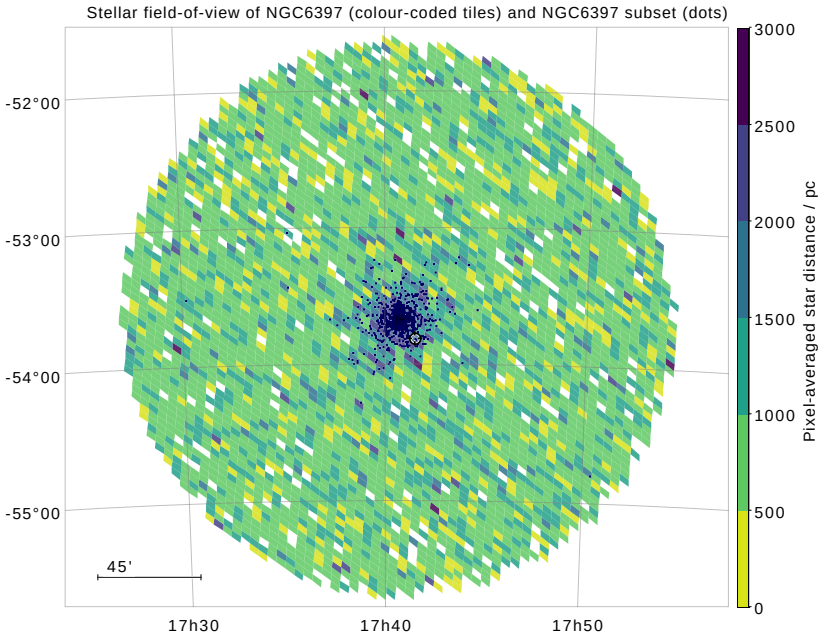
The Old Globular Cluster NGC6397



Geneviève Parmentier & Markus Demleitner

November 24, 2026

Abstract

NGC6397 is the second-closest Galactic globular cluster. Our line-of-sight to NGC6397 is crossing the Galactic disc, and thus its field of view contains a substantial fraction of foreground disc stars. Using Virtual Observatory facilities and in particular the Table Access Protocol (TAP), we find and isolate cluster stars, consider the relevance of errors, create colour-magnitude diagrams and compare them with galactic open clusters. We also introduce the HEALPix sky tesselation scheme and use it to consider extinctions as well as their utility in the context of aggregate functions as defined by the Astronomical Data Query Language ADQL.

Software: TOPCAT

1 Introduction

NGC6397 is an old globular cluster, a part of the Galactic halo. With a mass close to $10^5\,M_\odot$, an age of $\simeq 12$ Gyr and a metal fraction two orders of magnitude lower than that of the Sun, NGC6397 is significantly more massive, older and more metal-poor than the Pleiades open cluster (M45) discussed in the use case on Exploring the Pleiades with TOPCAT and Aladin (Demleitner & Heinl, 2024a), which you are welcome to explore as a prelude to the present document.

It is also much more distant than the Pleiades, which you can check by querying their respective parallaxes from the Simbad astronomical database: Go to `http://simbad.cds.unistra.fr/simbad/` and query for M45 and NGC6397 in turn. You will retrieve parallaxes of something like 7.364 mas for M45 and 0.416 mas for NGC6397. That is, NGC6397 is $\simeq 18$ times more distant from the Sun than the Pleiades.

An immediate consequence of NGC6397's larger distance is that stars of a given absolute magnitude appear fainter to the observer if they belong to NGC6397 than if they belong to the Pleiades. In turn, this implies that NGC6397 stellar photometry is of poorer quality than that of M45. Therefore, this tutorial pays a lot more attention to errors than the Pleiades one.

Our first goal will be to build a colour-magnitude diagram of NGC6397 following the prodedure from Demleitner & Heinl (2024a). We will find that we need more powerful tools in this case, and so we will turn to ADQL and TAP; a reasonably gentle introduction to them can be found in Demleitner & Heinl (2024b).

In contrast to a Cone Search, a TAP query allows one to finely control the parameters of the selected stars on the server side, without ever having to download stars clearly not belonging to the cluster. We will then see how the NGC6397 colour-magnitude diagrams obtained from a cone search and from an ADQL query compare to each other.

1.1 Before getting started ...

As you work through this tutorial – or any other TOPCAT tutorial –, do not forget to regularly save your TOPCAT session: File → Save Table(s)/Session. In the resulting window, select the tab Session as well as a directory of your choice (Location) and click OK. Not only will this back up your downloaded datasets, it will also save your own, user-defined, subsets and functions. Note that you can also save tables on a one-by-one basis.

2 A First Attempt Using Cone Search

To obtain a first idea of the kind of data we deal with here, follow the procedure outlined by Demleitner & Heinl (2024a) for the Pleiades. Refer to that tutorial for a bit of background on what happens here:

▷ **1** *Getting candidate members from Gaia* – In TOPCAT, hit VO → Cone Search, enter `gaia dr3` in keyKeywords and hit Find Services. Unless you want to experiment, pick `DR3 lite Cone` with a left click.

Enter `NGC6397` in Object Name, hit Resolve, and enter 0.35° (or 21 arcmin) as Radius. Send off the query. You should receive 96,524 objects.

▷ **2** *Plotting proper motions* – Go to Graphics → Plane Plot. The plot controls are in the bottom part of the plotting window. As X axis, choose `pmra`, as Y axis choose `pmdec`. Zoom into the overdensity around 3 mas/yr in pmra and −18 mas/yr in pmdec. That is our cluster in proper motion space.

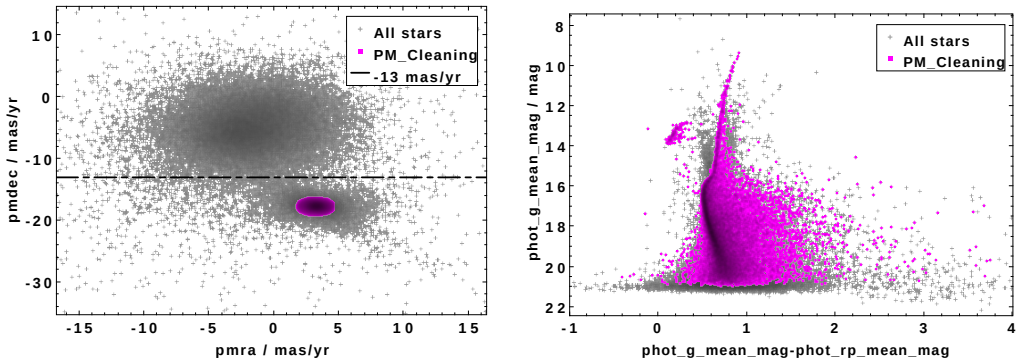


Figure 1: Distribution of the stellar dataset downloaded via a cone search in the proper motion space (left panel) and in the colour-magnitude diagram (right panel). There are the 96,524 stars present in the field-of-view of radius 0.35° centred on NGC6397 (grey symbols). Magenta symbols show stars likely to belong to NGC6397 based on their common proper motion ($\simeq 18,000$ stars).

▷ **3** *Plotting and drawing a subset* – Hit Subsets → Draw Blob Subset and draw a region around the overdensity in PM space. When done, do Subsets → Finish Drawing Blob. In the window popping up in response, in New Subset Name enter `PM cleaning` and click on Add Subset.

▷ **4** *Creating a CMD* – Open a second plane plot. This time,

- as X: enter `phot_g_mean_mag-phot_rp_mean_mag`,
- as Y: enter `-phot_g_mean_mag`.

Again go to the subsets tab in your plot and compare the CMDs of the whole dataset and of your NGC6397 subset.

Figure 1 shows how your plots might look like when you plot the full dataset as grey symbols in proper-motion space (left panel) and in the colour-magnitude diagram (Gaia G and RP passbands, right panel). In this case our selection for good candidate members of NGC6397 (magenta symbols) is about 18,000 stars.

Note that the structures in the colour-magnitude diagram widen strongly toward fainter apparent magnitudes due to increasing photometric errors. This is especially striking for the cluster main sequence.

This is nice, but obviously incomplete. We could scale things up, but for one this service limits responses to 100,000 rows by default, and then we already downloaded five times more objects than we wanted. It would be nice if we had better filters on the server side.

This is what TAP and ADQL let us do. With them, we will use limits in proper motion space. For instance $\mu_\delta < -13\,\mathrm{mas/yr}$, marked as the horizontal dash-dotted line in the left panel of Fig. 1, already is a first-order separation between stars more likely to belong to NGC6397 and stars belonging to the Galactic disc.

3 Using TAP and ADQL to Target a Cleaner Dataset

In this section, we write an ADQL query and use TAP (Table Access Protocol) to transport it in order to restrict, from the very beginning, the downloaded dataset to stars that

Reminder: Proper Motions

Proper motions are the motions along the celestial sphere, that is, perpendicular to the observer line-of-sight. The corresponding modification of the equatorial coordinates is usually resolved into two components, perpendicular and parallel to the celestial equator: μ_δ and $\mu'_\alpha = \mu_\alpha \cdot \cos(\delta)$, with the $\cos(\delta)$ factor accounting for the convergence of the meridians to the celestial poles (or think of it as a velocity in the tangential plane).
In the Gaia source catalogue, as in most modern astrometric catalogues, pmdec is just μ_δ, but pmra actually is $\mu_\alpha \cdot \cos(\delta)$. To check that pmra indeed is μ'_α rather than μ_α, check the metadata of the table you downloaded: in the TOPCAT's main window, try Views → Column Info or just click the  icon in the icon bar.

we will not have to discard as soon as we have downloaded them. Our proper motion criterion $\mu_\delta < -13\,\mathrm{mas/yr}$, for instance, already excludes many interlopers (Fig. 1).

If we select stars with more accurate proper motions, we will separate the stars of NGC6397 from those in the field more cleanly. More accurate magnitudes will yield a cleaner CMD (although this will come at a cost; see below). Last but not least, not retrieving "useless" stars decreases the sample size, thereby allowing us to extend the region of the sky we scrutinize. In what follows, the cone radius is widened from the 0.35° we used in the cone search to 2°.

3.1 Selection of the initial stellar dataset

▷ **1** *Prepare the service to pull astrometry from* – We want to work with Gaia DR3 and will look for a TAP service supporting it. So, in the TOPCAT main window, go to the VO → Table Access Protocol (TAP) dialog. In the input box Keywords:, type `gaia dr3` . Hit Find Services. While at least most of the services showing up should be fine for what we want to do, the concrete queries below are for the `GAVO DC TAP` service. For simplicity, select that, and then click Use Service near the bottom of the window.

We are then in the Use Service tab, where we will write our ADQL query into the text box near the window bottom.

▷ **2** *Write the ADQL query that will select the Gaia catalogue(s) and the Gaia stellar data you need* – As in the Pleiades tutorial (Demleitner & Heinl, 2024a), we use the catalogue `gaia.dr3lite`.

Since one of our goals is to build the CMD of NGC 6397 with absolute magnitudes, we also need the selected star distances. You might be tempted to compute them by inverting the Gaia parallaxes. But the parallax metadata in `gaia.dr3lite` points us to yet another catalogue, namely, `gedr3dist.main`: "If looking for a distance, consider joining with `gedr3dist.main` and using the distances from there".

In ADQL, when you want to combine data from two different tables, you will use the `JOIN` construct. It our case, we do not use the common crossmatch idiom (where you join on the distance between two objects) but instead on the identifier, `source_id`.

The relevant piece of ADQL looks like this:

```
SELECT source_id, ra, dec,
   pmra, pmdec, pmra_error, pmdec_error,
```

Graph reproducibility and STILTS

Although TOPCAT provides excellent plotting capabilities, you may wish to replot the content of TOPCAT-generated tables non-interactively (e.g., to reproduce plots a year from now). That is where the STILTS comes in handy. STILTS, which stands for *Starlink Tables Infrastructure Library Tool Set*, is the command-line counterpart of TOPCAT. Prompting TOPCAT to make a plot automatically generates the corresponding STILTS command lines. To access them, click on STILTS in the bottom pane of the Plane Plot window. You will find more information about how to use the STILTS commands on its homepage[a] and on GAVO's blog[b].

[a] https://www.star.bristol.ac.uk/~mbt/stilts/
[b] http://blog.g-vo.org/automating-tap-queries.html

```
   parallax, parallax_error,
   phot_g_mean_mag,  phot_g_mean_flux_over_error,
   phot_rp_mean_mag, phot_rp_mean_flux_over_error,
   r_med_geo, r_lo_geo, r_hi_geo
FROM gaia.dr3lite
LEFT OUTER JOIN gedr3dist.main
USING (source_id)
WHERE DISTANCE(POINT(ra, dec), gavo_simbadpoint('NGC 6397'))<2.00
      AND pmdec       < -13.
      AND pmra_error  < 0.5
      AND pmdec_error < 0.5
      AND phot_g_mean_flux_over_error  > 100.
      AND phot_rp_mean_flux_over_error > 100.
      AND astrometric_excess_noise     < 1.
      AND 0.5*(r_hi_geo-r_lo_geo) / r_med_geo < 0.1
```

The query above extracts from the `gaia.dr3lite` catalogue the Gaia `source_id` of the selected stars, their right ascension, declination, proper motions, parallax, apparent magnitudes in the Gaia G and RP bands along with their respective errors. Actually, a good way to come up with such select lists is to control-click on their rows in the Columns tab and then use the COLS button above the ADQL input box.

Additionally, the distance estimate (it is wise to stress on the word "estimate" here; see Sec. 3.3) of the selected stars, `r_med_geo`, is joined in from the table `gedr3dist.main`, along with `r_lo_geo` and `r_hi_geo`, its lower and upper 1-sigma-like bounds.

The keyword `LEFT OUTER JOIN` means that a star selected in `gaia.dr3lite` with no match in `gedr3dist.main` is nevertheless retained, with NULL values filling in the missing `gedr3dist.main` rows. Selected stars are imposed to lie within 2° from the centre of NGC6397.

Compared to a cone search, the ADQL query can impose much finer-grained constraints on the objects we will eventually download, going well beyond defining the size and location of the region of interest. For instance, in order to let us work with the proper motions and magnitudes without having to worry about errors too much, we impose a certain accuracy on the corresponding parameters, `pmra_error` < 0.5, `pmdec_error` < 0.5, `phot_g_mean_flux_over_error > 100` and `phot_rp_mean_flux_over_error > 100`. We

also constrain the relative error on stellar distances to be no greater than 10 %. How distance errors of the stars affect their absolute-magnitude estimates is discussed in Sec. 5.3.

▷ **3** *Execute the query* – As our ADQL query will run for some time, we go Async (Demleitner & Heinl, 2024b, p16). To do that, select Asynchronously next to Query mode, then hit Run Query. In the TAP window, under Job List, you can see what your job status is: `PENDING`, `QUEUED`, `EXECUTING` or `COMPLETED`.

▷ **4** *Inspect the data quality* – The resulting table appears in the Table list of the TOPCAT window. It contains 12,513 stars. View the contents (by double-clicking it). One thing you can do in that view is sort by the various columns (right click) to establish minimum and maximum values. Do that for `r_med_geo`, for instance, or for the magnitudes. Note that you can also search for specific values in specific columns (Rows → Search Columns).

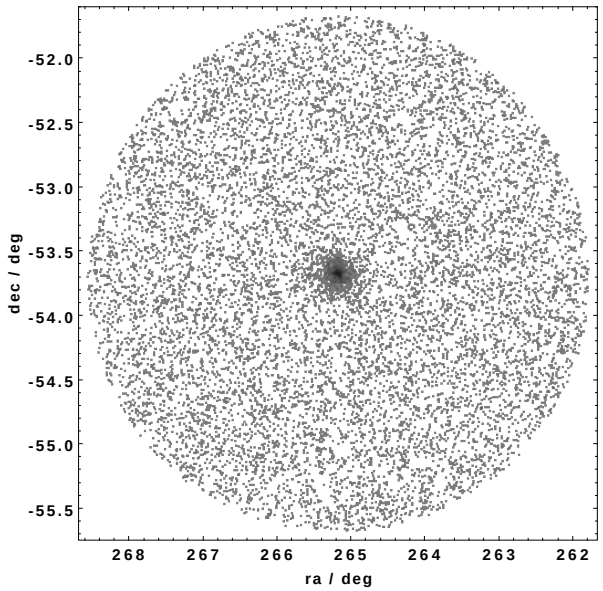


Figure 2: Distribution of the selected 12,513 stars in (α,δ) space.

▷ **5** *Plot the data in (α, δ) space* – Figure 2 shows the stellar sample distribution in the (α,δ) space. To obtain it, go back to the TOPCAT window and select Graphics → Plane Plot. Although not cleaned of its interlopers yet, the cluster already dominates the centre of the stellar distribution. We shall see later on that its actual extent is wider than what Fig. 2 suggests.

▷ **6** *A small riddle* – Use the compass icon of the Plane Plot window to measure the extent of the stellar distribution in your equivalent of Fig. 2. You will find that the "horizontal" (RA) diameter is greater than 4°. Can you figure out why that would be, given our cone-like condition?

3.2 Proper-motion based selection of the NGC6397 stars

▷ **1** *Plot the data in PM space* – Go to the Plane Plot window again and select `pmra` and `pmdec` as the X and Y coordinates.

▷ **2** *Locate the cluster and isolate a subset of candidate cluster stars in PM space* – Zoom

Note

The α-axis in Fig. 2 is oriented "leftward" (i.e., larger numbers left). Right ascensions are counted positively eastward. Picture yourself as an Earth-bound observer watching the sky, try to visualise the celestial equator and the right ascensions increasing eastward along it. For an Earth-bound observer whose north direction is upward, right ascensions increase leftward.
Incidentally, when you regard the celestial sphere from the outside, east would then be to your right and right ascensions would thus increase rightward. Fig. 13 will come back to this.

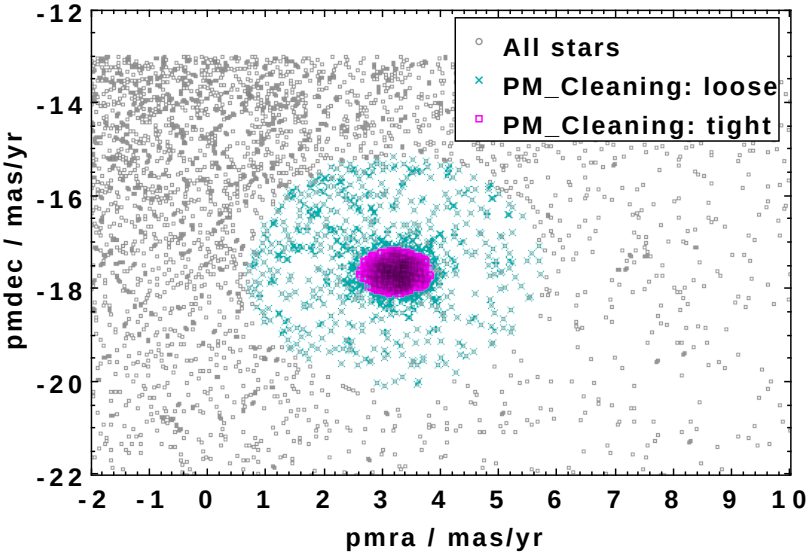


Figure 3: Selection of a "loose" subset of 1,555 stars and of a "tight" subset of 796 stars in the PM space. Both subsets, especially the second one, define our candidate cluster stars.

into the figure and the cluster eventually reveals itself as a tight concentration of points around $\mu'_\alpha \simeq 3\,\mathrm{mas/yr}$ and $\mu_\delta \simeq -18\,\mathrm{mas/yr}$. We build on the PM distribution of the stars to part the likely cluster members from the field interlopers.

In the Plane Plot window, click on Subsets → Draw algebraic subset → Rotated ellipse to draw an ellipse wrapping up the desired stellar subset (pick the centre first, then one half-axis and finally the other half-axis; click to fix a point). Once the ellipse has been drawn, do Subsets → Finish drawing subset (or use the equivalent toolbar button), give the subset a name and click OK (or Cancel to try again).

▷ **3** *Test several PM-space subsets* – Figure 3 highlights in colours two selected subsets, a "loose" one and a "tight" one. They contain 1,555 and 796 stars, respectively. To know the number of stars in your own subsets, go to the TOPCAT main window and, next to Row Subset (in the window to the right of the Table List), select your subsets. Take also a second to go back to the colour-updated (α, δ) distribution (i.e. your first Plane Plot window).

The two colour-highlighted subsets define our *candidate* cluster stars. Indeed, not all

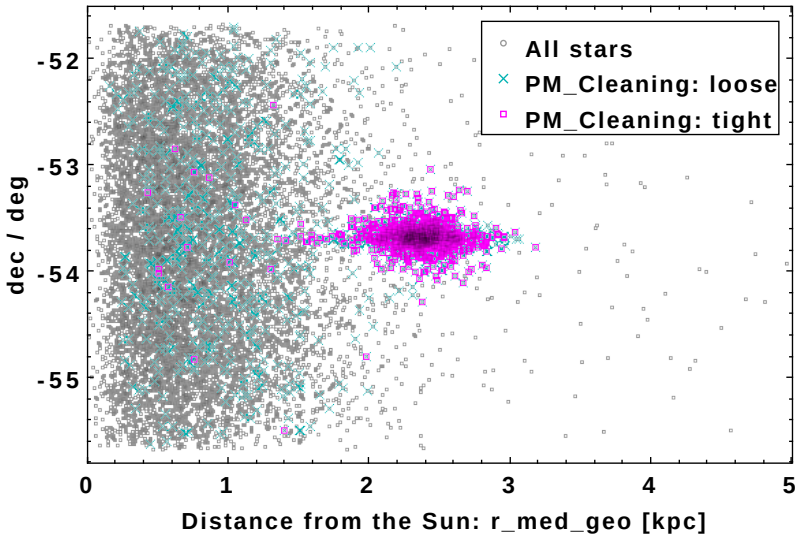


Figure 4: Declination of the 12,513 stars selected via the ADQL query on page 5 (open grey circles) as a function of their distance estimate from the Sun, *r_med_geo*. The two subsets of Fig. 3 are colour-and symbol-coded accordingly

of the 1,555, respectively, 796, stars belong to NGC6397, as Fig. 4 shows.

▷ **4** *Have a look at the distances of the candidate cluster stars* – Figure 4 plots the declination of the stars as a function of their distance to the Sun (i.e. `dec` as a function of `r_med_geo`). The cluster pops up at a distance of $D_{NGC6397} \simeq 2.4$ kpc. Most of the interlopers (open grey circles) are disc stars occupying the field-of-view foreground. As disc stars, they are closer to us than NGC6397.

Note the effect of tightening the PM-based cleaning of the stellar dataset: compared to the "tight" subset (magenta open squares), the "loose" subset (cyan crosses) contains more stars that likely belong to the Galactic disc. That is, the "tight" subset yields a less contaminated sample of NGC6397 stars.

▷ **5** *Get an estimate of NGC6397 distance* – To estimate the distance of NGC6397, we consider the "tight" subset and we take the median of its stellar distances `r_med_geo`. In the TOPCAT window, select Views → Column Statistics to generate a window entitled Row statistics. It contains a basic statistical analysis of the selected dataset, namely, the mean, standard deviation (SD), minimum and maximum of each stellar property. The number of rows, equivalently, the number of stars in the selected dataset, is also given (`nGood`). Note that you can swap one subset for another at the bottom of the window, next to Subset for Calculations.

For now, the Row statistics window does not provide us with the median. However, for samples containing interlopers (some foreground disc stars in our case), the median is a more robust measure than the mean. To get the median included in the statistical summary of a subset, click on Display and tick Median.

For the "tight" subset, the median of the stellar distances is 2.36 kpc, while the mean is 2.32 kpc (of course your values may differ slightly from ours as a result of your different PM-based selections). That the mean is slightly lower than the median stems from the few foreground interlopers. For this reason, it does not come as a surprise that the difference between mean and median is even larger if the "tight" subset is swapped for

the "loose" one (1.88 kpc vs 2.23 kpc, respectively).

In what follows, we consider the distance estimate of NGC6397 from the Sun to be 2.36 kpc. Its accuracy is discussed in the next section.

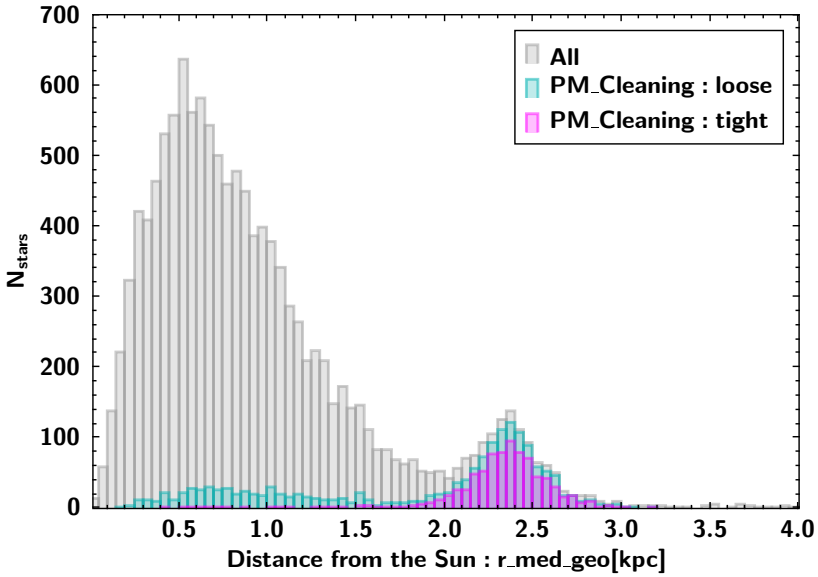


Figure 5: Histograms of the distances between the Sun and the stars for the whole dataset and its two subsets obtained via PM-cleaning (see Fig. 3). The distances are those established by Bailer-Jones et al (catalogue *gedr3dist.main*). The colour coding is as in Fig. 4

▷ **6** *Plot the stellar-distances distribution of your dataset and subsets* – Figure 5 presents the histograms of the stellar distances from the Sun for the whole dataset and its two PM-based subsets (Graphics → Histogram Plot). The dense stellar disc in the foreground dwarfs the cluster itself. The number of disc stars per bin first increases with increasing distance as our 4°-wide cone encompasses an ever wider cross-section of space. The histogram then plummets as the line-of-sight moves out of the Galactic disc into the much more diffuse stellar halo (the stellar halo mass is about two orders of magnitude smaller than the Galactic disc mass).

Note also how the "tight" PM-cleaning has excluded most (but not all) of the foreground stars still present in the "loose" subset (cyan histogram), as already shown in Fig. 4 (although one should keep in mind that the distance estimates may also be wrong and the odd cluster star may end up with an underestimated distance).

3.3 Do not confuse cluster size and distance error

Having estimated the cluster distance, we can turn the angular radius of the captured field-of-view (2°) into its linear radius at the cluster distance: For small angles, we do not even need trigonometric functions and can quickly estimate $r_{\rm fov} \simeq 2360\,{\rm pc} \times 2/57 \simeq 80\,{\rm pc}$ from our two-degree radius in radians.

Looking back to Fig. 4, we may thus be surprised by the seemingly strong elongation of NGC6397 along the line of sight. The cluster size measured (the word *measured* matters here!) along the line-of-sight is of order $2 \cdot 500$ pc. Yet, along the δ axis, the bulk of the cluster stars stretch over a small fraction of the field of view, i.e. $\ll 2 \cdot 80$pc.

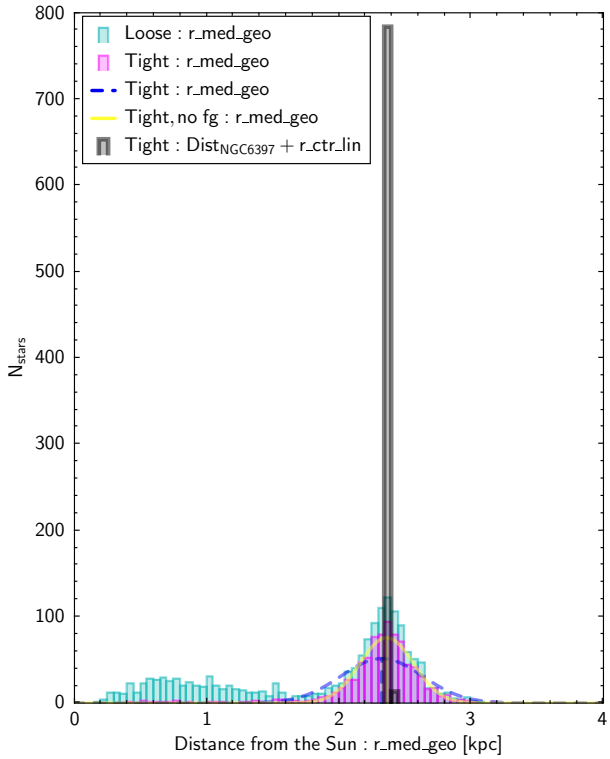


Figure 6: As in Fig. 5: histograms of the distance estimates between the Sun and the stars of the two subsets defined via PM-cleaning. Overlaid on the "tight"-subset (magenta) histogram are two Gaussians, the dashed one for the whole subset, the solid one excluding the 26 stars with heliocentric distances less than 1300pc ("no fg" ≡ no foreground). The black histogram depicts the distribution of the linear projected distances of the "tight"-sample stars from the cluster centre, shifted at the estimated cluster distance $D_{\rm NGC6397} = 2.36\,{\rm kpc}$.

Figure 6 quantifies both effects. The magenta histogram is the histogram of the stellar distances from the Sun. The black histogram is the histogram of the stellar distances from the cluster centre, projected in the plane of the sky, and shifted at the estimated cluster distance $D_{\rm NGC6397} \simeq 2.36\,{\rm kpc}$.

3.3.1 Accuracy on Stellar distances from the Sun

Figure 6 shows the "loose" and "tight" histograms of Fig. 5, completed by a dashed blue Gaussian function fitting the "tight" (magenta) histogram.

▷ **7** *Fitting and plotting Gaussians on the distance axis* – To fit a Gaussian function to a histogram, in the window Histogram Plot, click on the tab Form and there on Forms. The mean and standard deviation of the dashed blue Gaussian are, respectively, 2324 *pc* and 308 *pc* (the accuracy with which such numbers should actually be written is discussed in a short moment).

The presence of a few foreground disc stars in the "tight" subset widens the Gaussian compared to the bulk of the magenta histogram. You can be picky and define a tighter "tight" subset that excludes these foreground, likely disc, stars (keep in mind here that

our ADQL query has selected stars whose distance relative error is less than 10 %).

▷ **8** *Tightening up the cluster sample using distances* – For this, go back to Fig. 4, retain only the "tight" subset (in the bottom pane of the Plane Plot window, untick the whole dataset and the "loose" subset) and select a new subset that covers Sun distances from, say, 1.6 kpc to 3.0 kpc (Subsets → Draw algebraic subset → Draw algebraic subset: Box).

The solid yellow Gaussian corresponds to the "tight" subset with 26 foreground stars excluded (now 770 stars). The mean and standard deviation of this visually more satisfying Gaussian are, respectively, 2361 pc and 203 pc, unsurprisingly a slightly higher mean and a narrower standard deviation than for the dashed blue Gaussian.

Now consider our ADQL query again. It selects stars whose distance (estimated) relative error is 10 % at most. At the distance of NGC6397, this amounts to an error of $\simeq 240$ pc.

This shows that the standard deviation of either Gaussian in Fig. 6 reflects the errors impacting the measured distances of the selected stars. That is, in Fig. 4, the spread of the "tight" subset along the X-axis has **nothing to do with the cluster actual size**. Rather, it is driven by stellar distance errors.

3.3.2 Systematics and statistics in our distance estimate

At this stage, we need to discuss the accuracy with which we write the mean and standard deviation of the Gaussians shown in Fig. 6. As a reminder, the statistical error on the mean of a Gaussian obeys its standard deviation divided by the square root of the number of measurements. In our case, the measurements are the stellar distances of the "tight" subset, each such distance being an attempt to measure NGC6397 distance. The statistical error on the mean of the dashed blue Gaussian (796 stars) is 11 pc. This result, however, is not meaningful since yet another source of error must be considered, namely, the presence in the "tight" subset of several foreground stars. Foreground stars shift, systematically rather than randomly, the cluster distance estimate toward lower values. Compare the means of both Gaussians and the systematic error appears to be 3-4 times larger than the statistical error. To tame this systematic error, let us consider the solid yellow Gaussian, i.e. the one that has been purged of foreground stars, leaving it with 770 stars. The statistical error on its mean is 8 pc $\simeq$ 0.01 kpc.

We should thus write the mean as 2.36 ± 0.01 kpc and the standard deviation as 0.20 kpc (rather than 2361 pc and 203 pc).

Note that the mean is identical to the median of the stellar distances found in Sec. 3.2, which demonstrates that most foreground interlopers have indeed been excluded. One caveat, however; our distance estimate for NGC6397 rests on a small number of its stars. As we shall see in Sec. 4, all main sequence stars are excluded of our dataset by the distance and photometry accuracy imposed by the ADQL query.

3.3.3 Projected distances from the cluster centre

Now, let us compare the distribution of the stars along the line of sight with that across the field of view. For this purpose, we consider the distances between stars and cluster centre projected in the plane of the sky. For the "tight" subset, Fig. 6 shows in black their histogram, shifted at the estimated cluster distance $D_{NGC6397} = 2.36$ kpc. The black histogram cannot of course be wider than our 4°-wide field-of-view, $\simeq$ 160pc.

▷ **9** *Creating a computed column for the angular distance* – To plot this histogram, we must define, first, the *angular* distance and, second, the *projected linear* distance between the stars and the cluster centre. TOPCAT provides the first one. We define the second one.

In the TOPCAT window, select Views → Column Info (or hit) to display the column metadata, then Columns → New Synthetic Column (or hit). In the new window, Define Synthetic Column, you can define an expression for the column contents, give it a name, and provide further metadata (unit, description, UCD). Try it and make it a habit to fill out the extra info even if you don't *have* to do that right now. Your future self will be grateful to you.

To obtain the angular distance between a star and the centre of NGC6397, we use the TOPCAT-defined function `skyDistanceDegrees`:

```
r_ctr_ang = skyDistanceDegrees(ra,dec,265.175375,-53.674333)
```

where `ra` and `dec` are the star right ascension and declination, as labelled in the ADQL query. 265.175375° and −53.674333° are the right ascension and declination of the cluster centre, obtained from our preliminary `Cone Search` study. After hitting OK, inspect the column metadata, and you will find your new column as well as any extra metadata you added for the new column.

▷ **10** *Creating a computed column for the linear projected distance* – Next, let us define the *linear projected* distance between the cluster centre and the stars, in units of parsec:

```
r_ctr_lin = 2360* tan(degreesToRadians(r_ctr_ang))
```

Pay attention that, although we routinely use angles expressed in degrees, the argument of all TOPCAT trigonometric functions must be given in radians, hence the function `degreesToRadians` to convert `r_ctr_ang` from degrees to radians.

If you now inspect the table data, you should see the new columns with the values TOPCAT has computed using the formulae you entered.

The black histogram in Fig. 6 is, for the "tight" subset, the distribution of the projected linear distances `r_ctr_lin`, shifted for the estimated cluster distance $D_{NGC6397} = 2.36\,\mathrm{kpc}$. It depicts the cluster size more reliably than the distribution of the stellar distances from the Sun (magenta histogram): the errors on the positions from Gaia are *much* smaller that the one on the distances. As usual in astronomy: Distances are hard.

We shall come back to the actual cluster size in Sec. 10.

4 Colour-Magnitude Diagrams of NGC6397

We will build two types of colour-magnitude diagrams, one for the apparent magnitude of the stars (CmD) and one for their absolute magnitude (CMD).

Prior to this, we isolate two additional subsets containing, respectively, close-by foreground stars and more distant foreground stars (see Fig. 7, which is an update of Fig. 4). These subsets will help us understand how the heliocentric distances of the stars impact their location in the CmD and CMD.

▷ **1** *Define foreground and background subsets* – In Views → Row Subsets, hit the plus and then `Close_Foreground` as Subset Name and `r_med_geo<=400` as Expression. Similarly, define `Distant_Foreground` with `r_med_geo>=1400 && r_med_geo<=1800`.

▷ **2** *Plot the CmD* – To build the CmD of NGC6397, select Graphics → Plane Plot in the main TOPCAT window. In the command pane at the bottom, define

- the stellar colour `phot_g_mean_mag-phot_rp_mean_mag` as the X-coordinate,
- the stellar apparent magnitude `phot_g_mean_mag` as the Y-coordinate.

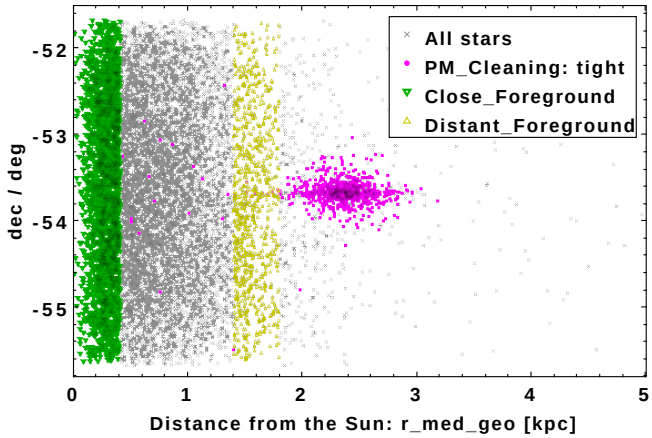


Figure 7: Same as Fig. 4 but with two additional stellar subsets, one covering close-by foreground stars (distances from the Sun ≤ 400 pc; green reversed triangles) and another one covering more distant foreground stars ($1400 \leq$ distance ≤ 1800 pc; yellow triangles).

`phot_g_mean_mag` and `phot_rp_mean_mag` are the stars' apparent magnitudes in the Gaia G and RP bands. To have a conventional orientation (bright stars at the top), click Axes and tick Y Flip. We shall work with the whole initial dataset of 12,513 stars, its close- and distant-foreground subsets and, of course, the "tight + no fg" subset, taken as representative of NGC6397.

All four samples are shown in Fig. 8 top panel, colour- and symbol-coded as in Fig. 7. To get the "tight + no fg" subset at the CmD forefront, go to the Subsets tab and drag it to the bottom of the subset list.

The cluster subgiant branch, red giant branch, horizontal branch and, even, the asymptotic giant branch pop up very clearly. The main sequence, however, is missing, as a result of the low astrometric and photometric errors imposed on the TAP-queried stellar sample (fainter stars have larger photometric errors). In between the hardly visible turn-off and the blue horizontal branch, a few stars wander: they are the so-called Blue Stragglers.

Close-by foreground stars appear, unsurprisingly, brighter than their distant foreground siblings (recall that we are dealing with the *apparent magnitude* here). Other than that, foreground stars do not draw any distinctive patterns. This will change if we move to the cluster CMD. Let us modify the definition of the Y-coordinate and swap `phot_g_mean_mag` for

```
5-5*log10(r_med_geo)+phot_g_mean_mag,
```

which reflects the definition of the *absolute* magnitude. Keep in mind that `r_med_geo` must be expressed in pc. Now most foreground stars fall in line, defining the main sequence. Recall that the green/grey/yellow sequences do not belong to the cluster.

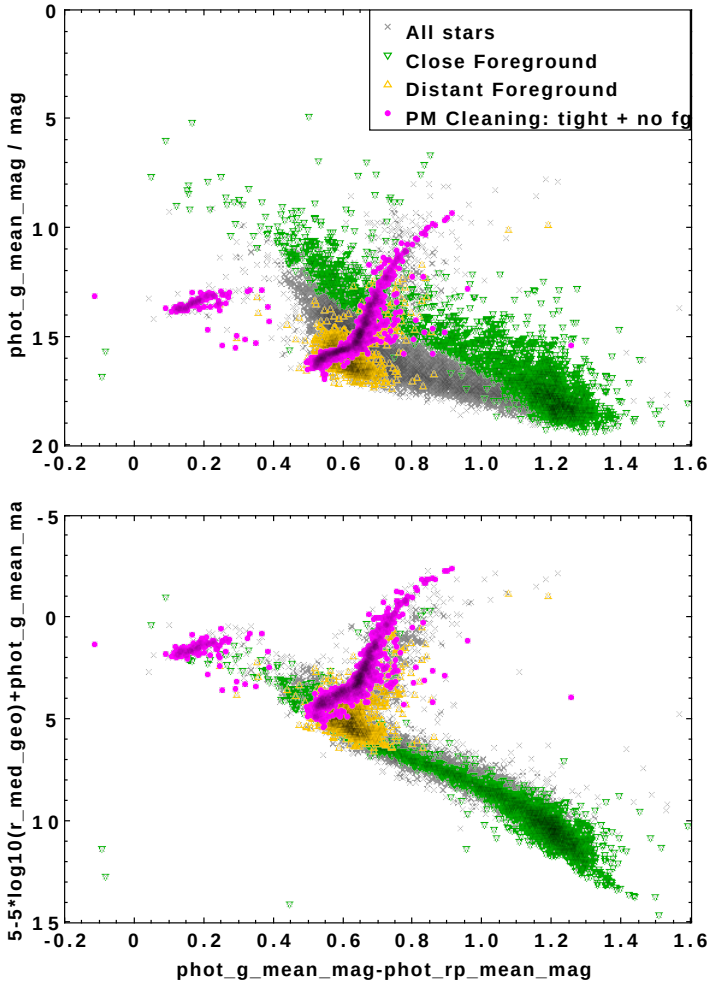


Figure 8: Top panel: Colour-magnitude diagram of the entire dataset of 12,513 stars, with the Y-axis showing the stars' apparent magnitudes. The subsets defined in Fig. 7 are colour-and symbol-coded accordingly. Bottom panel: same but with the Y-axis showing the absolute magnitude.

5 Magnitude and Colour Errors

5.1 Apparent magnitude error as a function of flux error

The `gaia.dr3lite` catalogue does not provide the error on the apparent magnitudes, either G or RP[1]. But it does provide for each star the ratio between the flux integrated in a given passband and its error. We refer to it as the F/σ_F ratio, writing F for the flux and σ_F for its error. For the G-band flux, the F/σ_F ratio is labelled `phot_g_mean_flux_over_error` in `gaia.dr3lite`. It is one of the quantities downloaded by our ADQL query.

Let us, for a given passband, infer the apparent magnitude error as a function of σ_F/F.

The relation between a flux F and its corresponding apparent magnitude m obeys

$$m = -2.5\log_{10}(F) + C\,, \tag{1}$$

with C a constant. We move the variable, namely, the flux, downwards and upwards by its error to obtain the magnitude m increased/decreased by its corresponding upper and lower error estimates, $\sigma_m^{\rm up}$ and $\sigma_m^{\rm low}$. That is, recalling that a flux decrease yields a higher magnitude, we successively get:

$$\begin{aligned} m + \quad \sigma_m^{\rm up} &= \quad -2.5\log_{10}(F-\sigma_F) + C \\ m - \quad \sigma_m^{\rm low} &= \quad -2.5\log_{10}(F+\sigma_F) + C\,, \end{aligned} \tag{2}$$

$$\begin{aligned} \sigma_m^{\rm up} &= -2.5\log_{10}\left(1-\frac{\sigma_F}{F}\right) > 0 \\ \sigma_m^{\rm low} &= \quad 2.5\log_{10}\left(1+\frac{\sigma_F}{F}\right) > 0 \end{aligned} \tag{3}$$

with σ_F/F known for each star of the selected sample. For the G band, $\sigma_F/F = 1/$`phot_g_mean_flux_over_error`.

Reducing Eq. 3 to the first order (i.e., the flux relative error is small, $\sigma_F/F \ll 1$), the upper and lower error estimates of the apparent magnitude are no longer unequal as they become:

$$\begin{aligned} \sigma_m^{\rm up} &\simeq \frac{2.5}{\ln 10}\frac{\sigma_F}{F} \\ \sigma_m^{\rm low} &\simeq \frac{2.5}{\ln 10}\frac{\sigma_F}{F}\,. \end{aligned} \tag{4}$$

Equation 4 is also what we find by propagating the flux error, assumed to be small, from the magnitude definition (Eq. 1):

$$\sigma_m^{\rm small} = \sqrt{\left(\frac{\partial m}{\partial F}\right)^2 \sigma_F^2} = \frac{2.5}{\ln 10}\frac{\sigma_F}{F}\,. \tag{5}$$

[1] In case you wonder why: the error model is roughly a Gaussian in the flux. For large relative errors, that translates into seriously non-Gaussian errors after taking the logarithm producing the magnitudes, and thus a single FWHM-like error is not a good measure. But see below why you're still fine for relative errors below 10%.

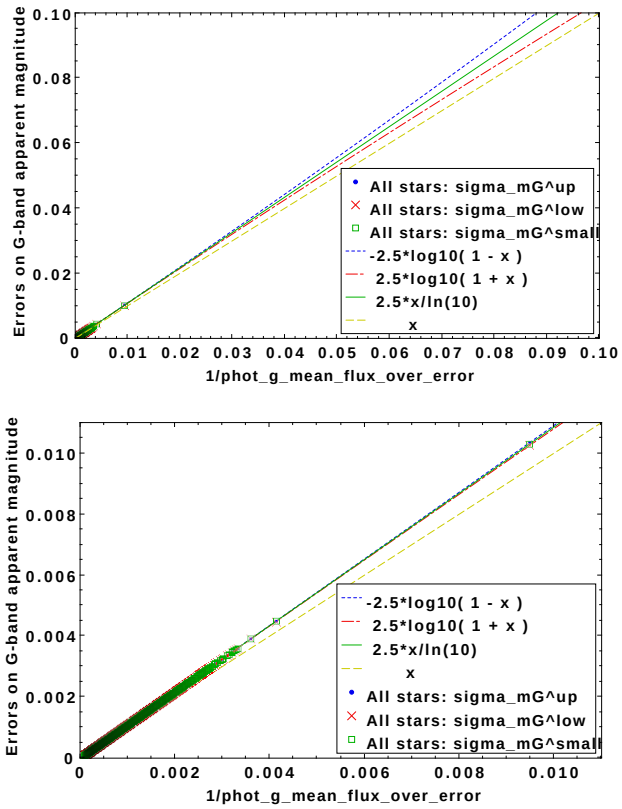


Figure 9: Top panel: Upper and lower error estimates of the G-band apparent magnitude as a function of the ratio σ_{F_G}/F_G (Eq. 3, dotted and dash-dotted lines). The solid line depicts Eq. 5, the apparent magnitude error for small flux relative errors. Symbols show the data for the same three cases. Bottom panel: same as top panel but with a zoom onto the data.

Note that

$$\sigma_m^{\rm small} = 1.09\frac{\sigma_F}{F} \simeq \frac{\sigma_F}{F}\ . \tag{6}$$

Therefore, for small flux errors σ_F in a given passband, the error $\sigma_m^{\rm small}$ on the corresponding apparent magnitude m almost follows a 1-to-1 relation with the ratio σ_F/F. The latter is provided by the `gaia.dr3lite` catalogue.

Equations 5-6 are directly relevant to the selected stellar sample since our ADQL query has imposed `phot_g_mean_flux_over_error > 100`, or equivalently $\sigma_F/F < 0.01$. With that cut, assuming symmetric errors in the magnitude is safe.

▷ **3** *Adding columns for errors* – Equations 3 and 5 can be added as new columns of our TAP-queried table just like we did for the projected distances between stars and cluster centre. In the TOPCAT window, click again on , then on . . In the popped-up window, do Define Synthetic Column, define columns `sigma_g_up` and `sigma_g_low` (Eq. 3), and `sigma_g_small` (Eq. 5).

▷ **4** *Convincing yourself of the validity of the error estimates* – Do another plane plot to reproduce something like Fig. 9. To have the three columns you created in the last step in one plot, do Layers → Add Position Control twice, and then do Layers → Add Function Control three times, entering the respective expressions from Fig. 9.

The top panel of Fig. 9 shows the upper and lower error estimates of the G-band apparent magnitude (Eq. 3, dotted and dash-dotted lines, respectively) as a function of the ratio $\sigma_{F_\mathrm{G}}/F_\mathrm{G} = 1/$`phot_g_mean_flux_over_error`.

The solid line depicts the G-band magnitude error when $\sigma_F/F \ll 1$ (Eq. 5). The bottom panel shows the same functions but zoomed in onto the region containing the data. For our sample of 12,513 stars, σ_m^{up}, σ_m^{low} and $\sigma_m^{\mathrm{small}}$ are indistinguishable from each other; this was to be expected given our small flux errors. Not only has our ADQL query imposed `phot_g_mean_flux_over_error > 100`, G-band magnitude errors are actually no larger than 0.003 mag for most selected stars anyway because of our other quality constraints. This is much smaller than the symbol size in the top panel Fig. 8.

If you repeat this exercise for the RP-band apparent magnitudes, you will find that a similar picture emerges; errors in RP are usually larger (up to three times) than in G, though. With this, we are ready to derive errors on the colour index $m_G - m_{\mathrm{RP}}$. This is the topic of the next section.

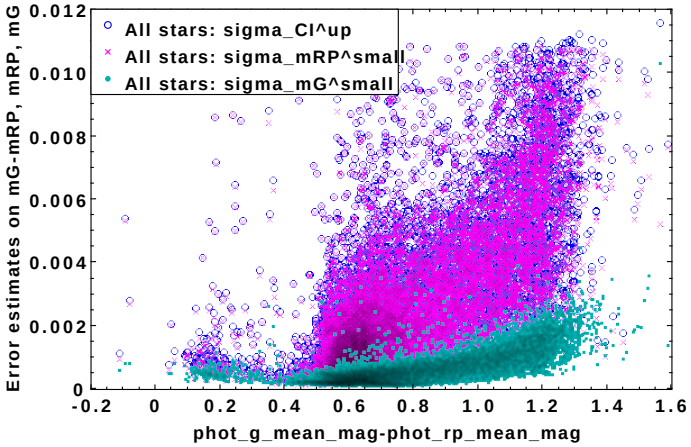


Figure 10: Error on the colour index $CI = m_\mathrm{G} - m_\mathrm{RP}$ in dependence of CI. The whole data set of 12,513 stars is shown. Also depicted are the errors on the G- and RP-band apparent magnitudes. CI errors are contributed to mostly by the RP band, with a negligible G-band contribution.

5.2 Colour-index error as a function of flux errors

Propagating the errors on m_G and m_RP, we find for the upper and lower error estimates of the colour index $CI \equiv m_\mathrm{G} - m_\mathrm{RP}$

$$\begin{aligned} \sigma_{CI}^{\mathrm{up}} &= \sqrt{(\sigma_{m_\mathrm{G}}^{\mathrm{up}})^2 + (\sigma_{m_\mathrm{RP}}^{\mathrm{low}})^2} \\ \sigma_{CI}^{\mathrm{low}} &= \sqrt{(\sigma_{m_\mathrm{G}}^{\mathrm{low}})^2 + (\sigma_{m_\mathrm{RP}}^{\mathrm{up}})^2}\,. \end{aligned} \tag{7}$$

In the small-error regime, appropriate for our dataset, we can write:

$$\begin{aligned} \sigma_{CI}^{\mathrm{small}} &= \sqrt{\left(\frac{\partial CI}{\partial m_\mathrm{G}}\right)^2 (\sigma_{m_G}^{\mathrm{small}})^2 + \left(\frac{\partial CI}{\partial m_\mathrm{RP}}\right)^2 (\sigma_{m_\mathrm{RP}}^{\mathrm{small}})^2} \\ &= \sqrt{(\sigma_{m_\mathrm{G}}^{\mathrm{small}})^2 + (\sigma_{m_\mathrm{RP}}^{\mathrm{small}})^2}\,, \end{aligned} \tag{8}$$

which is indeed what Eq. 7 provides when $\sigma^{\rm up}_{m_{\rm G}} \simeq \sigma^{\rm low}_{m_G} \simeq \sigma^{\rm small}_{m_{\rm G}}$ (see Fig. 9), with a similar argument for the RP band. The apparent-magnitude errors $\sigma^{\rm small}_{m_{\rm G}}$ and $\sigma^{\rm small}_{m_{\rm RP}}$ depend on their respective ratios σ_F/F as shown by Eq. 5.

$\sigma^{\rm up}_{CI}$ from Eq. 7 is shown in dependence of its colour index in Fig. 10. Given that $\sigma^{\rm up}_{CI} \simeq \sigma^{\rm low}_{CI}$, $\sigma^{\rm low}_{CI}$ is ignored for the sake of clarity. The figure also shows $\sigma^{\rm small}_{m_{\rm G}}$ and $\sigma^{\rm small}_{m_{\rm RP}}$. As expected, the colour index errors are driven mostly by the RP-band given that for most stars $\sigma^{\rm small}_{m_{\rm G}} \ll \sigma^{\rm small}_{m_{\rm RP}}$ and that individual errors are combined in quadrature. That the colour-index error roughly increases as the colour index gets redder results from the reddest stars being also – on average – the faintest ones (see Fig. 8 top panel).

5.3 Absolute magnitude error as a function of distance error

Compare both panels of Fig. 8: cluster sequences in the bottom panel are not as tight as in the top panel. For instance, the asymptotic giant branch hardly pops up in the bottom panel. This is not surprising since its Y-axis shows the absolute magnitude of the stars, which depends on both their apparent magnitude and distance. Since apparent magnitude errors are negligible for the vast majority of the stars (less than 0.003 mag, see Sec. 5.1 and Fig. 9), the mere inspection of the cluster sequences in Fig. 8 tells us that absolute magnitude errors must be dominated by distance errors.

In the absence of extinction, the G-band absolute magnitude of a star, $M_{\rm G}$, obeys `phot_g_mean_mag-5*log10(r_med_geo)+5`, equivalently

$$M_{\rm G} = m_{\rm G} - 5\log_{10}(d) + 5\,, \tag{9}$$

with stellar distances d expressed in pc. If we limit ourselves to the impact of distance errors, the $M_{\rm G}$ upper and lower error estimates, $\sigma^{\rm up}_{M_{\rm G}}$ and $\sigma^{\rm low}_{M_{\rm G}}$, follow from

$$\begin{aligned} M_{\rm G} + \sigma^{\rm up}_{M_{\rm G}} &= m_{\rm G} - 5\log_{10}(d - \sigma^{\rm low}_d) \quad +5 \\ M_{\rm G} - \sigma^{\rm low}_{M_{\rm G}} &= m_{\rm G} - 5\log_{10}(d + \sigma^{\rm up}_d) \quad +5\,. \end{aligned} \tag{10}$$

Combining with Eq. 9 yields

$$\begin{aligned} \sigma^{\rm up}_{M_{\rm G}} &= -5\log_{10}\left(\tfrac{d-\sigma^{\rm low}_d}{d}\right) \quad > 0 \\ \sigma^{\rm low}_{M_{\rm G}} &= \quad 5\log_{10}\left(\tfrac{d+\sigma^{\rm up}_d}{d}\right) \quad > 0\,. \end{aligned} \tag{11}$$

equivalently

$$\begin{aligned} \sigma^{up}_{M_G} &= -5\log_{10}\left(\tfrac{\rm r_lo_geo}{\rm r_med_geo}\right) \quad > 0 \\ \sigma^{low}_{M_G} &= \quad 5\log_{10}\left(\tfrac{\rm r_hi_geo}{\rm r_med_geo}\right) \quad > 0\,. \end{aligned} \tag{12}$$

`r_lo_geo` and `r_hi_geo` are the stellar distance lower and upper limits as provided by the catalogue `gedr3dist.main`. Look up their definitions in Views → Column Info: They are the 16th and 84th percentiles of the distance posterior, or something like a 1σ confidence interval.

▷ **5** *Visualising the absolute magnitude error as a function of the estimated distance* – Use Eq. 11 to add errors on the absolute magnitude in your table as above. Then open a plane plot and add a layer each for the lower and upper error estimates with `r_med_geo` on the abscissa. With a few style tweaks, this should result in something like Figure 11.

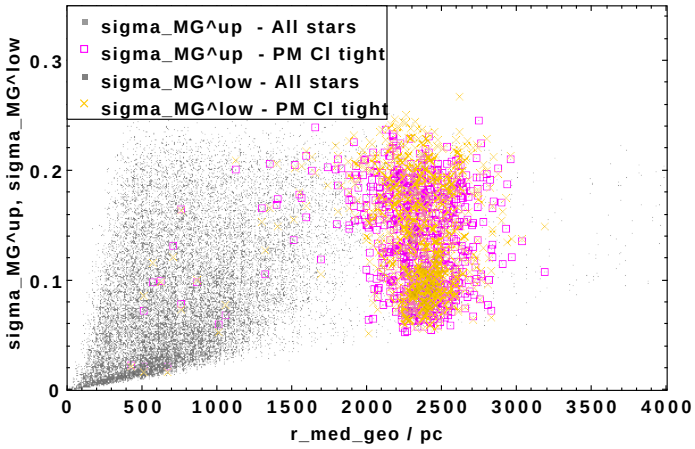


Figure 11: Upper and lower error estimates of the star absolute magnitude as a function of stellar distances as estimated by Bailer-Jones et al. The magenta open squares and yellow crosses depict the "tight" stellar subset, which, except for a few foreground stars, consists mostly of NGC6397 stars.

The likely members of NGC6397 (i.e., the "tight" subset) are highlighted as open magenta squares (for $\sigma^{\rm up}_{M_{\rm G}}$) and yellow crosses (for $\sigma^{\rm low}_{M_{\rm G}}$). The contribution of stellar distance errors to the absolute magnitude error ranges from $\simeq 0.05$ to $\simeq 0.25$. This is more than an order of magnitude higher than the G-band apparent magnitude error estimated in sect. 5.1. Therefore, the contribution of the apparent magnitude error to the absolute magnitude error can be neglected.

The slight blurring of the cluster CMD sequences in the bottom panel of Fig. 8 reflects the abolute magnitude errors $\sigma^{\rm up}_{M_{\rm G}}$ and $\sigma^{\rm low}_{M_{\rm G}}$ shown in Fig. 11, themselves fed by stellar distance errors.

6 Comparing the CMDs of Several Clusters

We now compare the CMDs of two star clusters of the Galactic disc, namely, M45 (the Pleiades) and M67, with that of NGC6397. These three clusters define sequences in age, distance from the Sun and metallicity, with NGC6397 the oldest, most distant and most metal-poor one, and M45 the youngest, closest and most metal-rich one.

Our CMDs follow from cone searching and PM cleaning the field-of-view of each cluster (for details, see Sec. 2). We use cone searches for the sake of simplicity. At this stage, you have in your hands the results of the NGC6397 cone search (Sec. 2) and, presumably, also that of the Pleiades. To download a M67 search dataset, follow the instructions outlined in Sec. 2, which we briefly summarize below:

▷ **1** *Getting candidate members from Gaia* – Do a Cone Search of the Gaia `DR3 lite Cone` catalogue with `M67` as the Object Name. For a field-of-view Radius of 1.0°, you should receive 20,276 objects.

▷ **2** *Plotting proper motions* – Build a Plane Plot of `pmra` versus `pmdec`. Zoom into the overdensity around −11 mas/yr in pmra and −3 mas/yr in pmdec. That is M67 in PM space.

▷ **3** *Plotting and drawing a subset* – Use a freehand or ellipse subset to collect the cluster PM overdensity in a subset.

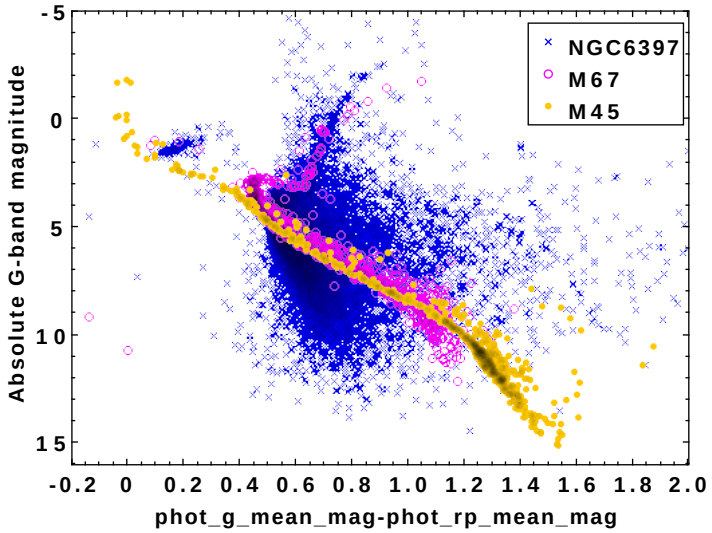


Figure 12: Colour-absolute magnitude diagrams of NGC6397, M67 and M45. These three clusters define a sequence in age, ranging from $\simeq 12$ Gyr for NGC6397, to 3.5-4 Gyr for M67 and 125 Myr for M45, which is reflected in their distinct main-sequence turn-off luminosities.

▷ **4** *Creating a CmD* – Open a second plane plot to show `-phot_g_mean_mag` as a function of `phot_g_mean_mag-phot_rp_mean_mag`.

▷ **5** *Turning the CmD into a CMD* – We are aiming for the cluster CMD, rather than the cluster CmD. Hence, we must turn the the vertical axis into the G-band *absolute* magnitude, which means we need distances or parallaxes. We cannot straightforwardly join with the server-side `gedr3dist.main` and therefore confine ourselves to distance estimates obtained by inverting stellar parallaxes. Note that you could still obtain better distance estimates using TAP upload joins as per Demleitner (2017). For us, most parallaxes in our data sets are good enough.

Modify your definition of the Y-axis to: `5+5*log(0.001*parallax)+phot_g_mean_mag`, not forgetting the factor 0.001 when taking the logarithm of the Gaia parallaxes since these are expressed in units of mas. Also do not forget to flip the Y-axis to get the brightest stars at the top of the diagram.

▷ **6** *Building a composite CMD of all three clusters* – Repeat the above procedure for both the M45 and NGC6397 PM-defined subsets. To add one cluster at a time, the PM-defined subsets of M45 and NGC6397 to the CMD, hit Layers → Add Position Control and select the corresponding table and subset. This should give you a diagram similar to Fig. 12. Sanity check: the absolute magnitude of the NGC6397 main-sequence turn-off must be around $M_G \simeq 4$, as expected for an old globular cluster.

The age estimates of NGC6397, M67 and M45 are, respectively, 12 Gyr, $3.5 \cdots 4$ Gyr, and 125 Myr. In the stacked colour-magnitude diagram (Fig. 12), this two-order of magnitude age range shows up as different luminosities for the cluster main-sequence turn-offs. While the Pleiades show the brightest turn-off, at $M_G \simeq -1$, the turn-offs of M67 and NGC6397 are fainter, with absolute magnitudes of $M_G \simeq 3$ and $M_G \simeq 4$, respectively.

NGC6397, M67 and M45 also define a sequence with respect to their distances to

the Sun: 2.36 kpc for NGC6397 (Sec. 3.3.1), $0.8 \cdots 0.9$ kpc for M67 and 136 pc for M45 (Melis et al., 2014).

While the range in cluster ages reflects itself in the range of main-sequence turn-off luminosities, different distances from the Sun are reflected in how deep the *observed* main sequence runs in the colour-magnitude diagram. The Pleiades being the closest of the three clusters, the stars at the faint end of its main sequence are more easily detected than in the more distant NGC6397. As a result, the observed main sequence of M45 reaches fainter absolute magnitudes than that of NGC6397 or M67. This does not mean that NGC6397 does not contain main-sequence stars intrinsically as faint as M45. But at NGC6397's distance, these stars are too faint to have been detected by Gaia.

Now compare NGC6397's CMD of Fig. 12 (the Y-axis shows stellar absolute magnitudes) and the NGC6397 CmD of Fig. 1 (the Y-axis shows stellar apparent magnitudes). Note how more poorly defined the main sequence and the red giant branch of the CMD are compared to those of the CmD. Echoing what we saw in Fig. 8 and what we studied in Sec. 5.3, the large scatter affecting the CMD reflects stellar parallax errors. Note also how the neat cut in apparent magnitude (at $m_G \simeq 21$, see Fig. 1) disappears in the CMD (Fig. 12), again reflecting the impact of uncertain stellar parallaxes, especially at the faint end of the main sequence.

7 Locating the Clusters on the Celestial Sphere

We now locate, in projection on top of the celestial sphere, our three clusters NGC6397, M67 and M45.

▷ **7** *Make a sky plots of one cluster* – In the TOPCAT main window, do Graphics → Sky Plot. In the resulting window, in the tab Position, select next to Table one of our cone-search datasets (say, M45; see Sec. 6).

The Data Sky System must be kept on `equatorial`, since we have RA and Dec as coordinates. TOPCAT has also correctly identified the columns containing them using their UCDs.

▷ **8** *Overplot the other clusters* – In the Sky Plot window, use Layers → Add Position Control and repeat the procedure for the two other clusters.

Figure 13 shows on the celestial sphere the fields of view of the M45 and M67 cone searches. The apparently disc-like structures are actually made of the many dots depicting the cone-searched stars. Check this by zooming onto any field of view (maybe after having reduced the symbol size to a dot, which you can do under the tab Form). Keep in mind that the raw dataset of a cone search consists of both clustered stars and interlopers (unless you configure this differently in the Subsets tab).

The respective diameters of the M45 and M67 fields-of-view are here $2 \times 1.9°$ and $2 \times 1°$. Recall that the angular extent of a field-of-view can be measured with the compass-like icon of the Sky Plot window (after clicking the icon, click, hold and drag from start to end). In contrast to Fig. 2, the right ascension extent equates with the extent of the field-of-view itself, not its projection onto the celestial equator. Therefore, no $\cos(\delta)$ factor comes into play here.[2]

To better appreciate the locations of the clusters, add galactic and ecliptic grids:

[2] If you happen to measure a field-of-view diameter smaller than the one imposed by your cone search, go back to the TOPCAT window and check the row number of your downloaded dataset. More than likely, the row number has hit the service's match limit (here, 100,000) for a cone search.

▷ **9** *Add an ecliptic grid* – In the Sky Plot window, do Layers → Add a new SkyGrid and select `ecliptic` as Sky System. Note that both M67 and M45 sit close to the ecliptic, as expected for clusters hosted by zodiacal constellations. This also explains why, with naked eyes, you occasionally witness Venus as it transits through the Pleiades (e.g., in April 2020, see https://apod.nasa.gov/apod/ap200402.html).

▷ **10** *Add a galactic grid* – Repeat with the Galactic-coordinate system (i.e., select `galactci` in Sky System).

▷ **11** *Configure the plot to work as a virtual sky globe* – In the `Sky Plot` window, navigate through Axes → Projection and there, *untick* Reflect Longitude Axis. This ensures a right-handed equatorial coordinate system, i.e., right ascensions increase from west to east, equivalently here, from left to right[3]. This is opposite to the x-axis of Fig. 2 because, now, you consider the celestial sphere from the outside rather than from the inside.

If you are familiar with the constellations, knowing that M45 lies in Taurus and M67 in Cancer will help you make sense of this. M45 lies to the left of M67 in Fig. 13. But Taurus and M45 lie to the right of Cancer and M67 when you look at the sky from the Earth, your head oriented to the north (as it is when you consider Fig. 13). For the sake of completeness, let us mention that NGC6397 is located in the southern constellation Ara.

Magenta lines in Fig. 13 mark the equatorial-coordinate system. The great circle $\alpha = 90°$ (or 6 h unless you uncheck Sexagesimal in the Grid tab in Axes), the one crossed by the Sun position on the day of the summer solstice, stands at the figure forefront, connecting both poles. It pops up in yellow, though, being overlaid by the great circle of ecliptic longitude $\lambda = 90°$. NGC6397, given its right ascension $\alpha \simeq 270°$, sits on the opposite side of the celestial sphere.

Keep in mind that the Galactic centre coordinates are $l = 0°$ (Galactic longitude) and $b = 0°$ (Galactic latitude). In Fig. 13, the Galactic centre is thus located on the background side of the celestial sphere. By contrast, the coordinates $l = 180°$ and $b = 0°$ point in the direction of the Galactic anti-centre, on the foreground of Fig. 13.

Those of you enjoying this kind of plot may notice that the position of the Sun on the day of the Summer solstice (ecliptic latitude $\beta = 0°$, ecliptic longitude $\lambda = 90°$, right ascension $\alpha = 90°$) is located almost exactly in the Galactic mid-plane (Galactic latitude $b \simeq 0°$) and close to the Galactic anti-centre (Galactic longitude $l = 180°$). You may wish to use Stellarium[4] or Aladin[5] to get the perspective of an Earth-bound observer on all of this.

Some may also find Fig. 13 somewhat arid. A great way of highlighting, even if coarsely, the Galactic mid-plane, which defines the fundamental circle (i.e., equator-like) of the Galactic-coordinate system, is by colour-coding the celestial sphere for Galactic extinction. Not only does the Galactic disc contain the vast majority of the stars of our Galaxy, it also contains lots of light-absorbing gas and dust. On an extinction map of the sky, the Galactic disc thus pops up as regions of high Galactic extinction combined to, as we shall see in Sec. 8, regions in which the Galactic extinction could not be measured based on the Gaia DR3 data.

[3] Any coordinate system added to a *Sky Plot* after unticking Reflect Longitude Axis will be right-handed as well.

[4] https://stellarium.org

[5] http://aladin.u-strasbg.fr/

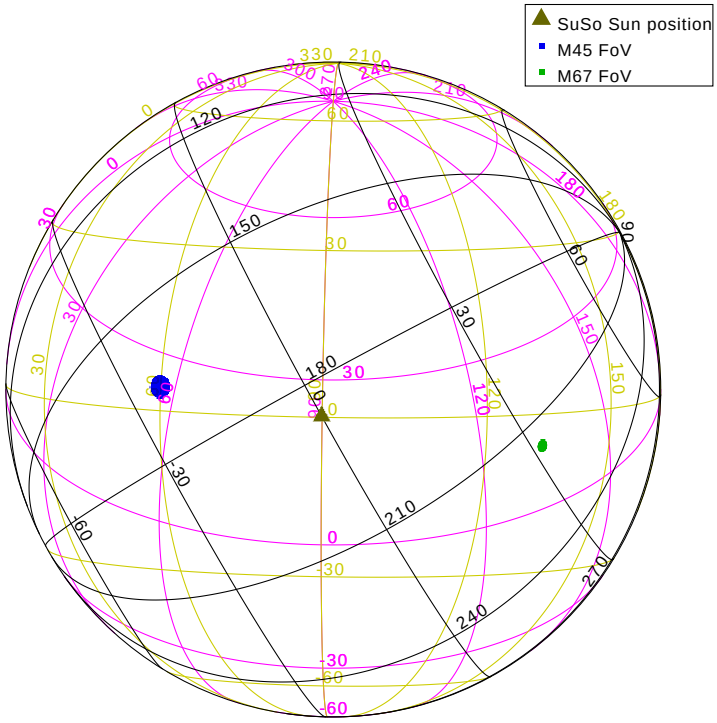


Figure 13: Outer surface of the celestial sphere gridded with the equatorial (magenta), ecliptic (orange) and Galactic (black) coordinate systems. The north celestial (equatorial) pole lies almost at the top of the figure. The Galactic mid-plane (i.e., $b = 0°$) is oriented from the bottom-right to the top-left of the figure. The fields of view used to cone-search for M45 (Sec. 2) and M67 (Sec. 6) are highlighted in blue and green, respectively. The brown triangle at the figure forefront highlights the Sun position on the day of the summer solstice.

8 Sky Tessalation and Total Galactic Extinction Map

To build an extinction map of the sky, we need a dataset consisting of the extinction caused by the Galactic Interstellar Medium (ISM) in dependence of sky position. We use the catalogue `gaiadr3.total_galactic_extinction_map`. It provides, as a function of sky position, the Total Galactic Extinction (TGE), namely, the "*full cumulative foreground extinction by the Milky Way on distant extragalactic sources*" (Delchambre et al., 2023, their Sec. 7). In this table, the TGE is the `a0` parameter. To first order, `a0` is similar to the text-book visual extinction A_V you may be more familiar with. But see Delchambre et al. (2023) for caveats.

Before downloading the dataset, let us consider how sky positions are defined and their TGE values are measured.

8.1 Positions on the celestial sphere

The table `gaiadr3.total_galactic_extinction_map` builds on the *Hierarchical Equal Area isoLatitude Pixelization* (HEALPix) of the sphere, a tessalation of the sky with

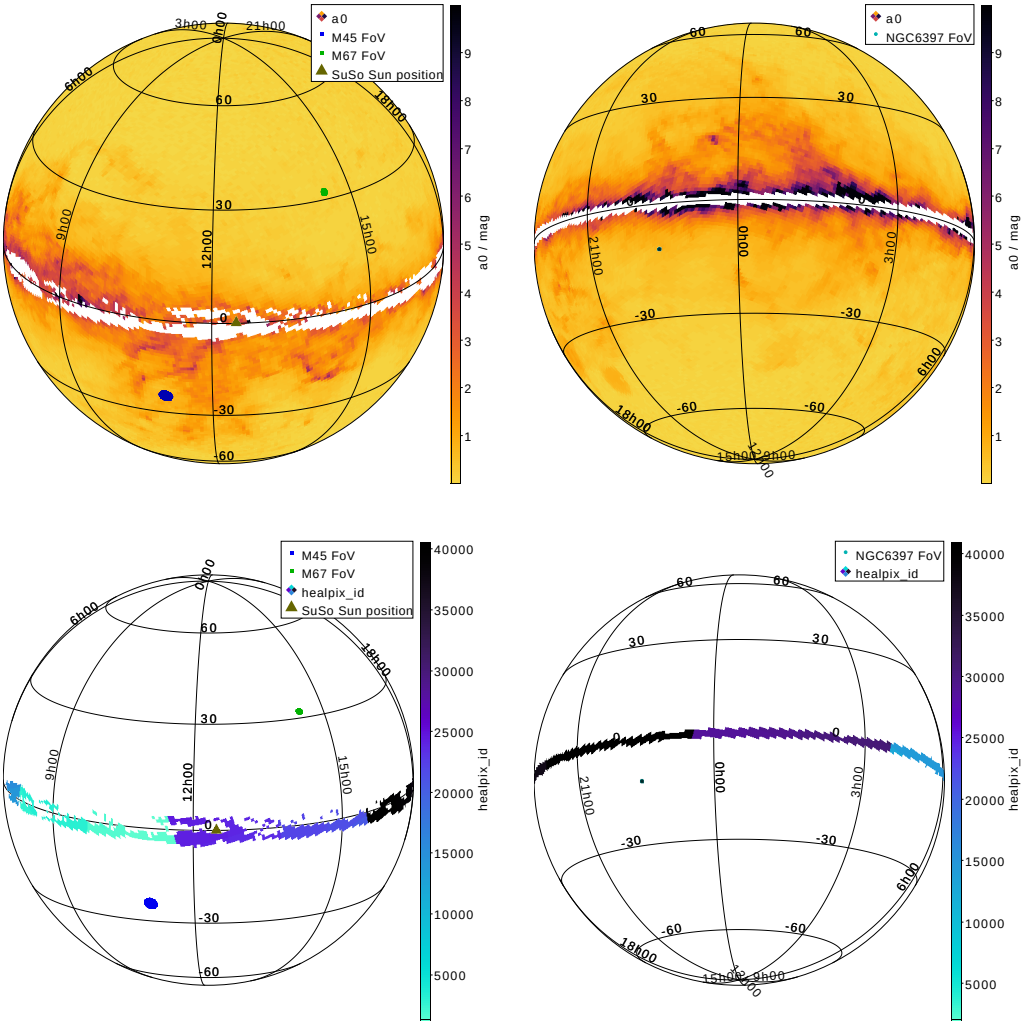


Figure 14: Top panels: Total Galactic Extinction maps of the sky at HEALPix level $n = 6$. They look towards the Galactic anticentre (left) and Galactic centre (right). The sky grid depicts the Galactic coordinates system, with the Galactic mid-plane (i.e., $b = 0°$) lying almost in the horizontal plane. Bottom panels: Sky maps highlighting the pixels with no a0 estimate, the colour coding being the pixel HEALPix index. The left panels also show the fields of view used to cone search for M45 and M67 (in blue and green, below and above the Galactic mid-plane respectively). Right panels: same for NGC6397 (cyan). Due to its small radial extent (0.35°), the NGC6397 field-of-view is barely visible. Its Galactic coordinates are $l \simeq 338°$ and $b \simeq -12°$.

square-ish equal-area pixels (Demleitner & Heinl, 2024b, page 8). The tessalation resolution is (somewhat indirectly) given by the parameter `healpix_level`. The lower `healpix_level`, the coarser the resolution and the larger the solid angle subtending the pixels covering the sky. In a HEALPix grid of level `healpix_level` $= n$, the pixel index ranges from 0 to 12×4^n; the pixels are built such that each pixel at level n contains 4 sub-pixels at level $n + 1$.

Look at the metadata of the `gaiadr3.total_galactic_extinction_map` table: its `healpix_level` ranges from 6 to 9 (see also Delchambre et al., 2023). A HEALPix level of 6 yields a sky coverage of $12 \cdot 4^6 = 49,152$ pixels. The angular area of the whole sky being

$$4\pi \,\mathrm{steradian} = 4\pi \left(\frac{180}{\pi}\right)^2 \simeq 41,253 \,\mathrm{deg}^2\,, \tag{13}$$

each level 6-HEALPix has an area of $41,253\,\mathrm{deg}^2/49,152 = 0.839\,\mathrm{deg}^2$.

The finest resolution admitted in Gaia DR3's extinction map at level 9 yields a sky coverage of 3,145,728 pixels and therefore a pixel area of $0.013\,\mathrm{deg}^2$.

8.2 Total Galactic Extinction (TGE)

The `a0` extinction provided by the `gaiadr3.total_galactic_extinction_map` catalogue table is the monochromatic extinction measured at 541.4nm (Delchambre et al., 2023). Extinction tracers are selected to be outside the dust layer of the Milky Way disc, such that measured extinctions are well-suited for extra-galactic studies. Because of their intrinsic brightness and large census beyond the Galactic-disc ISM, giant stars have been selected as extinction tracers.

To be assigned an `a0` value, a HEALPix pixel must contain at least three extinction tracers, whose median `a0` then defines the pixel `a0` value. At very low Galactic latitude, however, the number of available extinction tracers decreases as the greater distances covered by their line-of-sight imply higher extinctions. Therefore, low-latitude pixels are more likely to contain fewer than three tracers and to have no extinction value assigned to them. For a `healpix_level` $n = 6$, the sky map is complete only for $|b| > 5°$ (see Sec. 7 in Delchambre et al., 2023).

8.3 Getting and plotting the dataset

Let us go for a colourful TGE map of the sky, a map that will most naturally highlight the fundamental circle of the Galactic-coordinate system.

▷ **1** *Locating a service containing the extinction map* – In TOPCAT, again do VO → TAP, and if necessary go back to the Select Service tab. In the field Keywords, type `gaia dr3 extinction`. Hit Find Services. Open the `GAIA` sub-tree and double-click the entry for `gaiadr3.total_galactic_extinction_map` (or select the service and click on Use Service).

▷ **2** *Writing the ADQL query to download the a0 as a function of HEALPix index* – The complete level 6 map is small enough to simply download in full. So, run:

```
SELECT healpix_id, a0
FROM gaiadr3.total_galactic_extinction_map
WHERE healpix_level = 6
```

Hit Run Query; a new table should then appear in TOPCAT's table listing almost momentarily. Take note of the number of that table; you will need it in a moment.

▷ **3** *Inspect the downloaded dataset* – Double-click the new table entry to ascertain that the dataset indeed contains the two columns `healpix_id` (think of it as a 1D labeling of the 2D spherical coordinates) and `a0`.

A little reminder of what we saw in Sec. 3.2: Views → Column Statistics will help you confirm that the pixel numbering extends from 0 to 49,151. Note that `nGood`=49,152 for `healpix_id` but `nGood`=47,773 for `a0`. At `healpix_level = 6`, this difference corresponds to the 1,379 Galactic-disc pixels missing an `a0` value due to their low Galactic latitude (Sec. 8.2). Have another look at Figure 14 for a visualisation of these missing pixels.

Before going on, in the TAP window's table browser, view the column metadata of the `gaiadr3.total_galactic_extinction_map` and consider what additional columns you would need for which usage scenario.

▷ **4** *Configure plot axes to show Galactic coordinates* – In the Sky Plot window used for Fig. 13, untick the ecliptic SkyGrid layer (and the equatorial one if you happen to have one) as we now focus on the Galactic coordinate system. Follow Axes → Projection and select Galactic next to `View Sky System`.

▷ **5** *Doing the HEALPix plot* – To colour code the celestial sphere for extinction, in the SkyPlot window, navigate Layers → Add Healpix Control. Next to Table, select the table with the `a0` map (you just took note of its number) as your dataset and adjust the HEALPix Data Level to 6 (perhaps try a wrong HEALPix level to see the effect in case you ever inadvertently configure a bad level here; and yes, that you need to do this manually is the effect of missing interoperable metadata).

Set `healpix_id` as HEALPix index, Value is our `a0`. Also make sure that the Data Sky System is set on `equatorial`; only then will TOPCAT correctly transform to Galactic coordinates.

▷ **6** *Overplotting the clusters* – The plot still has the clusters we plotted above, but they are hidden by the opaque HEALPix map. To show them again, drag the layers of the M45 and M67 cone searches to the bottom of the layer stack in the Sky Plot window.

The top-left panel of Fig. 14 shows the result when looking into the direction of the Galactic anti-centre. Note that in contrast to Fig. 13, the Galactic mid-plane is now nearly horizontal.

▷ **7** *Inspect NGC6397 in the TGE map* – Rotate the sphere by 180° to look into the direction of the Galactic centre. The resulting TGE map is shown in the top-right panel of Fig. 14. Spot our pet globular cluster NGC6397, located under the Galactic mid-plane. Note also the higher extinctions characterizing the Galactic disc around the Galactic centre compared to the Galactic anticentre surroundings.

8.4 Visualizing pixels without TGE values

In the top panels of Fig. 14, pixels missing an `a0` value are left transparent. To better highlight these pixels – as discussed in Sect. 8.2, these contain fewer than three extinction tracers – let us create a table containing only these pixels.

▷ **8** *Get a data subset limited to the pixels with no a0 value* – In ADQL-queriable tables, missing values are marked with the magic value NULL. All "normal" comparisons with such a NULL evaluate to False.

```
SELECT healpix_id, a0
FROM gaiadr3.total_galactic_extinction_map
WHERE healpix_level = 6
  AND a0 IS NULL
```

Open the table this downloads and convince yourself that it actually consists of 1,379 rows whose `a0` (second) column is empty.

▷ **9** *Highlight the no-a0 pixels on a sky map* – Resume the previous Sky Plot. Get rid of the whole-sky TGE map by unticking it in the frame stack. Add a new HEALPix control layer to the stack in which to plot your newly-created subset. As above, select your new Table, but this time, configure the `healpix_id` itself as Value.

The result should look somewhat like Fig. 14. All of the 1,379 pixels missing an extinction value are located along the fundamental circle of the Galactic coordinate system (i.e., in the vicinity of $b = 0°$, Sec. 8.2). There are no exceptions.

Note that you could reach the same result without an extra download (but we wanted to remind you of the NULL magic). To try that, go back to your full HEALPix 6 table, in TOPCAT's main window do Views → Subsets, in the new Window Subsets → New Subset.

As Subset Name, type `empty`, as Expression use the somewhat funky `a0!=a0`. We told you above that NULLs compare unequal to everything including themselves, a property they, incidentally, share with the dreaded floating point NaNs.

That done, you can again add a HEALPix Control, plot the HEALPix id as the value and then choose `empty` as the Row Subset.

8.5 Moving the TGE map to higher resolution

The whole-sky maps in the top panels of Fig. 14 consist of 49,152 pixels, that is, the solid angle subtended by each pixel is $\simeq 0.8\,\mathrm{deg}^2$. This is about 4 times the area subtended by the full Moon and hence definitely not the finest resolution one can think of.

Let us therefore load an additional TGE map, one with a higher HEALPix level $n = 8$. You could go to $n = 9$, but since this is about 3 million pixels, we want to avoid that in a course situation. Feel free to try it yourself, though.

Back in the Table Access Protocol (TAP) Query window, we re-run the previous ADQL query for `healpix_level=8`:

```
SELECT healpix_id, a0
FROM gaiadr3.total_galactic_extinction_map
WHERE healpix_level = 8
```

This yields a table made up of $12 \times 4^8 = 786,432$ pixels (i.e., a 16-fold increase over what we had before. We shall build on it in Sec. 9.

9 Cluster Property Spatial Variations

Exactly as "tessalating" the sky allows you to characterize some physical property over it (e.g., in Sec. 8, the TGE), "tessalating" a cluster allows you to map how a property varies across the cluster. As a first example, let us highlight how the TGE varies across NGC6397 by cross-matching one of its subsets with the TGE sky map of Sec. 8.5. We will then address two cluster intrinsic properties: the star count per pixel and the mean stellar distance per pixel.

9.1 Matching a cluster dataset with a sky TGE dataset

To cross-match two tables, we use the TOPCAT tool Match Tables. Similarly to the ADQL `JOIN` instruction that allows one to merge data from two distinct catalogues/ser-

vices, Match Tables allows one to merge (part of) the data from two distinct tables (Taylor M., 2023).

The two tables we intend to match are the TGE dataset at HEALPix level 8 (Sec. 8.5) and the NGC6397 subset obtained via the ADQL query of Sec. 3.1, followed by the "tight" PM cleaning of Sec. 3.2 and the foreground removal of Sec. 3.3.1 (i.e., the cluster subset that here consists of 770 rows). The cross-match builds on the HEALPix index of both tables' pixels, thus implying that the HEALPix index of the cluster stars will have to be defined at some point.

▷ **1** *Crossmatch a cluster table and the TGE table* – In the TOPCAT window, click on the icon showing two ... matches (). A window entitled Match Tables pops up. There:

- In the field Algorithm under Match Criteria, select `Exact Values` so that pixel HEALPix indices from both datasets are matched exactly
- Under Table 1, select next to Table the relevant NGC6397 subset (or any other subset you wish to play with); we come back to the Matched Value column field in a moment
- Under Table 2, select next to Table your level-8 TGE table, and `healpix_id` next to Matched Value column
- Under Output Rows, next to Match Selection and Join Type, opt for Best match for each Table 1 row and All from 1, respectively; this will complete the cluster subset with the corresponding extinction values `a0` (rather than to complete the TGE map with the cluster data).

▷ **2** *Define the HEALPix index of the cluster stars* – Under Table 1, we still have to generate the HEALPix index of the pixels containing the cluster star positions. Fortunately, TOPCAT knows about computing HEALPix indexes from spherical coordinates.

So, in the field Matched Value column, type: `healpixNestIndex(8,ra,dec)`[6]. Knowing the pixel index of each star will straightforwardly allow Table 1 and Table 2 to be cross-matched, also yielding the TGE at each cluster-star location.

▷ **3** *Generate and inspect the matched table* – We are set. Hit Go to generate the new table. It is entitled `match(M,N)`, where M and N are the indices of the two matched tables (in the TOPCAT window, see the Table List numbering). The number of rows in `match(M,N)` is that of the dataset selected under Table 1 (since we selected *All from 1* above rather than *All from 2*). Compared to the cluster subset, `match(M,N)` contains 4 more columns, namely, `healpix_id` and `a0`, as expected, but also `GroupID` and `GroupSize`. We come back to the meaning of `GroupID` and `GroupSize` soon.

▷ **4** *Plot the match dataset* – Open a new Sky Plot window to mark the right ascensions and declinations of the cluster subset stars. To map the TGE `a0` across the cluster subset, add a HEALPix control layer (Layers → Add Healpix Control) and select the newly generated table `match(M,N)`. Assign `a0` to Value and adapt HEALPix Data Level if necessary ($n = 8$).

The left panel of Fig. 15 shows the NGC6397 subset along with the pixelated colour-coded TGE. NGC6397 being a halo GC, i.e. a cluster located beyond the Galactic disc, the Gaia-measured extinctions `a0` should be a good proxy of its foreground extinctions.

[6] If you expect to re-use HEALPix indexes like this, add them as a new synthetic column with just this expression.

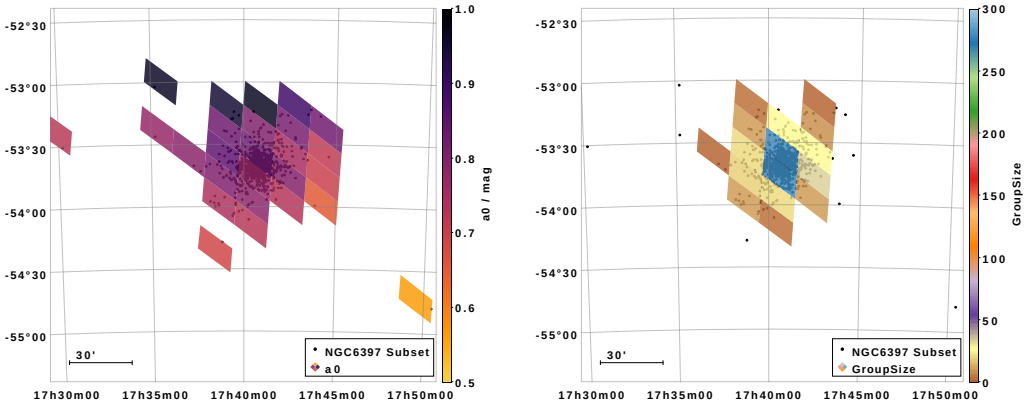


Figure 15: The NGC6397 subset (dots) obtained following "tight" PM cleaning and foreground removal (Sec. 3.2 and Sec. 3.3.1). Both panels are *Sky Plots* (i.e., the observer is outside the celestial sphere), not *Plane Plots* (note the slight curvature of both coordinate grids). Coordinates are right ascension and declination. Left panel: The cluster subset is overlaid with level-8 HEALPix pixels, colour-coded for the extinction parameter A_0. Right panel: same but with the colour-coding depicting GroupSize, the number of stars per pixel, obtained after cross-matching the cluster subset with the TGE dataset. Also note the sharp drop of the star count: from the two central pixels to the cluster outer regions, pixel colours jump from the top to the bottom of the colour palette. Pixels containing one source only are not highlighted.

9.2 Cluster stellar-density variations

An intrinsic property of a cluster that most strikingly varies across its field-of-view is its stellar density. To colour-code HEALPix pixels for the cluster stellar density, all we need is the number of stars per pixel since all HEALPix pixels have the same area.

9.2.1 From the cluster TGE map to the cluster star-count map

A few mouse clicks will turn the TGE map of Fig. 15 (left panel) into a star-count or stellar-density map. Back to our match table `match(M,N)`, let us consider its last two columns, namely, `GroupID` and `GroupSize`.

▷ **5** *Discover what GroupID and GroupSize are* – Open the column metadata of your match table to discover that GroupID is a *ID for match group*. That is, each pixel constitutes a match group of stars to which a GroupID is assigned (to convince yourself, you could make a plane plot of GroupID versus healpix_id and see there is a one-to-one match between both of them).

▷ **6** *Adjust the physical quantity to be displayed by your colour-coded map* – Back to the Sky Plot window, next to Value, swap `a0` for `GroupSize` to highlight the star count per pixel across the NGC6397 subset. The result is shown in the right panel of Fig. 15, where the colour-coding scheme has changed from `Inferno` to `Paired`. You would do this in Aux Axis → Map → Shader). Update the colour-palette range as well (Aux Axis → Range).

9.2.2 The SkyDensity option of the Position Control layer

The approach above builds on the `match(M,N)` table and, therefore, on the sky TGE map of Sect. 8.5, which kept us tied to its HEALPix level 8. If stellar densities and stellar counts are all you care about, go for the `SkyDensity` option of the Position Control layer. There, the HEALPix resolution can be freely adjusted and there is no need to define the HEALPix index of the pixels containing your stellar subset. All you need is a list of stellar coordinates, i.e., a dataset containing α, δ coordinates of stars.

To see the workflow, start a brand-new Sky Plot window and pick a stellar dataset that was *not* cross-matched with a TGE map. To do that, select the NGC6397 table we obtained in Sect. 3.1.

▷ **7** *Set up a sky density plot* – Do Graphics → Sky Plot in TOPCAT's main window, select the Form tab and then Forms. In the drop-down menu, select Add Sky Density[7].

▷ **8** *Vary the density plot* – Move your eyes to the right and, under Global Style, next to Combine, choose what you prefer to plot: star count (count) or stellar density (count per unit)?[8] In ADQL language, by the way, what TOPCAT offers you here are aggregate functions.

For our purposes, we opt for the star count to compare the new map to that already presented in Fig. 15. We therefore adjust HEALPix Level to `Abs` and 8 as well.

Although obtained differently, your new map (left panel of Fig. 16) and the previous one (right panel of Fig. 15) should be identical, apart from one detail: pixels containing one single star are now colour-coded as well.

▷ **9** *Vary the resolution of your Sky Density map* – Try varying the HEALPix level. At level 7, the pixel size is of order the full Moon area (check it). at level 6, one pixel becomes enough to cover almost all of the cluster subset.

In the right panel of Fig. 16, the tessalation resolution has been increased up to Level 9; notice how many pixels outside of the cluster centre are then unpopulated.

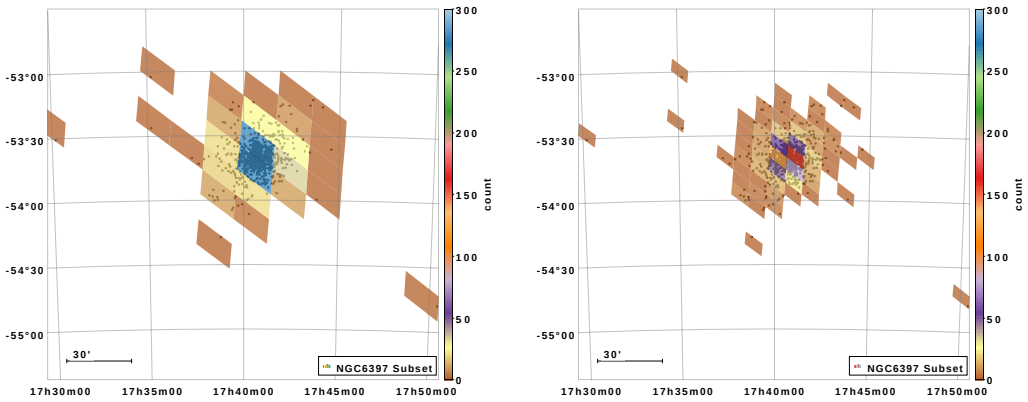


Figure 16: Left panel: Same as the right panel of Fig. 15 but the star count per pixel is here provided by the SkyDensity tool of the Position Control layer. The HEALPix level is 8. Right panel: same at HEALPix level 9 (i.e., each pixel of the left panel gets split into four subpixels).

[7] This should remind you of what we did to overlay Gaussians onto histograms in Sec. 3.3.1

[8] Regardless of your choice, keep in mind that your stellar subset represents but a fraction of the whole cluster population (less than 1 % in the case presented here).

9.3 What about other cluster properties?

We have so far restricted our cluster mapping to the TGE and star-count variations. But what about other cluster star properties? What about comparing the behaviour of the cluster stars with the behaviour of stars present in the cluster field of view but not belonging to the cluster? We can use ADQL's power to compute aggregate functions on the server using uploads.

The server-side equivalent of TOPCAT's healpixNestIndex function is the `ivo_healpix_index` ADQL "user defined" (actually, extension) function: It computes the HEALPix index (α, δ) coordinates (see the ADQL tab in the Use Service tab in TOPCAT's TAP window for documentation on what functions a service support).

Against our aggregation on uploads – which is more educational than anything else –, you will much more typically use aggregate functions before downloading anything; doing server-side aggregation greatly reduces the amount of data to transfer. To do a TAP upload, reference a magic table `TAP_UPLOAD.tx`, whose final "x" is to be substituted with the index of the table in TOPCAT's table list (in the control window). Try this query, adapting the table index to your situation:

```
SELECT ivo_healpix_index(9,ra,dec) AS hpx9,
COUNT (*)                          AS Star_Count_per_Pixel,
AVG(r_med_geo)                     AS AVG_Distance_per_Pixel
FROM TAP_UPLOAD.t1
GROUP BY hpx9
```

What is going on here?

- `TAP_UPLOAD.t1` references the uploaded table
- `ivo_healpix_index(9, ra, dec) AS hpx9` converts the equatorial coordinates of each object to a level-9 HEALPix and makes it available under the name hpx9 (without the `AS`, this would get an unintelligble name, and you could not reference the healpix).
- `GROUP BY hpx9` builds bins of objects that ended up in the same healpix (i.e., for which hpx9 has the name value). All aggregate functions now work per such bin.
- `COUNT (*) AS star_count_per_pixel` is such an aggregate function, yielding the number of objects in the bin.
- `AVG(r_med_geo) AS avg_distance_per_pixel` is another aggregate function application, this time taking the average of the distance in each bin.

After you run this query, back in the TOPCAT control window, double-click the newly created table to see that it contains the expected three columns:

- `hpx9`,
- `star_count_per_pixel`, and
- `avg_distance_per_pixel`.

Double-check its row number, too. Since each row corresponds to one HEALPix, the number of rows returned depends on the spatial extent of the objects you uploaded

(and their distribution; you will not get back aggregate rows for HEALPixes in which no object exists) in the ADQL query, not the number input rows.

For instance, with a (fully populated) table with a diameter of 4°, you will get 1'014 rows. For our cluster subset after PM cleaning and foreground removal, the field of view is much smaller, and you will only get something like 56 result rows.

Try it: Run the ADQL query for both the full (raw) dataset and a cluster-restricted subset; simply replace the table index in the `TAP_UPLOAD.tn` table reference. Incidentally, in TOPCAT's control window, you can configure a Row Subset; TAP uploads will honour that setting.

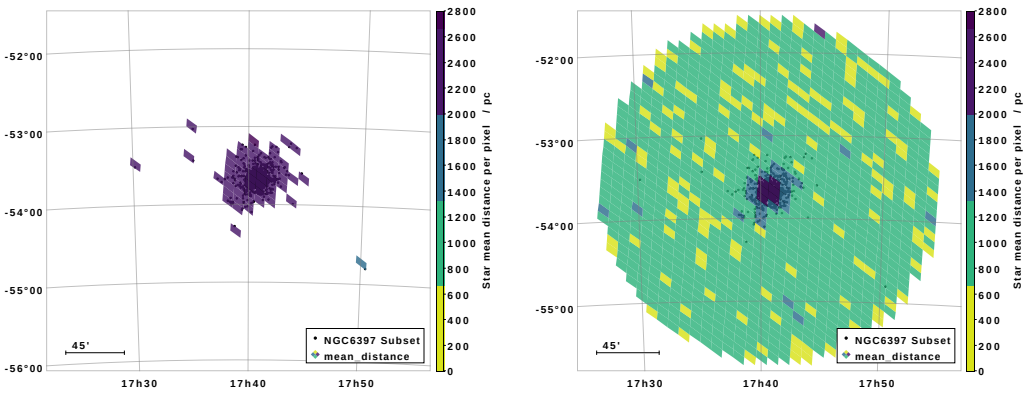


Figure 17: The NGC6397 subset (dots) obtained following "tight" PM cleaning and foreground removal (Sec. 3.2 and Sec. 3.3.1). Left panel: Tessalated view at HEALPix level 9 of the cluster subset. Each pixel is colour-coded for the mean distance of the stars it contains. Right panel: Tessalated view at HEALPix level 9 of the full dataset the subset was drawn from (colour coding as in left panel). But for its cluster-dominated centre, most of the 4°-wide field-of-view is occupied by foreground stars of NGC6397 (compare with the histogram of Fig. 5).

Figure 17 shows how the pixel-averaged stellar distance varies across the NGC6397 subset (left panel[9]) and across the 4°-wide field-of-view covered by the ADQL query on page 5 (right panel).

In the left panel, stellar distances averaged per pixel span a fairly uniform range, from $\simeq 2,000$ pc to $\simeq 2,800$ pc. These stars correspond to the "`PM_Cleaning: tight`" histogram of Fig. 5 minus the 26 foreground contaminants. They thus define a subset of the NGC6397 stars. But do keep in mind how much error-contaminated their distance *estimates* are. The black histogram of Fig. 6 reminds you that the cluster actual extent is much smaller than the $\simeq 800$ pc-wide interval suggested by the solid (yellow) Gaussian of Fig. 6 and by the uniform pixel colour-coding in Fig. 17 left panel. We come back to the cluster actual extent in Sec. 10.

The right panel of Fig. 17 shows as symbols the same cluster subset (770 stars), overlaid with the pixelated whole dataset (12,513 stars; pixel colour coding as in left panel). The right-panel pixelated view is thus the counterpart of the `All` histogram of Fig. 5, both illustrating that most stars across the 4°-wide field-of-view are stars in the foreground of NGC6397.

The right panel of Fig. 17 also suggests that NGC6397 – or more exactly the stellar subset we have isolated – sits in the field-of-view innermost regions. Can we estimate the

[9]Of course, you could also map again the star counts and check that your new map is identical to the right panel of Fig. 16.

actual radial extent of NGC6397? Since stellar distances measured along the line-of-sight are heavily affected by errors (at the distance of NGC6397, even a less than 10 % relative error is much larger than the cluster size; see Sec. 3.3 and Fig. 6), we will consider the star distances measured from the cluster centre, in projection in the plane of the sky, and plot their cumulative radial distribution.

10 The Actual Extent of NGC6397

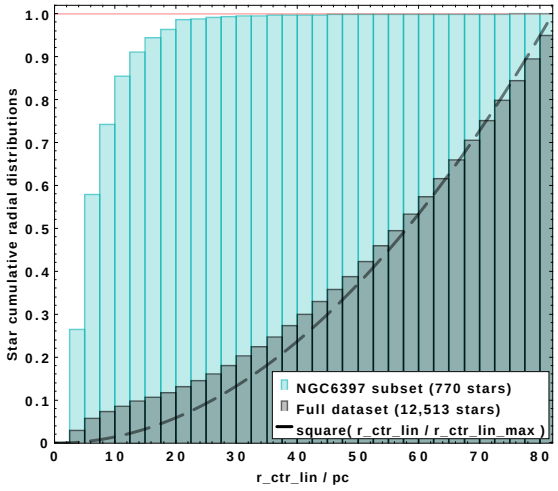


Figure 18: Cumulative radial distributions normalised to their upper limit for the full dataset obtained from the ADQL query in sect. 3.1 (12,513 stars) and for the NGC6397 subset obtained following PM cleaning and foreground removal (770 stars). The thick dashed line depicts the expectation for a sky uniform stellar density. The x-axis extent (82pc) corresponds to the radius of the searched field-of-view (2°) at the NGC6397 estimated distance of 2.36 kpc

To estimate the actual radial extent of NGC6397, we plot the cumulative radial distribution of our stellar subset in the plane of the sky, that is, the cumulative histogram of the `r_ctr_lin` parameter introduced in Sec. 3.3.3.

▷ **1** *Start a histogram plot of the NGC6397 subset* – In TOPCAT's control window, select the result of the original query again. Click Graphics → Histogram Plot, use `r_ctr_lin` as the X coordinate. Recall that `r_ctr_lin` is the linear projected distance between stars and cluster centre (i.e., the distance you used to build the black histogram of Fig. 6).

▷ **2** *Switch to a normalised cumulative distribution* – To turn the histogram into a cumulative distribution, select the Bins dialog, and in it the General tab. In it, check forward in the Cumulative line. To normalise the distribution (which we want for later comparsion), select `Maximum` next to Normalise.

▷ **3** *Overplot the cleaned distribution* – Repeat the sequence, this time configuring the tightly cleaned subset in the Subsets tab.

▷ **4** *Plot the expectation of a uniform distribution* – If the stars in our field were uniformly distributed, the density would rise with the area covered, i.e., quadratically. With the normalisation we used, this means the expectation is $(r_{\rm lin}/2°)^2$.

Add this expectation to the plot by doing Layers → Add Function Control, and

then configuring Independent Variable Name as `r_ctr_lin` and the function expression as `square(r_ctr_lin/2)`.

This should result in something like Figure 18. The x-axis runs up to the edge of the field-of-view, that is, 2° or 82 pc at the adopted cluster distance of 2.36 kpc.

In the field-of-view outer regions (`r_ctr_lin`$\gtrsim$ 50pc), the full-dataset distribution is well described by the expectation for a uniform stellar density on the sky (thick dashed line). In the field-of-view inner regions, this expectation underestimates the actual distribution owing to the cluster contribution.

The cumulative distribution of the cluster alone saturates at a distance of $\simeq$ 35pc from its centre, which could thus be considered an estimate of its outer boundary. However, this boundary estimate is only a lower limit owing to the nature of our stellar subset. Go back to Fig. 8 to remind yourself that our NGC6397 stellar subset consists mostly of subgiant and red-giant stars, that is, stars more massive than main sequence stars. As star clusters age, higher-than-average mass stars sink toward the cluster inner regions, a phenomenon that stellar dynamics refers to as *mass segregation*. Giant stars alone therefore yield an underestimated cluster boundary. The actual outer boundary of NGC6397 as a whole is estimated as $\simeq$ 50 pc by Baumgardt et al[10]. This is its tidal radius.

The radius of the field-of-view searched by the ADQL query on page 5 was chosen to be 2° with the purpose of covering NGC6397 in full.

Acknowledgements

GP acknowledges funding by the Bundesministerium für Bildung und Forschung (BMBF: German Ministry of Research) – Project-ID e-Inf-Astro 3 / 05A23VH4 (PI: Prof. Dr. J. Wambsganss). She is also most grateful to Mrs. D. Schwalbe, Head of the Heidelberg *Astronomisches Rechen-Institute* Administration, for having initiated her e-Inf-Astro employment contracts.

References

Delchambre, L., Bailer-Jones, C. A. L., Bellas-Velidis, I., et al. 2023, A&A, 674, A31. doi:10.1051/0004-6361/202243423

Demleitner, M. 2017, Adding catalog data to object lists using the VO – http://doi.org/10.21938/LvJ43o1bcSnFO94vnnsqQA

Demleitner, M., Heinl, H. 2024a, Explore the Pleiades with TOPCAT and Aladin http://www.g-vo.org/tutorials/pleiades.pdf

Demleitner, M., Heinl, H. 2024b, A short course on ADQL - https://docs.g-vo.org/adql/notes.pdf

Melis, C., Reid, M. J., Mioduszewski, A. J., Stauffer, J. R., & Bower, G. C. 2014, Science, 345, 1029 - 2014Sci...345.1029M.

Taylor, M., 2023, National Astronomy Meeting: TOPCAT and Gaia DR3, slide 10 - https://www.star.bristol.ac.uk/mbt/talks/nam2023/tcgaia.pdf.

[10] https://people.smp.uq.edu.au/HolgerBaumgardt/globular/parameter.html

Find me on VO Text Treasures http://dc.g-vo.org/VOTT



This document is in the public domain.