MW-Gaia STSM

Catalogue of reference QSOs within 3° of M31

Charis Tsakonas from University of Athens (Greece) spent one month (08/01-08/02/2023) at the Paris Observatory to extend and secure QSO candidates in 3 degrees round the centre of M31 using Gaia EDR3.

To begin with, we query Gaia EDR3 for all the sources inside a circle of 5-degree radius from the center of M31. The first step was to clean the data from spurious sources by applying a series of quality cuts. The first filter is to drop all duplicated sources, and sources brighter than G<13.8 mag and fainter than G>20.8 mag, in order to select a region that is populated by QSO candidates (Riello et al. (2021), fig.21). Additionally, we inspect the data for strong deviations in the distribution of the renormalised unit weight error factor and we drop all sources with ruwe>1.4 as well as all sources without color measurements. We also demand the astrometric parameters solved with 5 or 6 parameters equations, so that we have complete measurements for our sources. After applying these quality cuts we end up with 811598 sources, from the initial 4352517 we had after querying Gaia EDR3.

Before proceeding to the next phase we run a series of quality test to inspect our data. We did a useful coordinate transformation following Helmi et al. (2018) and we test that the error in each component of the proper motion is below an acceptable threshold and it behaves as expected (i.e. it diverges for fainter sources).

After acquiring our own sample of sources we cross match it with Gaia table “aux_allwise_agn_gdr2_cross_id”, which is still the least contaminated QSO candidate sample (see the left panel of figure 1) because it is obtained from mid-Infrared data of WISE (Secrest et al. 2015). Gaia EDR3 also provides a new QSO candidate table (gaiadr3.qso_candidates). However, after checking the spatial distribution and color magnitude diagram of the QSO candidate around M31 region, we find that this new QSO table may be significantly contaminated by the M31 blue stars and MW stars (see figure 2). In principle, QSOs in any field of view on the sky should be evenly distributed.

Our ultimate goal is to obtain the most reliable QSO sample to M31 studies, as a step in our aim to measure the proper motion of M31. Thus, we decide to start from the QSO sample of DR2, and, then, carefully select the QSO candidate using EDR3 data.

After analysing figure 1 and 2, we notice that we have to limit the selection of QSO by colours, e.g, between the green and orange line, to avoid contamination from MW stars. We also need to avoid source with color bluer than 0.4 because they are dominated by blue giant stars in M31 disk (this is why EDR3 QSO candidates show a good trace of M31 disk.)

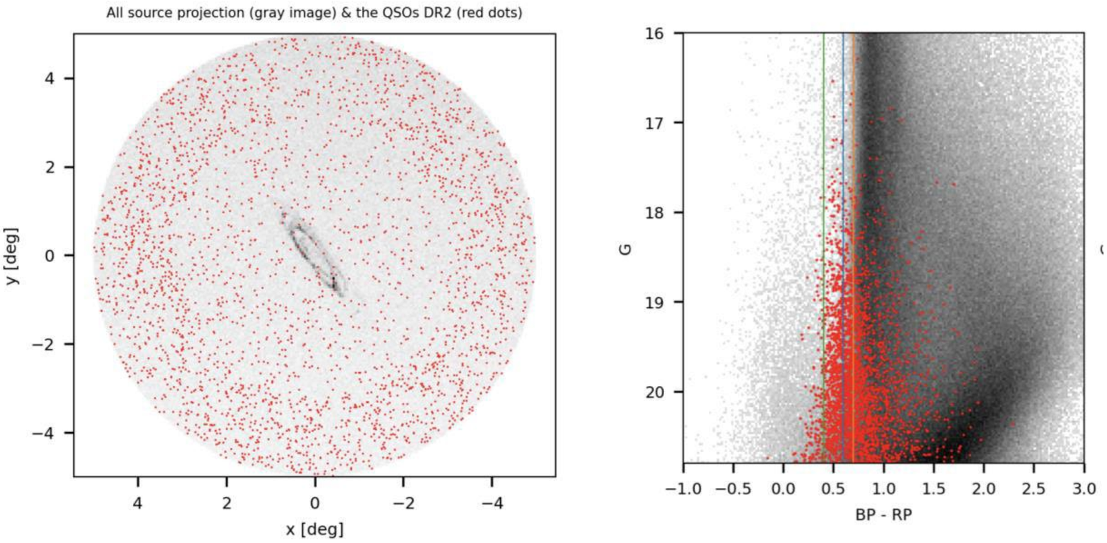


Figure 1: On the left, all sources are scattered with grey dots and QSO in Gaia DR2 are scattered with red dots. On the right, the positions of these QSO on the CMD (in red), overscattered on all sources (in grey).

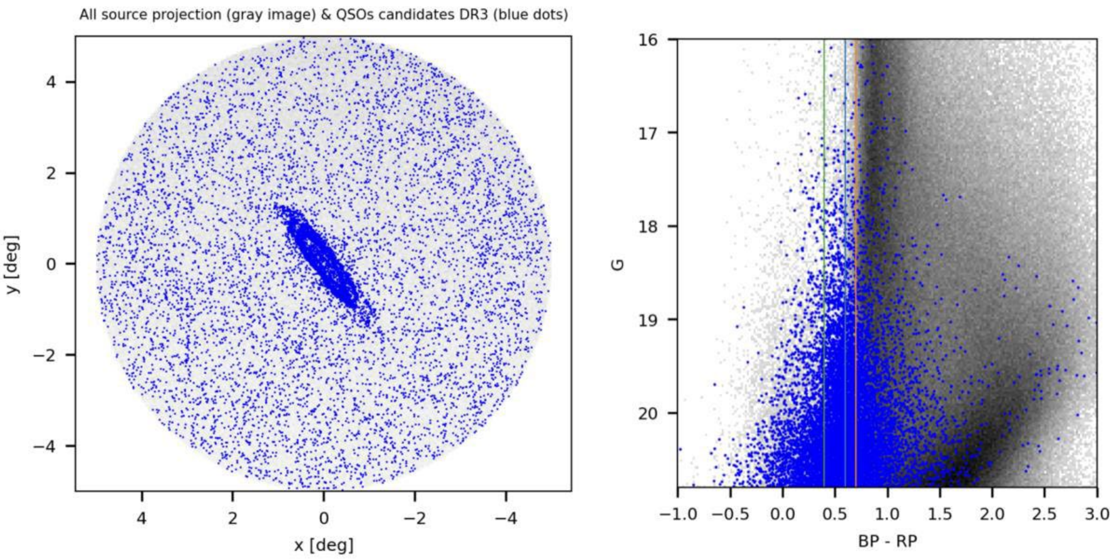


Figure 2: With blue color, scattered are the QSO in Gaia EDR3 catalogue. On the left, scattered, on a 5-degree circle around the center of M31. On the right, all sources in the same region are plotted in grey, and with blue, the QSO candidates on a CMD.

QSO are extragalactic source, their parallax measurements should appear as noise in Gaia data.

According to the above analysis, we did the following further selection of QSOs from Gaia DR2 sample. We utilize the cross match by only keeping sources (as our QSO candidates) that have a distance greater than 0.5 arcsec from any source on “qso_candidates” catalogue in Gaia DR2. So we manage to drop all the sources that are already found in “qso_candidates” catalogue. The region where we expect to find more QSO candidates for our extended sample is in a ring outside of the disc of M31 and inside a circle within a 3-degree radius.

The next step is to try to identify the contamination sources for our sample. To this end, we create two subsamples. The first one is a circle of 1-degree radius around the centre of M31, that probes mostly the disc (and the blue stars inside it). The second one is a ring between 3 and 4 degrees from M31’s centre, that probes a region outside the disc of M31.

We plot the proper motion for both of our subsamples. On the left, is the proper motion for the sources in our sample in the circle of 1 degree from the centre of M31 and on the right the same plot for the (3,4) ring.

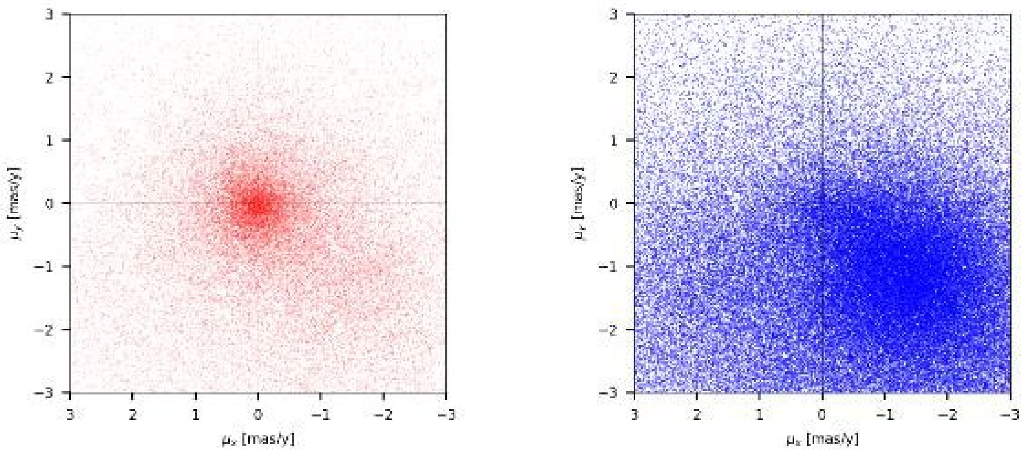
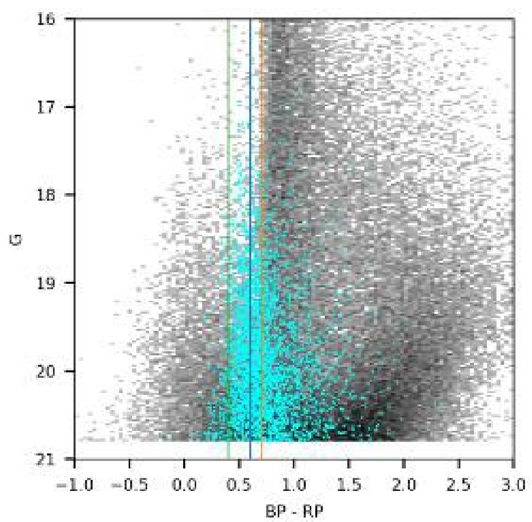


Figure 3: Left: Proper motion of all sources within 1 degree from the centre of M31. Right: Proper motion of all sources within a ring between 3 and 4 degrees from M31’s centre.



We also plot a color magnitude diagram for our subsample inside the 1-degree circle, with the confirmed from Gaia (table “qso_candidates”) scattered in cyan.

Figure 4: All sources (grey points) inside a 1-degree circle from the centre of M31. In cyan all the sources from the table “qso_candidates” of Gaia DR2 data.

What we infer from the above plots is that our sample will be contaminated both from Milky Way stars (the above, right panel for the (3,4) ring clearly shows an overdensity of proper motion typical for Milky Way's stars) and, judging from the CMD, it includes many M31's disc stars. The vertical lines one the CMD will be the cuts we will apply and will be discussed. These figures shape the criteria we have to apply in order to create an augmented catalogue of QSO candidates.

We, then apply the following cuts:

I) Parallax: The absolute parallax to be less than 1/50, so that the distance to the sources is greater than 50kpc. Though this cut, we exclude most of Galactic contamination. Also, we need to account for the error distribution of the parallax measurement, so we also apply parallax_over_error to be less than 3.5.

So our first cut is:
(np.abs(df.parallax) < 1./50) & (df.parallax_over_error < 3.5)

II) **BP-RP and G:** As we can see from the plot of bp-rp versus magnitude, most of the confirmed QSO lie in a region approximately of 0.4<bp-rp<1.0. So, in order to avoid contamination both from Milky Way stars and from M31 disc stars we apply a quality cut in the bp-rp column for our sources, by centering our target at 0.55 and taking all the sources in the region (|bp_rp| - 0.55) < 0.15, as the line in the graph inticate. After some trial and error, we decided to apply another cut on the G magnitude so that our sample suffers less from galactic contamination (Riello et al. (2021), fig.21). We end up only with sources of photometric g mean magnitude greater than 19.5.

References

Riello, M., De Angeli, F., Evans, D. W., Montegriffo, P., Carrasco, J. M., Busso, G., ... & Yoldas, A. (2021). Gaia early data release 3-photometric content and validation. *Astronomy & Astrophysics*, *649*, A3.
Helmi, A., Babusiaux, C., Koppelman, H. H., Massari, D., Veljanoski, J., & Brown, A. G. (2018). The merger that led to the formation of the Milky Way's inner stellar halo and thick disk. Nature, 563(7729), 85-88.
Secrest, N. J., Dudik, R. P., Dorland, B. N., Zacharias, N., Makarov, V., Fey, A., ... & Finch, C. (2015). Identification of 1.4 million active galactic nuclei in the mid-infrared using WISE Data. The Astrophysical Journal Supplement Series, 221(1), 12.

Description of the STSM main achievements and planned follow-up activities

By this procedure we managed to get 213 new QSO candidates that are not in the Gaia catalogue and reside outside of the disc. Their positions on the sky is shown in the figure below.

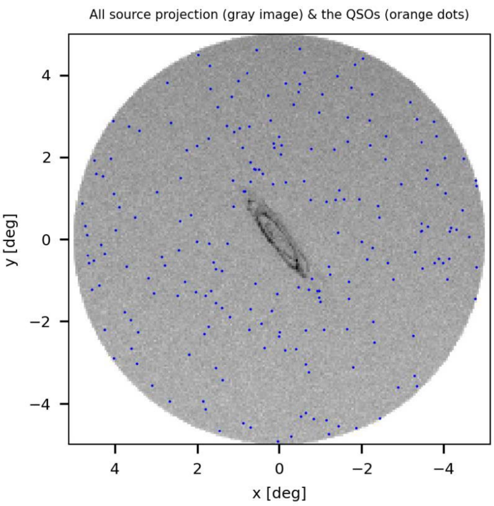


Figure 5: Scattered with blue color, the new QSO candidates we managed to acquire. All sources scattered with grey color.

Their position in the proper motion landscape, scattered in blue, above the confirmed QSO from Gaia in the figure below.

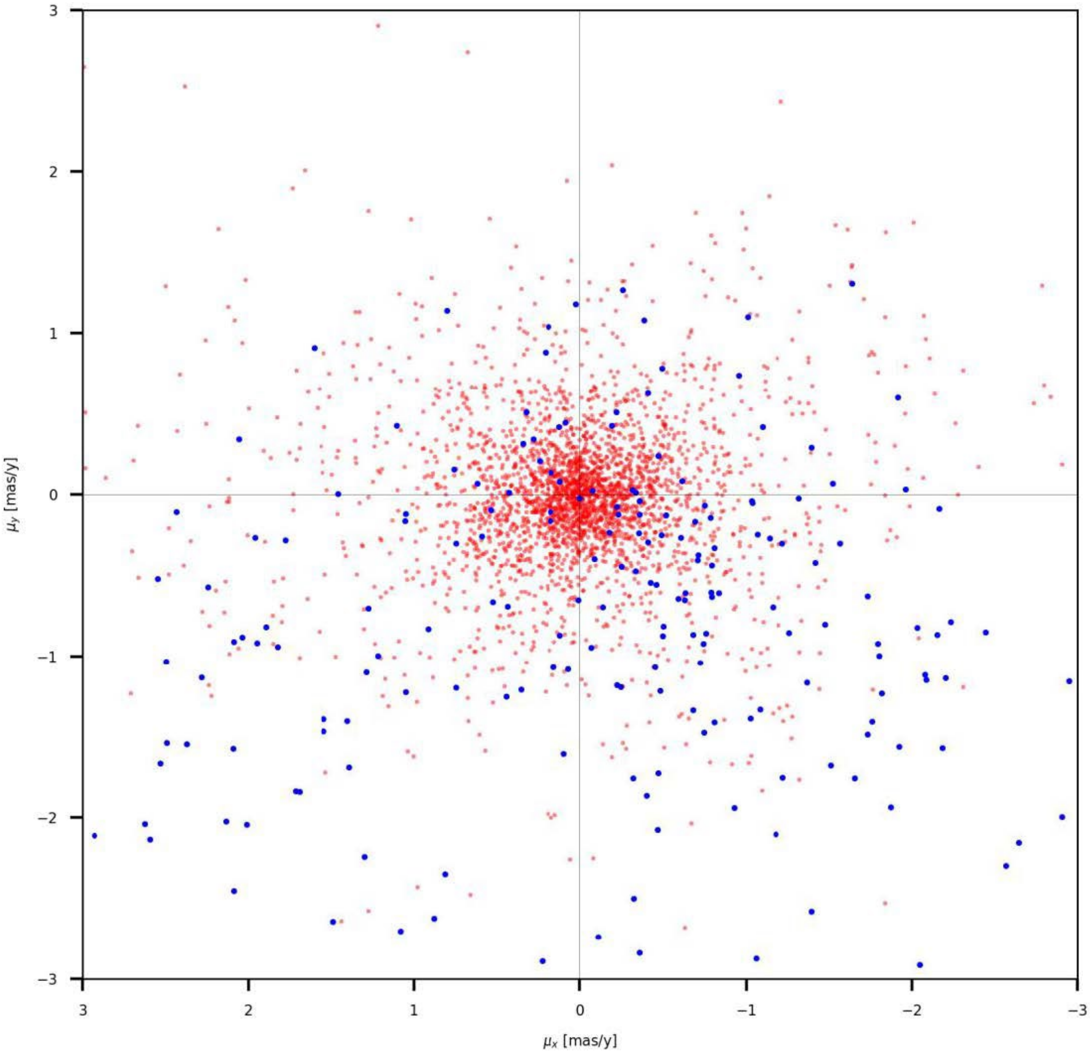


Figure 6: Proper motion of our new candidates, in blue points.

Their position in the CMD, scattered in blue above the complete sample of sources for the same region.

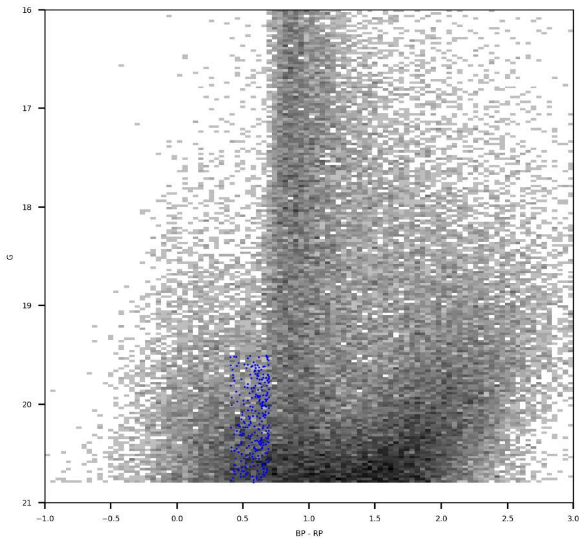


Figure 7: CMD of our new QSO candidates.

Finally, to assess the quality of our work, we plot the following diagram that indicates the locus of the g magnitudes that we managed to find more QSO candidates. The grey histogram shows the QSO candidates from Gaia, while the salmon colored is our enhanced sample. As planned, we managed to populate the catalogue of QSO candidates with more faint sources.

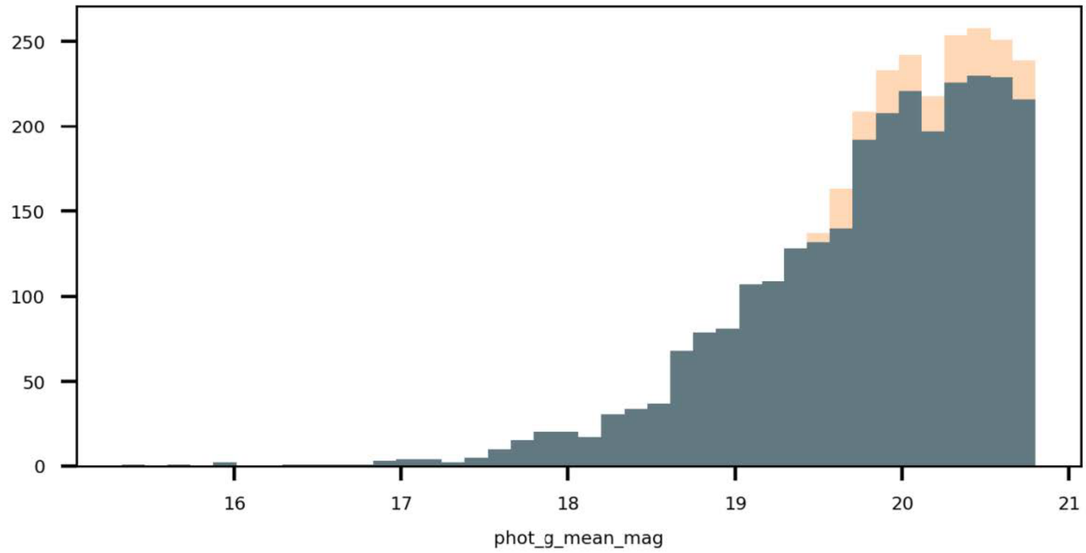


Figure 8: Histogram of the magnitudes of QSO candidates from Gaia DR2 with cyan and of our sample with salmon.

Concluding, the STSM achieved its preliminary goal, to acquire a secure sample of QSO candidates in a region around M31 circle, using, as a first step, Gaia DR2. After we build our sample by carefully selecting sources Gaia EDR3 sample we will have our final sample from Gaia data. The next step will be to fold in the photometric information provided by the u-band measurements of the Canada-France Imaging Survey (CFIS) and the 1 million quasar candidates from Pan-STARRS 1 and use a Deep Learning network to learn the photometric colors and variability of genuine quasars given the full PS1+Gaia+CFHT data set, and propagate that information to all the sources in the region around M31

References

Riello, M., De Angeli, F., Evans, D. W., Montegriffo, P., Carrasco, J. M., Busso, G., ... & Yoldas, A. (2021). Gaia early data release 3-photometric content and validation. *Astronomy & Astrophysics*, *649*, A3.

In a further collaboration with the host (Dr. Francois Hammer), we are working on constraining the formation history of M31 through hydrodynamical, N body simulations. We compare the predictions of simulated analogues with observations of M31 (oxygen abundances in the disc of M31, metallicity measurements in the giant stream).