offsets (in mas) between reference sources in the AllWISE catalog and their *Gaia* counterparts, using the default catalog AllWISE positions. In Figure 5, we show the offsets using our re-derived positions.

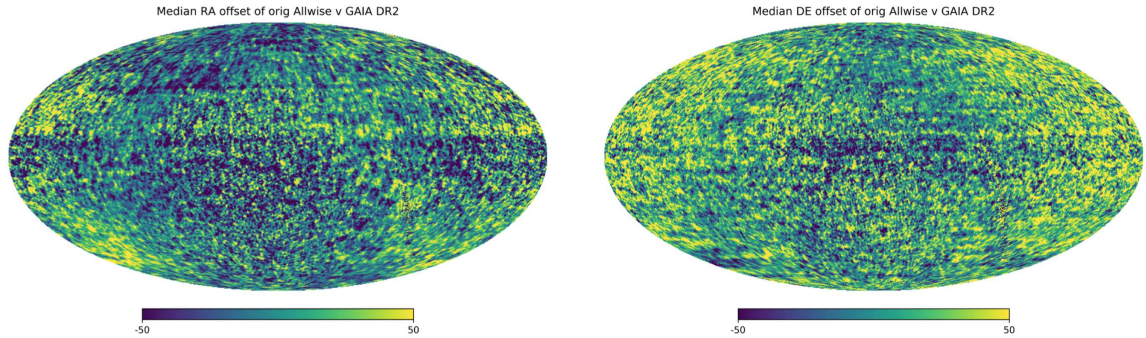

Figure 4: Offsets (in mas) in R.A. and Decl. between AllWISE and *Gaia* counterparts using default AllWISE catalog source positions.

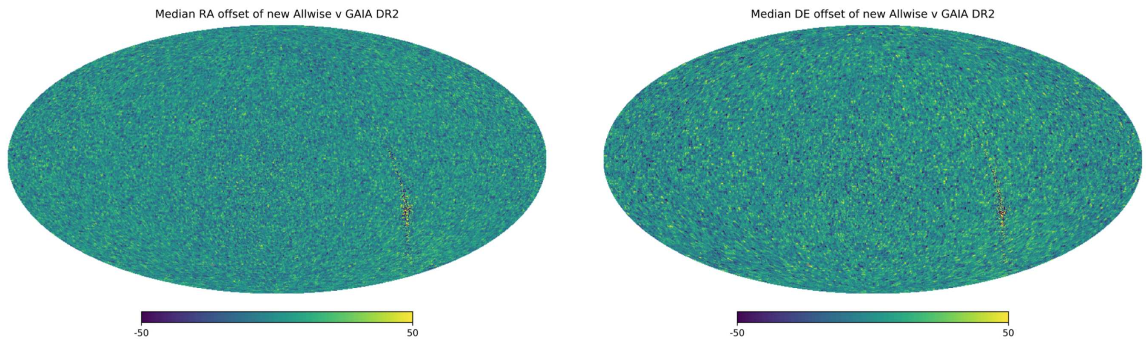

Figure 5: Offsets (in mas) in R.A. and Decl. between AllWISE and *Gaia* counterparts using re-derived AllWISE catalog source positions.

As can be seen, the new version of the AllWISE catalog is now on the *Gaia* coordinate system after using *Gaia* as the astrometric reference catalog. This essentially de-biases the WISE data with respect to position, allowing for large, all-sky studies using catalog data unaffected by positional systematics. We emphasize that these results are preliminary, and we are currently working on further validating them. Nonetheless, with the impressive depth of WISE data (over 2 billion objects in the latest unWISE release; Schlafly et al. 2019) and its wide-ranging utility, from quasar and AGN studies to studies of brown dwarfs and asteroids, efforts to produce optimized astrometry are of utmost importance. Serendipitously, in the preparation of this work the new CatWISE catalog (Eisenhardt et al. 2020) was released,[3] which uses *Gaia* for astrometric validation, further underscoring the reciprocity between these two missions.

[2] `http://wise2.ipac.caltech.edu/docs/release/allwise/expsup/sec2_5.html`.

[3] `https://catwise.github.io`.

4. REFERENCES

Altamimi, Z., Rebischung, P., Métivier, L., et al., 2016, Journal of Geophysical Research (Solid Earth), 121, 6109

Arias, E. F., Charlot, P., Feissel, M., et al., 1995, A&A , 303, 604

Colin, J., Mohayaee, R., Rameez, M., et al., 2017, MNRAS , 471, 1045

Charlot, P., et al., 2020, "The third realization of the International Celestial Reference Frame by very long baseline interferometry", A&A , doi: 10.1051/0004-6361/202038368.

Donley, J. L., Koekemoer, A. M., Brusa, M., et al., 2012, ApJ , 748, 142

Eisenhardt, P. R., Marocco, F., Fowler, J. W., et al., 2020, ApJS , 247, 69

Fey, A. L., Gordon, D., Jacobs, C. S., et al., 2015, AJ , 150, 58

Hobbs, D., Brown, A., Høg, E., et al., 2019, arXiv e-prints, arXiv:1907.12535

Lindegren, L., Hernández, J., Bombrun, A., et al., 2018, A&A , 616, A2

Lucas, P. W., Hoare, M. G., Longmore, A., et al., 2008, MNRAS , 391, 136

Ma, C., Arias, E. F., Eubanks, T. M., et al., 1998, AJ , 116, 516

Mateos, S., Alonso-Herrero, A., Carrera, F. J., et al., 2012, MNRAS , 426, 3271

Rameez, M., Mohayaee, R., Sarkar, S., et al., 2018, MNRAS , 477, 1772

Schlafly, E. F., & Finkbeiner, D. P., 2011, ApJ , 737, 103

Schlafly, E. F., Meisner, A. M., & Green, G. M., 2019, ApJS , 240, 30

Secrest, N. J., Dudik, R. P., Dorland, B. N., et al., 2015, ApJS , 221, 12

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al., 2006, AJ , 131, 1163

Spoto, F., Tanga, P., Bouquillon, S., et al., 2017, A&A , 607, A21

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al., 2010, AJ , 140, 1868

Yuan, H. B., Liu, X. W., & Xiang, M. S., 2013, MNRAS , 430, 2188

Zacharias, N., Finch, C. T., Girard, T. M., et al., 2013, AJ , 145, 44