Astronomy & Astrophysics

Gaia colour-magnitude diagrams of young open clusters

Identification in the UBC catalogue and comparison between manual and automated analyses[⋆]

I. Negueruela[1,2] and A. de Burgos[3,4]

[1] Departamento de Física Aplicada, Facultad de Ciencias, Universidad de Alicante, Carretera de San Vicente s/n, 03690 San Vicente del Raspeig, Spain
e-mail: ignacio.negueruela@ua.es
[2] Instituto Universitario de Investigación Informática, Universidad de Alicante, 03690 San Vicente del Raspeig, Spain
[3] Instituto de Astrofísica de Canarias, c/ Vía Láctea s/n, 38200 La Laguna, Tenerife, Spain
e-mail: abel.burgos@iac.es
[4] Universidad de La Laguna, Departamento de Astrofísica, 38206 La Laguna, Tenerife, Spain

ABSTRACT

Context. Automated analyses of *Gaia* astrometric data have led to the discovery of many new high-quality open cluster candidates. When a good determination of their parameters is available, these objects become excellent tools for investigating the properties of our Galaxy.
Aims. We explore whether young open clusters can be readily identified from *Gaia* data alone by studying the properties of their *Gaia* colour-magnitude diagrams. We also want to compare the results of a traditional cluster analysis with those of automated methods.
Methods. We selected three young open cluster candidates from the UBC catalogue, ranging from a well-populated object with a well-defined sequence to a poorly populated and poorly defined candidate. We obtained classification spectra for the brightest stars in each. We redetermined members based on EDR3 data and fitted isochrones to derive the age, distance, and reddening.
Results. All three candidates are real clusters with ages below 100 Ma. UBC 103 is a moderately populous cluster, with an age around 70 Ma. At a distance of ∼3 kpc, it forms a binary cluster with nearby NGC 6683. UBC 114 is a relatively proximal (∼1.5 kpc) poorly populated cluster containing two early-B stars. UBC 587 is a dispersed and very young (≤10 Ma) cluster located at ∼3 kpc, behind the Cygnus X region, which may be a valuable tracer of the Orion arm.
Conclusions. The `OCfinder` methodology for the identification of new open clusters is extremely effective, with even poor candidates resulting in interesting detections. The presence of an almost vertical photometric sequence in the *Gaia* colour-magnitude diagram is a safe way to identify young open clusters. Automated methods for the determination of cluster properties give approximate solutions, but these are still subject to difficulties. There is some evidence suggesting that artificial intelligence systems may systematically underestimate extinction, which may impact in the age determination.

Key words. stars: evolution – Hertzsprung-Russell and C-M diagrams – open clusters and associations: general – Galaxy: structure

1. Introduction

Open clusters are valuable natural laboratories, allowing for the study of stellar evolution, Galactic structure, and chemical gradients, among other current topics of interest. The advent of *Gaia* astrometric and photometric data (Gaia Collaboration 2018) has resulted in substantial advances in our knowledge of open clusters. Current astrometric data allow for a much better definition of membership than any previous datasets, leading to stronger constraints on stellar evolution, while future releases will (in all likelihood) improve our understanding of dynamical processes in these stellar systems.

In addition, the wealth of high-quality data has resulted in the discovery of many new cluster candidates which, if confirmed, could lead to the doubling of the number of clusters known. Most of these cluster candidates have been found by means of automated methods, which use a variety of strategies, from visual searches (e.g., Ferreira et al. 2020) and super-

vised explorations of regions where a result is anticipated (e.g., Maíz Apellániz et al. 2020a) to friends-of-friends algorithms (e.g., Liu & Pang 2019) and machine-learning Bayesan algorithms (e.g., Hao et al. 2020; He et al. 2021) over large regions of the sky. Until now, the most successful exploration by far has been carried out at the University of Barcelona by using the `OCfinder` method (Castro-Ginard et al. 2019). Based on DR2 data, Castro-Ginard et al. (2020) presented almost 600 high quality new[1] open cluster candidates, representing an increase by more than one-third in the number of open clusters known. An even higher number was recently discovered in EDR3 data (Castro-Ginard et al. 2022), while other authors have reported similar numbers of new discoveries (e.g., Hao et al. 2022; He et al. 2023). These new cluster candidates discovered in *Gaia* data now rival in number the clusters known before its launch. In most cases, however, the candidates have not been validated with other techniques. Here, we will concentrate on the candidates found by the `OCfinder` method. As these are named

---

[1] Although a few of them can be identified with previously named, but unstudied, clusters, the vast majority represent new detections.

UBC (University of Barcelona Cluster) followed by a sequential number, we will collectively refer to them as UBC clusters.

Beyond the value of discovery, the new clusters may be used to study the structure, kinematics, and/or chemical properties of the Milky Way, depending on their age and the number and nature of their members. Thus, the new large catalogues of cluster candidates only become truly valuable when the basic parameters of the clusters are estimated. To handle such large numbers of objects, methods based on machine-learning algorithms are almost mandatory. As an example, Cantat-Gaudin et al. (2020) estimated parameters in a homogenous way for a very large sample of Milky Way open clusters, which included a sizable fraction of the DR2 UBC clusters (Castro-Ginard et al. 2020). Their method is based on comparison to clusters of known parameters and, in most cases, results in a fair agreement with previous determinations. Its accuracy and the possibility of systematic effects have not yet been assessed.

Independently of their age and nature, the reasons why these new clusters had not been found before are easy to identify (in most cases). Some lie in areas of very high stellar density (cf. Ferreira et al. 2020; or see the case of Valparaiso 1 = UBC 106, independently (and serendipitously) discovered by Negueruela et al. 2021). Although Castro-Ginard et al. (2022) have concluded that their method preferentially identifies objects of limited size in the sky, most are dispersed groups, with no obvious central concentration that might appear to the eye. Among these, some are low-density groups at moderately large distances, which may more readily be identified as young associations than open clusters; for instance, there is UBC 609, which includes the previously known compact clusters Waterloo 1 and Meyer 2, as well as many stars in the surrounding area.

Many other UBC clusters are relatively nearby, small groups of intermediate age that are likely remnants of clusters that have already lost all their massive stars and are in the process of dissolution. A typical example of this kind of candidate may be UBC 94, whose *Gaia* colour-magnitude diagram (CMD) is shown in Fig. 1 for illustration. With their machine-learning approach, Cantat-Gaudin et al. (2020) derived a distance modulus $DM = 10.5$ mag (driven by the *Gaia* DR2 parallax), with $A_V = 1.32$, and an age of 280 Ma, but it is difficult to evaluate the goodness of such a fit, because the CMD is completely featureless. There is not a single evolved star and no guide to the position of the turn-off. As an example, we include a 150 Ma isochrone, with $A_V = 1.7$, which provides an equally valid (if not better) "fit" to the data for the same distance. Indeed, younger ages could also be accommodated simply by assuming that the upper main sequence is not populated, given the low number of members. The *Gaia* CMD on its own does not contain enough information to determine an age. In this particular case, the three brightest stars are sufficiently bright to be included in catalogues. Two of them have HD classifications of A0 and A2, which supports the older age, even if we assume that they are giants. However, such information is not available in most other cases, leading to an indetermination in the cluster properties. An artificial-intelligence algorithm, on the other hand, will always deliver an age for the cluster.

As discussed by Cantat-Gaudin et al. (2020), the presence of red clump stars breaks the degeneracy between age and reddening and offers a quick and easy way to correctly assign ages. When no such stars are present, age determinations are more difficult. Unfortunately, this is generally the case not only for poorly populated clusters, but also for young clusters, which are of particularly high value, as they can be used to trace recent star formation and provide information on Galactic structure that
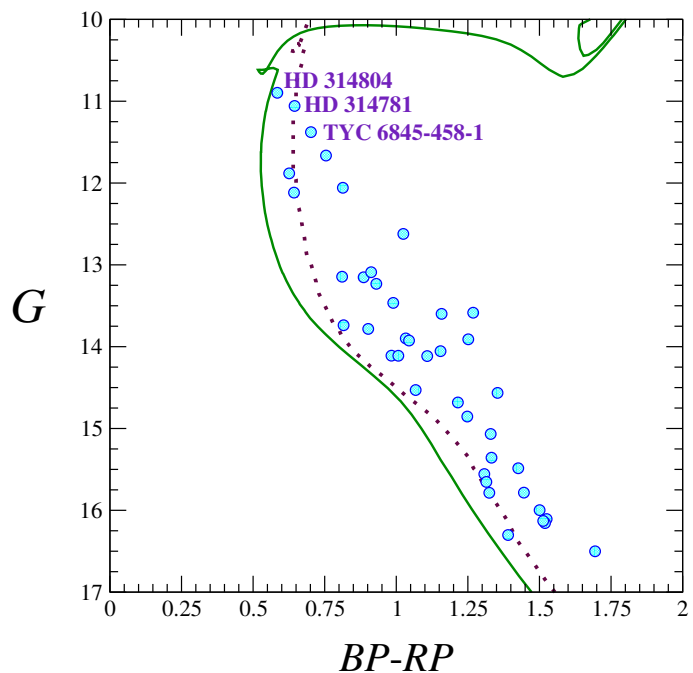
Fig. 1. *Gaia* DR2 CMD for high-probability members of a "typical" UBC cluster, UBC 94. For reference, catalogued stars are named in the diagram. The green solid line represents the preferred isochrone fit according to Cantat-Gaudin et al. (2020, 280 Ma, $A_V = 1.32$). The purple dotted line is an alternative fit with a 150 Ma, $A_V = 1.7$ isochrone. The lack of evolved stars allows for very different solutions.

can be used to test different models for the formation of spiral arms (e.g., Castro-Ginard et al. 2021). Moreover, the most luminous stars in young clusters will be OB stars, which are of a high value in themselves, as probes of stellar evolution. In this paper, we explore the possibility that young open clusters can be identified among the new candidates by the shape of their *Gaia* CMDs alone. In Sect. 2, we discuss the features that allow the identification of young open clusters in a *Gaia* CMD. In Sect. 3, we describe how we selected three young cluster candidates from the UBC database based on the shape of their CMDs and obtained spectra of their brightest members. We then derive parameters for these three clusters and proceed to a full characterisation in Sect. 4, along with a discussion of the limitations of photometric data.

2. Identification of young clusters in the *Gaia* CMD

For the sake of simplicity, in this work, we mostly use the PARSEC isochrones (Chen et al. 2015) retrieved from the CMD 3.6 interface[2]. These isochrones correspond to non-rotating models with assumed solar metallicity $Z_\odot = 0.015$. They were generated with the passbands from Maíz Apellániz & Weiler (2018) and reddened automatically by assuming a Cardelli et al. (1989) extinction law, which may be inadequate at high reddening values because of its polynomial nature; however, in most cases, it is sufficiently accurate for moderate reddening.

To illustrate the shape of a young open cluster in *Gaia* photometry, Fig. 2 shows a 30 Ma isochrone affected by different amounts of extinction. As can be seen, the shape of the isochrone is not significantly modified by reddening. The
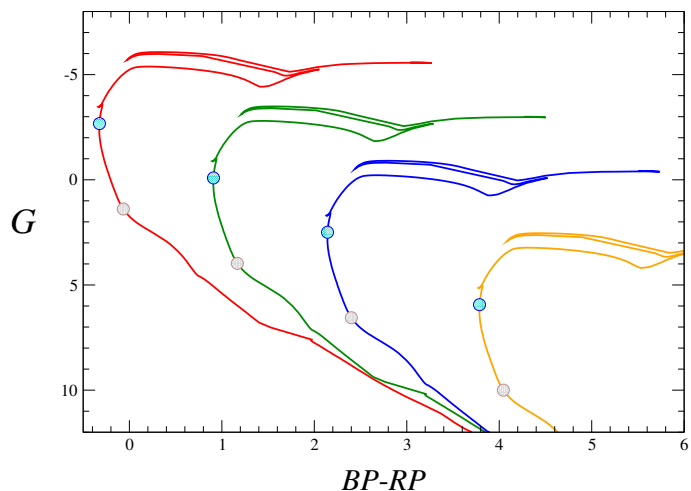
---

[2] http://stev.oapd.inaf.it/cgi-bin/cmd_3.6

isochrone presents three well-differentiated parts. The lower main sequence (MS) runs diagonally through the diagram, close to a straight line. An important feature of the *Gaia* photometric system is that reddening moves stars roughly along the same line, so that the different isochrones end up running together. Because of this, determining the extinction (and, hence, the age) for an intermediate-age cluster with no evolved stars is almost impossible without additional information, as illustrated above by the case of UBC 94. The upper main sequence is much more vertical. In all the isochrones shown in Fig. 2, we have marked by circles the position of stars with $8\,M_\odot$ (a B2 V star at $Z_\odot$) and $2.3\,M_\odot$ (around A0 V). The almost vertical stripe thus corresponds to the position of B-type stars. The change in slope between these two sections, at the position of early A-type stars, is perhaps the most easily identified feature in the MS of well-populated clusters. The post-MS evolution, with the almost horizontal track towards the red and the blue loop, is very well defined, but this region is (in most cases) very poorly populated.

These features are illustrated in Fig. 3, where two well-studied and well-populated clusters are shown as an example of the shape of a young cluster CMD in *Gaia* photometry. The left panel shows the CMD for the young cluster NGC 457. We selected objects with membership probability $P \geq 0.5$ in Cantat-Gaudin et al. (2018). Given its moderate, uniform reddening and the absence of significant foreground or background population (cf. Negueruela 2017), NGC 457 is a very good template for the shape of a real CMD. The exquisite quality of *Gaia* photometry is manifest in the narrowness of the upper main sequence. The turn-offs of young open clusters are in general ill-defined, because of the combined effects of binary evolution (which generates blue stragglers) and a range of initial rotational velocities (which leads to broadened sequences). In fact, all the bright objects falling to the right of the sequence are known binaries or Be stars. We must point out that there are a few objects, all fainter than $G = 12$, whose position is incompatible with the isochrone, which we consider to be astrometric interlopers (same astrometric parameters as the cluster, but not related)[3]. Ultimately, these objects should be ignored.

By assuming a distance modulus $DM = 12.2$, after Cantat-Gaudin et al. (2020), the sequence is very aptly fit by a 20 Ma isochrone reddened by $A_V = 1.5$, parameters that are directly comparable to those given by Cantat-Gaudin et al. (2020), who found the same age and a slightly lower extinction of $A_V = 1.42$. The age is determined by resorting to the location of the turn-off and the red supergiant HD 236697. We must note that two other cluster supergiants do not fit the isochrone. One is the B5 Iab supergiant HD 7 902, seen well above the isochrone in Fig. 3. The other is the F0 Ia standard ϕ Cas, too bright to have *Gaia* photometry. This discrepancy is discussed in Negueruela (2017), where a fit of Geneva (Ekström et al. 2012) isochrones with $A_V = 1.55$ and $DM = 12.0$ (i.e., parameters compatible within errors) to UBV photometry returns the same age estimate, demonstrating a rather high degree of consistency between different models and observed datasets.

Not all real cluster sequences are as well defined as that of NGC 457. In many cases, differential reddening contributes as significantly as intrinsic effects to the broadening of the stellar sequence. As an example, the right panel of Fig. 3 shows the older cluster IC 4725, which spans well over one degree on the

---

[3] The number of such interlopers will partly depend on the richness of the field, but also on how close the cluster astrometric parameters are to those of the surrounding field. See the case of UBC 103, discussed in Sect. 4.1.1.

Fig. 2. Padova PARSEC isochrones for 30 Ma and solar metallicity, affected by A_V (from left to right) of 0, 3, 6, and 10 mag. The circles mark the position of stars with initial masses of $8\,M_\odot$ (top, blue) and $2.3\,M_\odot$ (bottom, grey).

sky in an area of variable obscuration. Accepting the distance modulus $DM = 9.2$ from Cantat-Gaudin et al. (2020), the best fit is obtained by an 85 Ma isochrone reddened by $A_V = 1.4$. These parameters are very different from the 110 Ma and $A_V = 0.84$ given by Cantat-Gaudin et al. (2020). Our younger age is determined by the location of the two yellow supergiants on the blue loop. The reason for the much higher extinction is obvious from Fig. 3, where we have also plotted the isochrone favoured by Cantat-Gaudin et al. (2020). While our value of the reddening is driven by the bulk of the stars (and could indeed be even a little higher), the value given by the automated analysis represents a lower envelope for the set, determined by a few outliers with low extinction. In any event, the broadening of the upper main sequence still preserves the shape of the sequence and the inflection in its slope is perfectly noticeable, driving the right combination of distance modulus and reddening.

3. Observations and astrometric data

3.1. Sample selection

Three representative clusters were selected from the UBC database by resorting to a visual inspection. The selection required the clusters to contain at least a moderate number of stars brighter than $BP = 13$ (so that spectra could be obtained with not very long exposure time), display an approximately vertical sequence in the CMD, which we have shown to be the identifying feature of young open clusters, and have good visibility in the early summer.

The raw CMDs of the clusters selected are displayed in Fig. 4. The three chosen clusters cover a variety of situations that can be found among the new cluster candidates. UBC 103 (left) is a well populated cluster, with a moderately broad main sequence, hinting at differential extinction. It is thus classified as an A category (easily identifiable) cluster by Castro-Ginard et al. (2020). Its slope inflection is clearly seen around $G \approx 15$. UBC 114 (centre) is not so well populated, but has a narrow, very well-defined sequence, which still permits its classification as an A category cluster in Castro-Ginard et al. (2020). The slope inflection can be guessed slightly above $G \approx 14$, but the number of stars above it is so small that the identification as a young
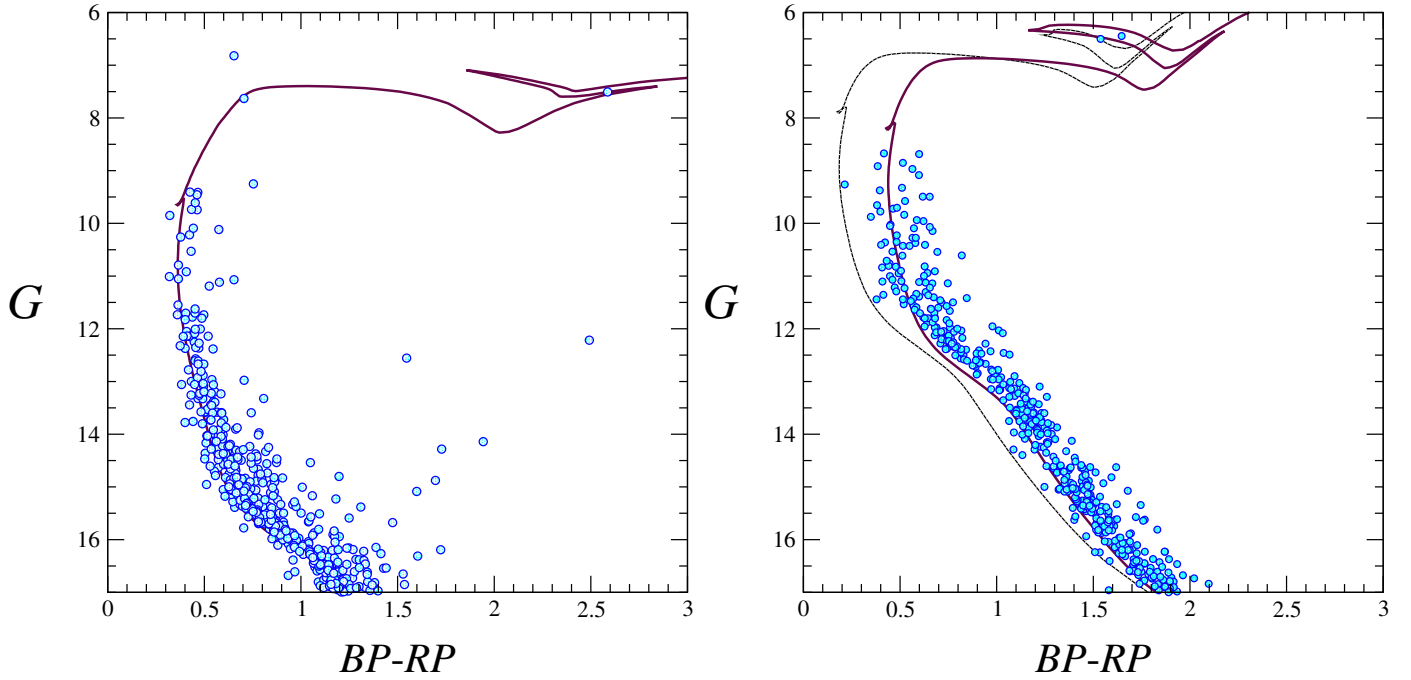
Fig. 3. Padova Parsec isochrone fits for the *Gaia* DR2 CMDs of two well-studied young clusters. Left: NGC 457 fit with a 20 Ma solar metallicity isochrone. Right: IC 4725 fit with an 85 Ma solar metallicity isochrone (the dotted line shows the isochrone proposed by Cantat-Gaudin et al. 2020). See main text for details.
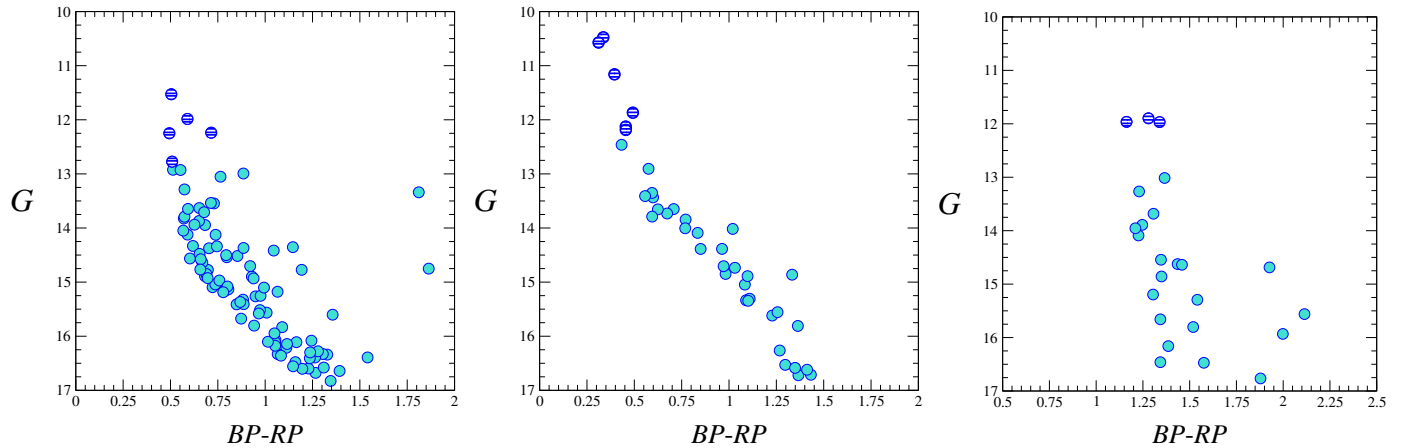
Fig. 4. Raw CMD diagrams for the three clusters observed, based on *Gaia* photometry. From left to right: UBC 103, UBC 114, and UBC 587 (noting the shift in the horizontal axis due to higher extinction). The large striped circles indicate the stars whose spectra were collected. Cluster members are from Cantat-Gaudin et al. (2020). See main text for details.

cluster is uncertain. UBC 587 (right) is a very poorly populated cluster. The stellar sequence is broad and ill-defined, leading to its classification as a C category (doubtful) cluster candidate in Castro-Ginard et al. (2020). The sequence, however, is decidedly vertical over five magnitudes, suggesting a very young cluster.

3.2. Observations

Spectroscopic observations were taken using the Alhambra Faint Object Spectrograph and Camera (ALFOSC), attached to the 2.56 m Nordic Optical Telescope (NOT) at the Observatorio del Roque de los Muchachos (La Palma, Spain) of the Instituto de Astrofísica de Canarias. The instrument was equipped with CCD #14, a 2048 × 2064 e2v camera. The configuration used consists of grism #18 and a 0″.75 long slit. Grism #18 provides

a wavelength range of 345–535 nm, with a central wavelength at 436 nm and an approximate resolving power of $R \sim 2000$.

With this configuration, we gathered 19 spectra of different sources selected from the three clusters described above. Observations were taken as part of the NOT Fast-Track program during two consecutive bright nights on June 6 and 7, 2020. A summary of the target properties and exposure times used can be found in Table 1. Weather conditions were very good, with a seeing of $\sim$0″.9 measured on the acquisition images. We aimed at a signal-to-noise ratio (S/N) of $\sim$150 in the 500–600 nm range.

The reduction was carried out using Python programming language and standard tasks within the IRAF software environment. The full reduction of the spectra included bias and flat-field correction, cosmic-ray removal, sky-subtraction, and wavelength calibration by using Th-Ar arc-lamps and flux calibration

Table 1. Some basic parameters of the stars observed in the field of UBC 103, UBC 114, and UBC 587, together with the exposure times used.

ID	Name	RA	Dec.	BP (mag)	RP (mag)	Exposure time	Spectral type
			UBC103				
1	TYC 5125-2315-1	18:42:29.5	−06:34:16	11.70	11.20	360	B4 III
2	–	18:42:33.8	−06:37:39	12.20	11.60	440	B5 IV
3	–	18:42:43.6	−06:39:56	12.50	11.79	400	B5 V
4	TYC 5125-2466-1	18:42:39.1	−06:35:20	12.42	11.92	440	B6 IV
5	–	18:42:31.7	−06:35:15	12.92	12.42	590	B5 V
			UBC114				
1	HD 177631	19:05:32.1	−00:16:15	10.59	10.26	200	B2 V
2	TYC 5128-116-1	19:05:37.8	−00:24:29	10.68	10.37	140	B2.5 V
3	TYC 5128-702-1	19:06:11.9	−00:30:05	11.29	10.90	300	B5 V
4	–	19:05:50.7	−00:22:35	12.28	11.83	450	B8.5V
5	–	19:05:46.5	−00:25:32	12.34	11.89	470	B8 V
6	TYC 5128-593-1	19:05:16.3	−00:35:22	12.04	11.55	440	B9 V
			UBC587				
1	BD +43°3710	20:45:34.7	+43:32:27	10.24	8.21	180	B1 Ia
2	BD +43°3714	20:46:46.3	+43:47:34	10.79	9.39	240	B0 II
3	ALS 11533	20:46:13.8	+43:44:59	11.53	10.22	270	O9.7 V
4	–	20:46:18.3	+43:35:04	12.45	11.29	440	B1 V
5	TYC 3178-867-1	20:46:18.6	+43:44:13	12.45	11.17	410	B0.5 IV
6	–	20:46:48.3	+43:47:30	12.56	11.22	410	B0.7 V

using spectra of standard stars observed with the same setup during the same nights. The spectral classification was carried out by following the traditional classification procedure through a comparison to a new, large grid of standard stars (Negueruela et al. 2018, and in prep.), degraded to the same resolution as the target spectra.

3.3. Astrometric data

As a first step, astrometric data from Castro-Ginard et al. (2020) were downloaded from the CDS/Vizier server. These include information on parallaxes (ϖ) and proper motions (μ_α, μ_δ) from *Gaia* DR2 (Gaia Collaboration 2018; Lindegren et al. 2018) for the individual members of each cluster as determined by their algorithm. We then used the Topcat[4] Virtual Observatory tool (Taylor 2005) to cross-match the coordinates of all these stars with the *Gaia* EDR3 database (Gaia Collaboration 2021). For the cross-matches, we used a radius threshold of 1 arcsecond. The EDR3 data for this set of likely members were used to recalculate average cluster values in EDR3.

After this, we used Topcat again to retrieve all stars in *Gaia* EDR3 matching two conditions: firstly, they must be within half a square degree of the average coordinates of the stars in each cluster; secondly, they must be within 2σ from the average value of ϖ, μ_α, and μ_δ in EDR3 for their corresponding cluster. The associated errors of each individual value were not considered. This procedure may leave out some stars that are in fact members, but have divergent astrometric values with large uncertainties[5]. However, this does not affect the average cluster values that we are deriving (see e.g., Maíz Apellániz et al. 2020a, for a procedure to detect all likely members and slow ejections from a given cluster).

---

[4] http://www.star.bris.ac.uk/~mbt/topcat/
[5] In fact, the uncertainties provided in the database are certainly underestimations for the parallax values. See Maíz Apellániz et al. (2021), for the best way to estimate realistic parallax uncertainties.

4. Results

In order to test the validity of the cluster parameters and cluster membership obtained with the (semi)automatic procedure of Castro-Ginard et al. (2020), we used both the spectroscopic and astrometric information available (see Sect. 3.2).

We found that some stars that were considered cluster members by Castro-Ginard et al. (2020), based on their *Gaia* DR2 parameters, lie outside 3σ of the new average *Gaia* EDR3 values in ϖ, μ_α, and μ_δ. We removed them and used the remaining stars to recalculate these averages. Once the new EDR3 cluster parameters were thus fixed, we selected all *Gaia* EDR3 sources within 2σ from these average values in all astrometric parameters. In the case of ϖ, we first corrected the values from zero-point offset by using the procedure described in Lindegren et al. (2021)[6].

From the sources selected, we took those brighter than $G = 17.4$ with an on-sky distance $<0°.25$ deg from the nominal cluster location in Castro-Ginard et al. (2020). This was felt to be a sensible compromise between the need to probe a larger spatial extent than determined by Castro-Ginard et al. (2020) and the danger of including other nearby stellar systems. Finally, we filtered out the sources with renormalised unit weight error (RUWE) >1.4 or with an error in parallax higher than 33%. This gave us a final list of stars whose parameters are compatible with cluster membership out to a large distance from their putative centres.

In a final step, we built the *Gaia* CMD, namely, G vs. ($BP - RP$), with all the stars selected through this procedure. We marked the members selected by Castro-Ginard et al. (2020), unless they had been removed and compared their position in the diagram with the rest of the stars that have compatible EDR3 astrometric values, but had not been selected as members by their algorithm. In each of the clusters, we estimated the age

---

[6] We used the algorithm provided by these authors at https://gitlab.com/icc-ub/public/gaiadr3_zeropoint

Table 2. Summary of parameters of UBC 103, UBC 114 and UBC 587 derived from the isochrone fitting compared with the results from Cantat-Gaudin et al. (2020).

	Age (Ma)	A_V (mag)	DM (mag)
	UBC103		
Cantat-Gaudin	22	1.4	12.37
This work	70 ± 10	1.6	12.41
	UBC114		
Cantat-Gaudin	50	1.0	10.9
This work	<30	1.3	10.84
	UBC587		
Cantat-Gaudin	14	3.1	12.2
This work	>3 & <10	>3.3 [(1)]	12.44

Notes. [(1)]Although many cluster members form a sequence that can be fitted by $A_V = 3.3$, many others have an obviously higher exinction.

of the population and the extinction (A_V) by visually fitting the isochrones, paying special attention to the position of the stars that we had observed spectroscopically and exploring the combination of both parameters. For the sake of completeness, in addition to the Padova (CMD 3.6) isochrones, we also downloaded MIST isochrones from their website[7] (version 1.2). In all cases, the distance modulus assumed for the fit corresponds to the distance obtained by simple inversion of the average *Gaia* EDR3 parallax, calculated as described at the start of this section. We then explored if the observed spectral types of the brightest members were compatible with their location in the CMD. A summary of the results derived from the isochrone fitting for each of the clusters compared to those from Cantat-Gaudin et al. (2020) is included in Table 2.

4.1. UBC 103

UBC 103 is a well populated cluster. It lies on the edge of the Scutum Star Cloud (at $l = 26.0$, $b = -1.0$) and is thus projected on an area of very high stellar density. Its CMD (see Fig. 5, top-left panel) is suggestive of a moderate degree of differential reddening. The mean parameters found by Castro-Ginard et al. (2020) are $\varpi = 0.28 \pm 0.03$ mas, pm = $(-0.40 \pm 0.09, -2.27 \pm 0.09)$ mas a^{-1}. The automated procedure of Cantat-Gaudin et al. (2020) finds a $DM = 12.37$, $A_V = 1.4$ and an age of 22 Ma. It is not possible, however, to obtain a good fit with these parameters, as the extinction needs to be higher for any sensible age. In *Gaia* EDR3, the average cluster values, after applying a zero-point offset correction to the parallax, are $\varpi = 0.33 \pm 0.06$ mas, pm = $(-0.38 \pm 0.08, -2.28 \pm 0.10)$ mas a^{-1}.

We observed the five brightest members according to Castro-Ginard et al. (2020). Their spectra are displayed in Fig. 6. The spectral types have been derived by assuming that they are all single stars. However, star #4 is almost certainly a binary, based on its broad asymmetric lines. Aside from that, all their spectral types (Table 1) are very similar, with an apparent turn-off at B5 V and some slightly evolved stars. These spectral types are inconsistent with the young age proposed by Cantat-Gaudin et al. (2020), as a turn-off at B5 V occurs at ages in the 60 to 100 Ma range at solar metallicity.

Amongst the EDR3 new candidate cluster members, we found three catalogued bright objects. One of them, TYC 5125-2068-1,

was photometrically classified as a B5 giant by Roslund (1963). It has very similar magnitude and colour to the brightest member for which we have a spectrum, TYC 5125-2315-1 (B4 III). Two other bright stars with redder colours occupy a position in the CMD compatible with cluster membership (top-left panel of Fig. 5). They are TYC 5125-368-1, whose location in the CMD suggests it is a yellow supergiant, and TYC 5125-2890-1, which occupies a position expected for a red supergiant. There are three strong reasons to believe that these two objects are cluster evolved members: Firstly, their position in the 2MASS CMD (($J - K_S$) vs K_S) is also compatible with the cluster isochrone (see Fig. 7). Secondly, their EDR3 radial velocities are 44.6 ± 0.2 and 45.6 ± 0.1 km s^{-1}, confirming that they belong to the same population. The only other red star with similar brightness has an incompatible value, and velocities around 45 km s^{-1} are not particularly common in this field[8]. Thirdly, stars of such brightness at this distance are in all likelihood supergiants, and such objects are much more likely to be associated to the young cluster.

These two candidate supergiants allow a firm determination of the cluster age. In Fig. 5, we show the fit of a 70 Ma isochrone reddened with $A_V = 1.6$, which gives a rather good fit for the distance modulus $DM = 12.4$ derived from the *Gaia* parallax[9]. Without the evolved stars, it is possible to obtain good fits for younger ages with slightly higher extinctions. The observed spectral types at the top of the main sequence, however, are incompatible with ages much younger than ~60 Ma. Ages up to 100 Ma would also be compatible with a turn-off at B5 V (cf. the case of IC 4725).

4.1.1. A population of interlopers

The search for new members among *Gaia* EDR3 sources out to 15 min from the centre adds a substantial population to the main sequence (top-left panel of Fig. 5), but it also includes a second sequence at much redder colours – about one magnitude in ($BP - RP$) – that was already hinted at in the sample of Castro-Ginard et al. (2020). The majority of these redder sources lie at large distances from the main cluster concentration (top-right panel of Fig. 5), suggesting that they represent a different population. To confirm this, we cross-matched the selection with the 2MASS database. The field is very crowded, and the possibility of spurious matches is high. We therefore kept only 2MASS sources with distances smaller than 0.′′2 from the corresponding DR3 sources. Also because of crowding, only a small fraction of sources have good flags and acceptable photometric errors. Although we were lax with the photometric precision required, allowing for errors up to 0.08 mag in each of the three (JHK_S) bands, we were left with only 75 sources, and many of them with moderately large errors.

The CMD for this sample of objects is shown in Fig. 7, where we have separated the stars that fall on the cluster sequence in the *Gaia* CMD from those in the second sequence at redder colours.

---

[7] http://waps.cfa.harvard.edu/MIST/interp_isos.html

[8] Castro-Ginard et al. (2020) quote a RV of -3.99 km s^{-1} for UBC 103. This is based on the RV of a single star that, given its position in the CMD, is not a cluster member. In fact, it is part of the redder population with similar proper motions discussed below.

[9] Nevertheless, we have to note that in the absence of a *Gaia* parallax, photometry alone would have favoured a distance modulus $DM = 12.0$, implying a slightly older age. It is difficult to estimate how significant this difference is, as, at this large distance, the small dispersion of the parallax values (0.33 ± 0.06) would translate into a broad range of distance modules (12.4 ± 0.4) and a huge range of distances ($3.0^{+0.7}_{-0.4}$ kpc), if directly propagated.
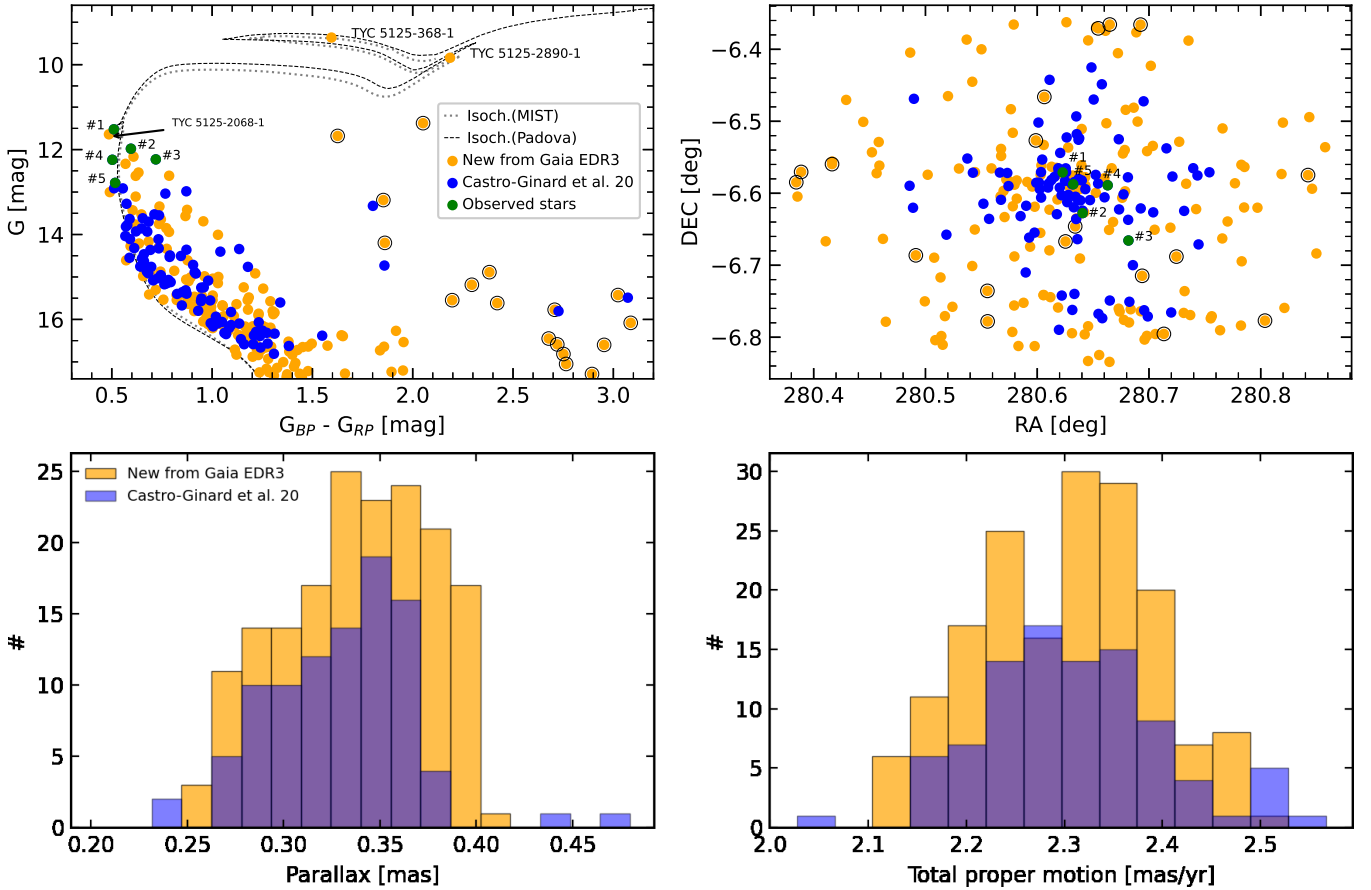
Fig. 5. Summary figure for UBC 103. *Top two figures:* colour-magnitude diagram (left) and sky coordinates of the sources (right). *Bottom two figures:* histograms of the parallax (left) and total proper motion (right). In *all figures*: blue circles or bins indicate cluster stars listed in Castro-Ginard et al. (2020), here represented by their *Gaia* EDR3 data; filled green circles are the observed stars; filled orange circles or bins indicate other stars found in *Gaia* EDR3 within 2σ from the average cluster values (see Sect. 4). Additionally, empty black circles indicate a potential separate population of stars (see Sect. 4.1); the gray dashed and dotted lines correspond to the best fitted isochrones from Padova and MIST databases, respectively, both for a value of 70 Ma and $A_V = 1.6$.
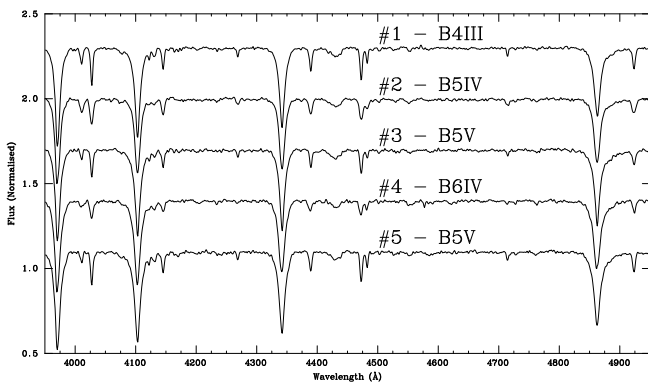
Fig. 6. Classification spectra of the brightest members of UBC 103 from Castro-Ginard et al. (2020).

We see that these objects also have much redder 2MASS colours. In addition, the objects in the red sequence are much brighter in K_S than the stars on the blue sequence of the same G, confirming that they are intrinsically red stars. For the brightest objects, which have accurate 2MASS photometry, we calculated their Q_{IR} reddening-free indices. While the brightest objects in the blue sequence have $Q_{IR} \lesssim 0.1$, typical of early-type stars, the objects in the red sequence have $Q_{IR} \gtrsim 0.2$, again indicating

that they are intrinsically redder objects (Negueruela & Schurch 2007). The only possible interpretation is the presence of a different population with the same astrometric parameters as cluster members. We may speculate on a population of red giants, which would be compatible with the red colours and bright infrared magnitudes.

4.1.2. Twin cluster and an association

Even if we remove this second population, there are many stars whose astrometric and photometric parameters are compatible with membership at high distances from the nominal cluster centre (top-right panel of Fig. 5). If we increase the radius of our search, rather than finding a decrease in the density of candidate members in all directions, we come across a very strong concentration located ~22′ to the north of UBC 103, coincident with the position of the known open cluster NGC 6883. Unfortunately, this cluster is not included in the catalogue of Cantat-Gaudin et al. (2020). In Appendix A, we calculate its astrometric parameters, which are fully compatible with those of UBC 103, making them a couple of twin clusters.

The distribution of astrometric members in the sky suggest that NGC 6883 and UBC 103 are the main concentrations of a rather large population spread over a large area, which extends well beyond the 15′ covered by our initial search. To estimate
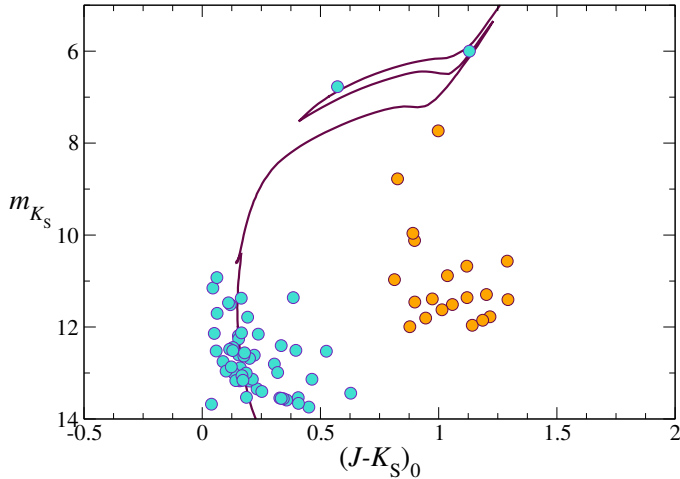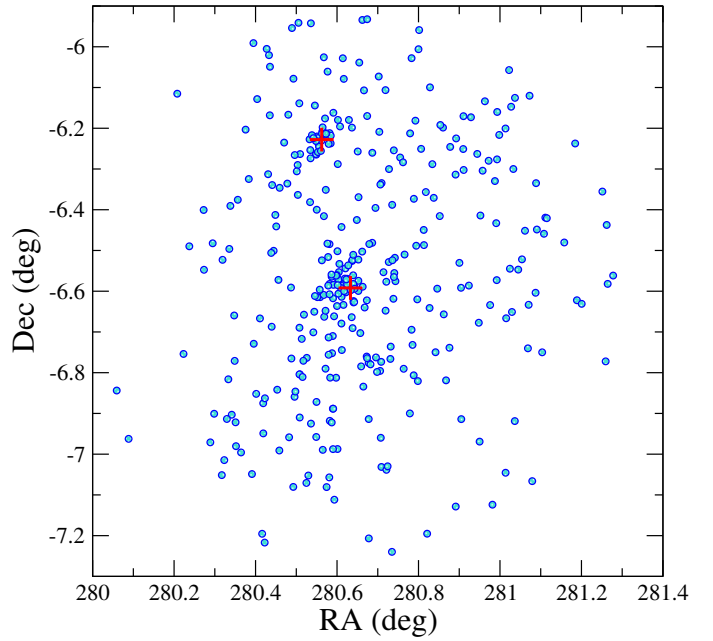
Fig. 7. Infrared CMD for the stars selected as astrometric members of UBC 103. Blue circles are the 2MASS counterparts of objects that fall in the cluster sequence and the two objects that we interpret as cluster supergiants. The orange circles are the counterparts to the objects that present redder colours in the *Gaia* CMD. They are also clearly displaced to redder colours, as well as to brighter magnitudes, in this plot, indicating a separate population.

Fig. 8. Spatial distribution of likely members of the stellar association surrounding UBC 103 out to 40′ from its nominal centre. High-probability members in the space of proper motions found by Clusterix 2.0 have been pruned by the parallax and position in the CMD (only stars with $10 \le G \le 16$ are included). The red crosses mark the nominal position of UBC 103 (in the centre of the search area) and the position of NGC 6883 found in Appendix A. We note the lack of members to the west, most likely due to intervening dark clouds causing heavy extinction.

its size, we used the Virtual Observatory tool Clusterix 2.0 (Balaguer-Núñez et al. 2020), an interactive web-based application that calculates the grouping probability of a list of objects by using proper motions and the non-parametric method proposed by Cabrera-Cano & Alfaro (1990) and described in Galadi-Enriquez et al. (1998). We run Clusterix on circles of increasing radius around UBC 103, restricting the search to stars brighter than $G = 16$ for ease of computation. We find significant numbers of stars with compatible proper motion and parallax out to 40′, especially towards the north and northeast of NGC 6883 and to the south of UBC 103. There are few stars belonging to this population to the west, but this is in all likelihood due to the presence of dark clouds, which render the density of bright stars much lower. Towards the east, we move deep into the Scutum Star Cloud, where crowding becomes an issue, even for *Gaia*. We cannot continue the search at larger radii, as we run into the very massive open cluster Valparaiso 1 (= UBC 106; $d \approx 2.3$ kpc, $\tau \approx 75$ Ma; Negueruela et al. 2021), whose centre is located slightly more than 1° due North from UBC 103, and its larger population starts to dominate the frequency distribution in the vector point diagram.

The results of this search can be seen in Fig. 8. We took EDR3 data for stars in the $10 \le G \le 16$ range within a circle of radius 40′ around the nominal position of UBC 103 and then we selected the objects identified by Clusterix 2.0 as high-probability members of a homogenous population in proper motion space. Afterwards, we pruned the sample by requiring a parallax consistent with that of UBC 103 and a position in the CMD compatible with the sequences of UBC 103 or NGC 6883. The result is a list of 358 stars, very strongly concentrated towards UBC 103, but including several other clumps, the most populous of which is NGC 6883. By comparison, there are 34 objects within 5′ of the centre of NGC 6883 and 58 objects within 5′ of the nominal position of UBC 103.

Extending (for completeness) the search to brighter stars, up to $G = 9$, recovers the two likely supergiant members of UBC 103 mentioned above and adds a single, very interesting source: the 12.9 d Cepheid Z Sct. This object, located close

to the northernmost edge of our search area, has astrometric parameters fully consistent with membership and average photometric colours almost identical to those of TYC 5125-368-1 (it is about 0.1 mag brighter in all bands). Its *Gaia* DR3 RV is 53.3 ± 3.6 km s^{-1}, where the large dispersion is due to intrinsic variability. Anderson et al. (2016), however, calculated a systemic velocity of 29.9 km s^{-1}, which is inconsistent with the RVs of the two members of UBC 103. The membership of this object is therefore uncertain.

At a distance of ∼3.0 kpc, a diameter of 80′ corresponds to a ∼70 pc span. The existence of such large stellar complexes in this direction is not unexpected. In fact, recently Kuhn et al. (2021) identified some larger complexes of star forming regions towards the inner Galaxy, including a huge high-pitch structure extending over almost 1 kpc between $l = 4°$ and 18°. In the direction to UBC 103 ($l \sim 26°$), the recently identified open cluster UBC 1035, located about 1° to the west, has the same proper motions, parallax, and estimated age as Valparaiso 1, which is ∼80′ away. The projection of several large clusters with ages between ∼50 and ∼100 Ma and slightly different distances could indicate that a structure similar to that found by Kuhn et al. (2021) has produced abundant star formation in the past.

4.2. UBC 114

UBC 114 is a poorly populated cluster at a relatively low distance. The grouping appears projected on top of a dark cloud, at $l = 32°6$, $b = −2°7$. The mean parameters found by Castro-Ginard et al. (2020) are $\varpi = 0.66 \pm 0.03$ mas, pm = $(−1.50 \pm 0.10, −1.88 \pm 0.09)$ mas a^{-1}. The automated procedure of Cantat-Gaudin et al. (2020) finds a $DM = 10.91$, $A_V = 1.0$
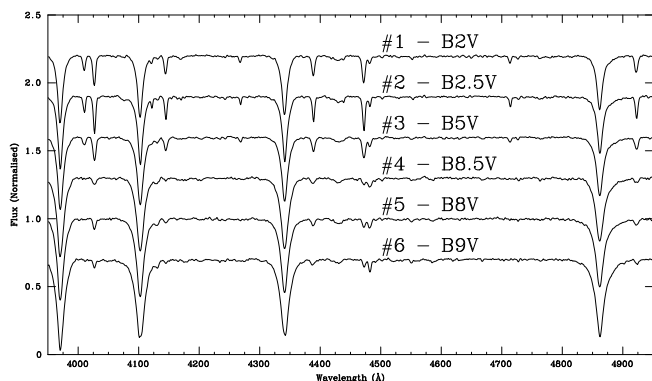
Fig. 9. Classification spectra of the brightest members of UBC 114 from Castro-Ginard et al. (2020).

and an age of 50 Ma. In *Gaia* EDR3, the average values for the same stars, after applying a zero-point offset correction for the parallax, are $\varpi = 0.68 \pm 0.03$ mas, pm = $(-1.48 \pm 0.06, -1.95 \pm 0.10)$ mas a^{-1}.

We observed the six brightest members from the list of Castro-Ginard et al. (2020), whose classification spectra are displayed in Fig. 9. Their spectral types (Table 1; mid panel) cover most of the range of B-type dwarfs. Again, all the spectral types assume that the stars are single. Star #5 is very likely a binary, given its broad, flat-bottomed He I lines. Star #6 is also possibly a binary, as it is brighter than stars of earlier spectral type.

The search in EDR3 only adds faint members to the cluster sequence, confirming that there is no obvious contaminating population. The upper main sequence is very poorly populated, and there are no evolved stars to fix the cluster age. The observed distribution of spectral types allows for the age cluster to be as young as wished, with an upper limit on the age given by the turn-off age of a B2 V star, around 25 Ma. As an example, in Fig. 10 we plot an isochrone of 26 Ma and $A_V = 1.3$, where the *Gaia* parallax provides $DM = 10.8$.

Alternatively, one could consider that the turn-off regions of the CMDs of young clusters are, in many cases, dominated by the effects of binary interaction and assume that the most luminous stars are blue stragglers of some sort. In support of this interpretation, the EDR3 astrometric values for star #3 are very different from those of DR2 and the star is no longer selected as a cluster member, given that its μ_δ (-2.34 ± 0.02 mas a^{-1}) is significantly different from the cluster average. If we assume that this object is not a member, then we find a huge gap of almost two magnitudes between the two stars at the top of the sequence and the faintest stars for which we have spectra, at types B8–9. The presence of such a gap may lead us to suspect that the actual age of the cluster is given by these later type stars, because the sequence is only continuous from that point, implying that the two early-type stars are some sort of blue stragglers.

This scenario, however, is unlikely, because a B2 V star has a mass of $\sim 8\,M_\odot$ and thus cannot be formed even by the direct merger of two B8 V stars ($\sim 3.3\,M_\odot$; Harmanec 1988). Therefore, we would have to assume that the blue stragglers have formed by the interaction of stars more massive than any that we see, which again requires a younger age. Moreover, turning back to star #3, if we consider that its parallax and μ_α are fully consistent with the cluster average, it is still possible to assume that the object is connected to the cluster. Given that it lies about 9 arcmin south of the nominal centre (see Fig. 10, right-top panel), this object may be a slow runaway in the process of being ejected. Therefore, we consider that this is indeed a very young cluster. Finally, we

note that even for older ages, the $A_V = 1.02$ value derived by Cantat-Gaudin et al. (2020) is – once again – too low to permit a good fit to the bulk of the stars. Our fits require $A_V = 1.3$ for ages younger than ~ 30 Ma and $A_V = 1.2$ for older ages, up to ~ 150 Ma.

4.3. UBC 587

UBC 587 is an extreme case, as its raw CMD does not show evidence of any distinctive shape. The distribution of stars in the CMD, however, suggests a vertical stripe affected by strong differential reddening. The mean parameters found by Castro-Ginard et al. (2020) are $\varpi = 0.28 \pm 0.02$ mas, pm = $(-2.87 \pm 0.08, -4.84 \pm 0.13)$ mas a^{-1}. The high distance implied further suggests a young open cluster. The automated procedure of Cantat-Gaudin et al. (2020) finds a $DM = 12.2$, $A_V = 3.1$ and an age of 14 Ma. In *Gaia* EDR3, the average values for the same stars, after applying a zero-point offset correction for the parallax, are $\varpi = 0.33 \pm 0.04$ mas, pm = $(-2.85 \pm 0.09, -4.86 \pm 0.13)$ mas a^{-1}.

The field lies in the area of the Cygnus X complex ($l = 83°6$, $b = +0°3$), less than 2° away from Deneb, in a gap between two dark clouds. We observed the three brightest members from the list of Castro-Ginard et al. (2020), two catalogued OB stars in their immediate vicinity, LS III +43°11 (= ALS 11533) and BD +43°3714, and a more distant catalogued OB star, BD +43°3710, whose proper motions are very similar to those of the cluster. Their spectra are displayed in Fig. 11. Their spectral types are presented in Table 1; bottom panel). The two BD stars are early-B supergiants, while the other four are dwarfs with spectral type close to B0.

Our search for new members in *Gaia* EDR3 adds a considerable number of faint members and a few objects with redder colours than the main stellar sequence (see top-left panel of Fig. 12). Most of these red objects lie to the West of the field (RA < 311°5 in the top-right panel of Fig. 12), where the dark cloud starts to be noticeable; thus, they are likely more reddened members. According to our criteria, BD +43°3710, which is located in the same area, is a cluster member. This object is much brighter than any other member, in good agreement with its spectral classification as a B1 Ia supergiant. Its red colour indicates that it is much more heavily reddened than the other members with spectra, although it is not redder than other EDR3 members in its vicinity.

BD +43°3710 is surrounded by a complex nebula seen in *Spitzer*/MIPS 24 μm images (Kraemer et al. 2010, their Fig. 1). The nebula, which is markedly bipolar, was analysed by Kraemer et al. (2010) under the erroneous assumption that BD +43°3710 was a nearby Carbon star. Flagey et al. (2014) reported the mistake and used low-resolution near-IR spectra to classify the star as a B-type supergiant. They estimated a spectral type B5 by comparison to a very limited number of standards. To check the possible effect of circumstellar material on its colours, we use its 2MASS photometry, $J = 6.60 \pm 0.02$, $H = 6.14 \pm 0.02$, and $K_S = 5.88 \pm 0.02$. Its $Q_{IR} = -0.01$ is typical of normal early-type stars, while emission-line stars and other objects with circumstellar discs usually have a more negative Q_{IR} (e.g., Negueruela & Schurch 2007; Maíz Apellániz et al. 2020b). In addition, the nebula is not detected at wavelengths shorter than 24 μm and the *Gaia* colours are not redder than those of other members in its vicinity. Therefore, there is no evidence for circumstellar emission in the near-IR. We can thus use the 2MASS photometry to estimate its intrinsic brightness. Assuming a standard (Rieke & Lebofsky 1985) extinction law, the intrinsic
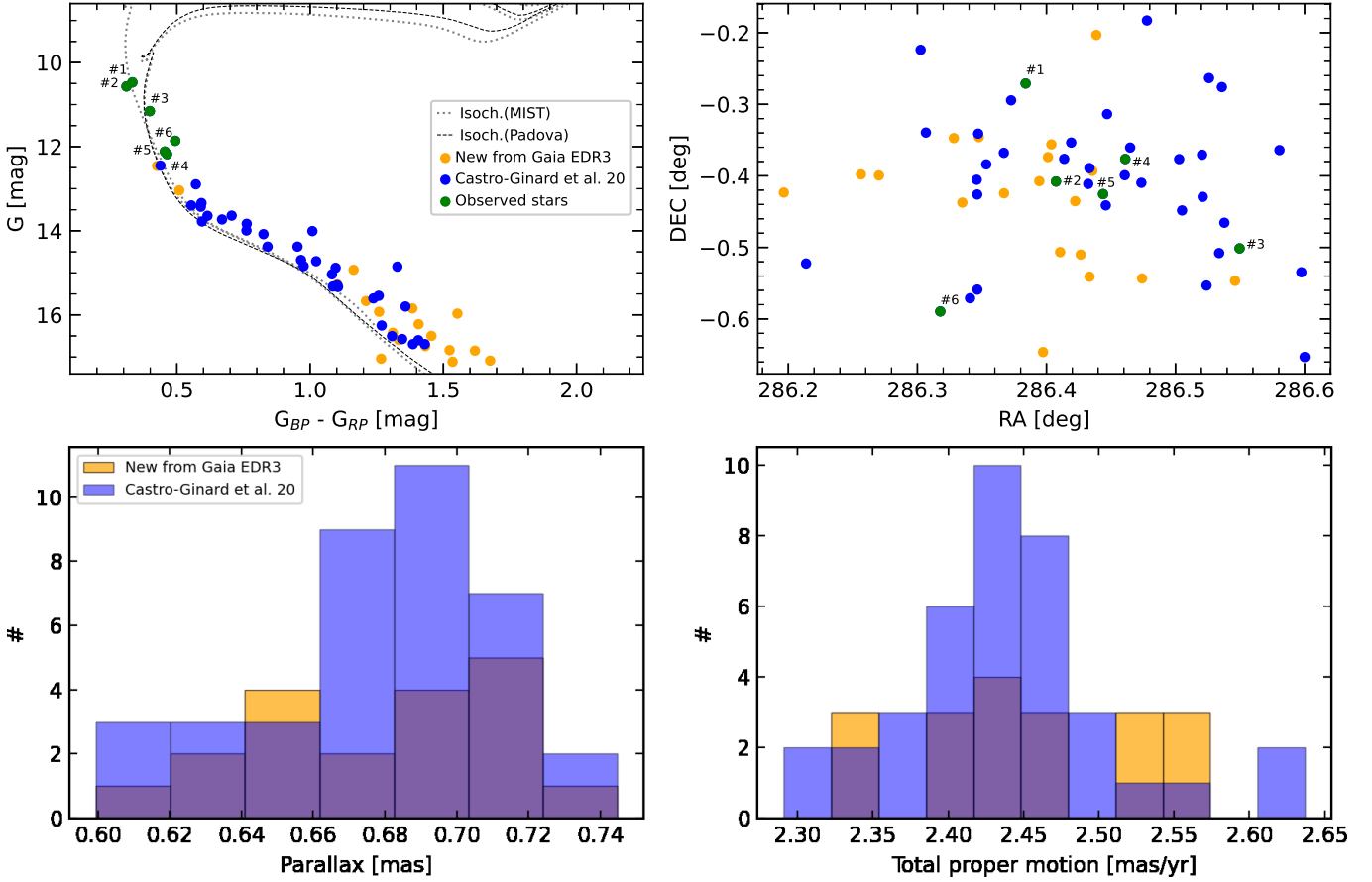
Fig. 10. Summary figure for UBC 114. *Top two figures:* colour-magnitude diagram (left) and sky coordinates of the sources (right). *Bottom two figures:* histograms of the parallax (left) and total proper motion (right). See Fig. 5 for a description of the symbols and colours. The two grey lines next to each other are isochrones of 100 Ma with $A_V = 1.2$ (the dashed line is a CMD 3.6 isochrone, while the dotted line is from the MIST database). Dotted line to the left of the others corresponds to a 26 Ma isochrone with $A_V = 1.3$, taken from MIST. The data seem to favour the younger age.
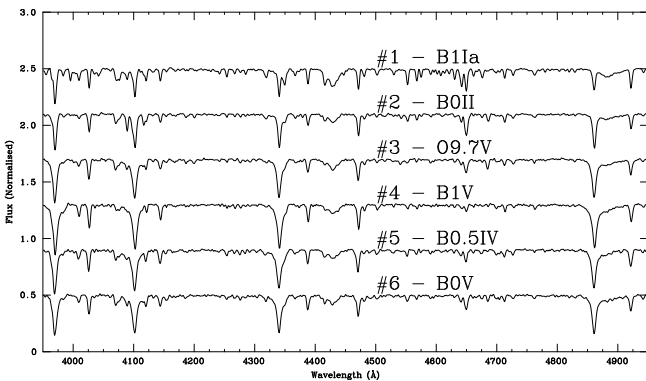
Fig. 11. Classification spectra of the three brightest members of UBC 587 according to Castro-Ginard et al. (2020) and three brighter catalogued OB stars in the area.

colours of Winkler (1997) and the $DM = 12.4$ found for the cluster, the intrinsic M_K is a phenomenal -7.1, fully consistent with a Ia luminosity class.

The second supergiant in the field, BD +43°3714, was not selected as a cluster member. Although its parallax, $\varpi = 0.28 \pm 0.02$ mas (uncorrected) is fully compatible with the cluster average, its proper motions deviate significantly. For this reason, it is not included in Fig. 12. Nevertheless, its observed colour

($BP - RP = 1.39$) and G magnitude (1.8 mag brighter than the main sequence stars observed) are fully consistent with a star of its spectral type at the cluster distance. We can thus speculate with a cluster member whose kinematic properties are due to a past binary interaction. In the case of the other catalogued OB star, LS III +43°11, its ϖ (0.22 ± 0.04 mas) and μ_δ (−5.18 ± 0.04 mas a^{-1}) are just outside the 2-σ threshold used for member selection (but its high RUWE = 2.2 suggests that the errors may be underestimated, even without applying the correction of Maíz Apellániz et al. 2021, and so it is likely a cluster member). We might speculate that this is a background star, but then it is far too bright for its spectral type. Even at the cluster distance, its $M_K = −4.0$ (derived as above) is too bright for the spectral type, although it is, in fact, compatible with a binary, for example.

The age of UBC 587 cannot be strongly constrained beyond stating that it is a very young open cluster. The spectral types of the main sequence stars present force it to be younger than $\sim$10 Ma, while the presence of a B supergiant, namely, a moderately evolved massive star, makes it older than $\sim$3 Ma. The extinction $A_V = 3.3$ used in the fit (Fig. 12) is adequate for the less reddened members, but much higher values should be used for the more obscured objects.

With a *Gaia* distance of 3.0 kpc, UBC 587 is decidedly behind the Cygnus complex. Parallax distances to different star-forming regions in the Cygnus X complex are compatible with a
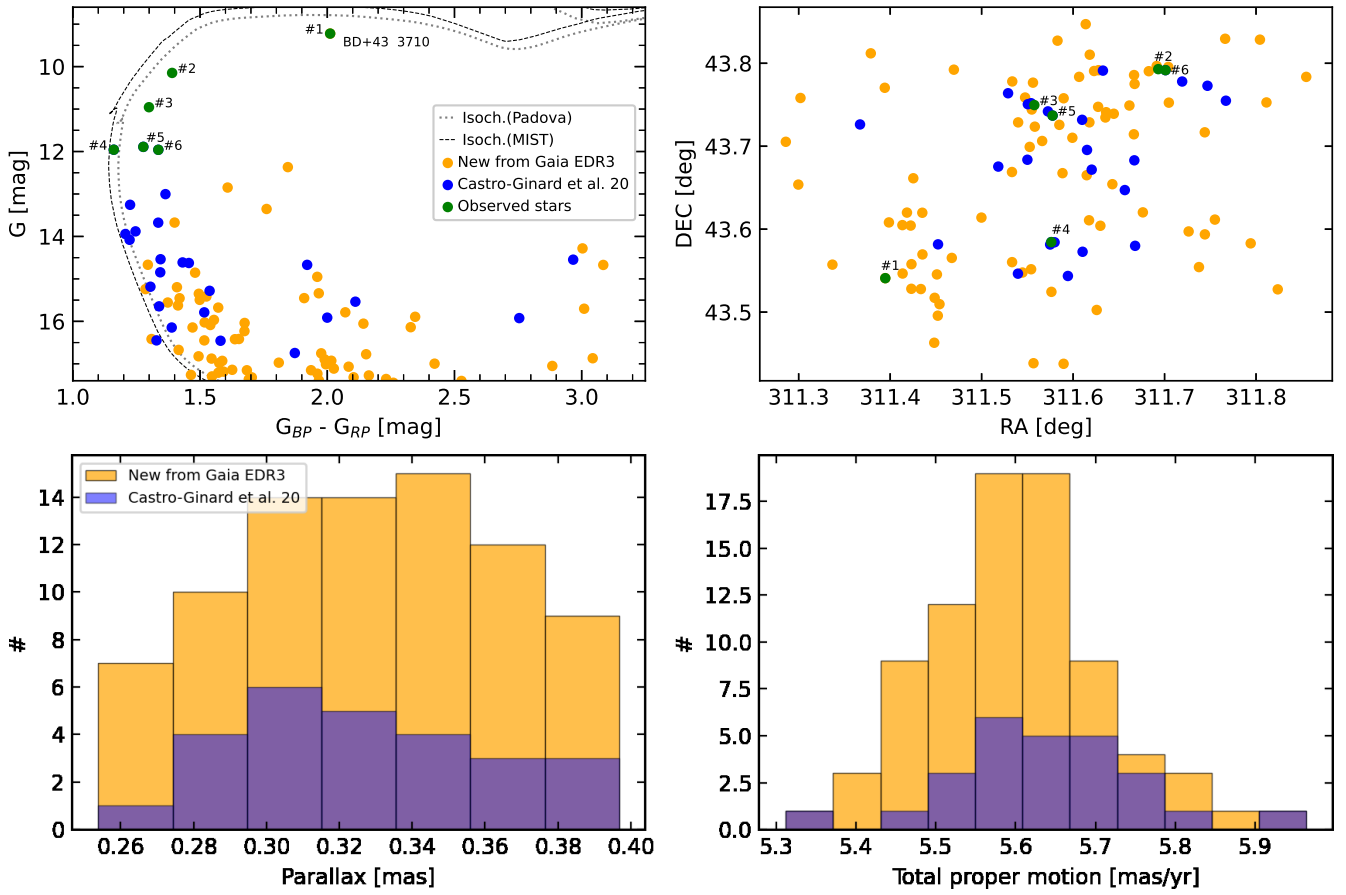
Fig. 12. Summary figure for UBC 587. *Top two figures:* colour-magnitude diagram (left) and sky coordinates of the sources (right). *Bottom two figures:* histograms of the parallax (left) and total proper motion (right). See Fig. 5 for a description of the symbols and colours. The grey dashed and dotted lines correspond to isochrones from Padova and MIST databases, respectively, both for a value of 16 Ma and $A_V = 3.3$.

distance around 1.4 kpc (Rygl et al. 2012). An analysis of *Gaia* DR2 data suggests that the bulk of Cyg OB2 is considerably farther away, at around 1.8 kpc (Berlanas et al. 2019), but this is still much closer than UBC 587. Our line of sight in this direction ($l \approx 83°$) is believed to run along the extent of the Local Arm for several kpc (Xu et al. 2013). As a very young open cluster, UBC 587 must be a tracer of more distant stretches of this arm than are usually considered (of all the tracers used by Xu et al. 2013, only two have comparable distances).

Interestingly, Castro-Ginard et al. (2022) recently reported a second young open cluster that overlaps on the sky with UBC 587. UBC 1113 has different proper motions, but a very similar parallax, and is affected by about the same amount of extinction. In particular, average values of $(BP - RP)$ for both clusters are very similar, with $(BP - RP)_{\text{UBC 587}} = 1.73 \pm 0.45$ and $(BP - RP)_{\text{UBC 1113}} = 1.68 \pm 0.27$. In Fig. 13, we show the spatial distribution of members of the two clusters. Although they clearly overlap in space and reddening, UBC 1113 is clearly more spatially concentrated. Inspection of sky images suggests that both clusters are seen through a gap in the dense clouds of the Cygnus X complex due to nearby star formation.

5. Conclusions

We have explored the use of *Gaia* photometry to identify young open clusters among the many new catalogues of candidates generated by means of artificial intelligence algorithms. We find that the presence of an almost vertical sequence in the $(BP - RP)$ ver-
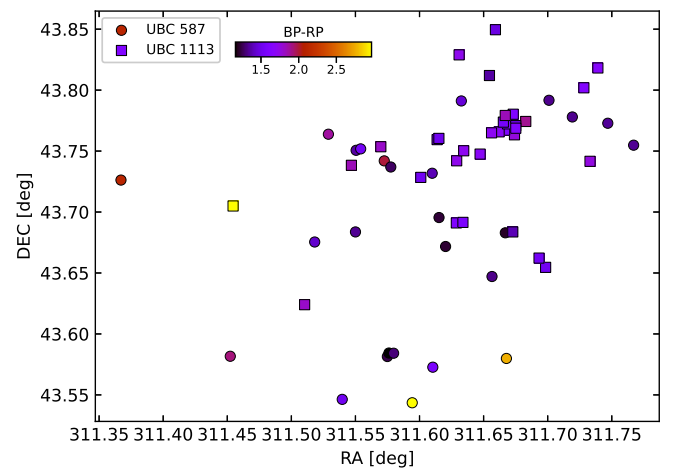
Fig. 13. Spatial distribution in sky coordinates of members in UBC 587 and UBC 1113 from Castro-Ginard et al. (2020; circles) and Castro-Ginard et al. (2022; with squares), respectively. The colour code corresponds to the $(BP - RP)$ colour. Compare to Fig. 12, top-right panel, where a group of EDR3 members with much higher reddening can be seen around $l = 311°4$ (close to star #1).

sus G CMD is a clear-cut signature of a cluster containing B-type stars. The vertical sequence is more readily identifiable when the cluster sequence extends to fainter magnitudes, as in the examples displayed in Fig. 3, but it is a reliable signal even when

dealing with poor sequences affected by differential reddening, as is the case of UBC 587. Automated algorithms to derive parameters will always give back a cluster age, even in those cases where the data do not provide a definitive answer, as in the example of UBC 94 (Fig. 1), and therefore the possibility of confirming the young nature of an open cluster from direct inspection of the CMD is reassuring.

The length of the vertical stripe is directly related to the age of the cluster. In very young open clusters, containing O-type stars and/or blue supergiants, it can extend over more than six magnitudes, while Fig. 3 illustrates the case of older clusters. The vertical sequence can even be detected in well-populated clusters of 200–300 Ma. At this age, the main-sequence turn-off should be around B8 at solar metallicity, and the red clump should already be significant, allowing an independent determination of the age. A good example of such a cluster (if somewhat extreme in terms of its mass) is M 11, which is around 220 Ma old (Bavarsad et al. 2016) and presents a vertical sequence extending for about two magnitudes.

Beyond confirming the validity of this approach, two of our test targets have turned out to possess considerable astrophysical interest of their own. UBC 103 is a moderately massive cluster of around 70 Ma, forming the core of an extended association together with its twin NGC 6683. At a distance around 3 kpc, this region may be associated to the Scutum Arm, although perhaps it represents the result of a high-pitch angle structure, like the one found by Kuhn et al. (2021), as it is aligned with the similarly-aged massive cluster Valparaiso 1, which (at 2.3 kpc) seems to lie in between the arms. UBC 587 is also a valuable spiral tracer, lying well behind the known structures in the Cygnus complex and perhaps marking the outer edge of the Local Arm. It contains a very bright B1 Ia supergiant, BD +43°3710, which is surrounded by a complex mid-IR nebula. UBC 184, on the other hand, represents an example of the difficulty in assigning ages to clusters if their stellar sequences are not well populated.

Automated determinations of cluster parameters with artificial intelligence algorithms, as used by Cantat-Gaudin et al. (2020) and Castro-Ginard et al. (2022), give results that do not differ very strongly from those obtained by means of visual fits. However, in most cases where a comparison has been carried out, the value of extinction found by the algorithms has been lower than that found by humans, with no example in our sample of the opposite case. As shown by the case of IC 4725 in Fig. 3, the algorithms may be misinterpreting the spread caused by differential reddening. If this is the general case, then extinctions obtained from these artificial intelligence methods are usually underestimated, likely resulting in an overestimation of the cluster age. The sample studied here is too small to claim a systematic effect, but a trend in exactly the same sense can be seen in other clusters that we have recently studied by using Gaia photometry, such as NGC 6649 and NGC 6664 (Alonso-Santiago et al. 2020) or Valparaiso 1 (Negueruela et al. 2021). This is particularly relevant if we consider the study of Dias et al. (2021), who, by using a different isochrone fitting algorithm, find values of extinction that are systematically smaller than those of Cantat-Gaudin et al. (2020), except for very high values ($A_V \gtrsim 3$), when they become systematically larger.

Intriguingly, all the clusters studied in the direction of the Scutum Star Cloud (UBC 103 and NGC 6683 in this work, as well as Valparaiso 1 in Negueruela et al. 2021) would give a better isochrone fit for a shorter distance modulus than derived from Gaia. We have not seen a similar effect for any other young cluster, with Gaia distances generally providing

the best isochrone fits. This is an issue that deserves further investigation.

References

Alonso-Santiago, J., Negueruela, I., Marco, A., Tabernero, H. M., & Castro, N. 2020, A&A, 644, A136
Anderson, R. I., Casertano, S., Riess, A. G., et al. 2016, ApJS, 226, 18
Balaguer-Núñez, L., López del Fresno, M., Solano, E., et al. 2020, MNRAS, 492, 5811
Bavarsad, E. A., Sandquist, E. L., Shetrone, M. D., & Orosz, J. A. 2016, ApJ, 831, 48
Berlanas, S. R., Wright, N. J., Herrero, A., Drew, J. E., & Lennon, D. J. 2019, MNRAS, 484, 1838
Cabrera-Cano, J., & Alfaro, E. J. 1990, A&A, 235, 94
Cantat-Gaudin, T., Jordi, C., Vallenari, A., et al. 2018, A&A, 618, A93
Cantat-Gaudin, T., Anders, F., Castro-Ginard, A., et al. 2020, A&A, 640, A1
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Castro-Ginard, A., Jordi, C., Luri, X., Cantat-Gaudin, T., & Balaguer-Núñez, L. 2019, A&A, 627, A35
Castro-Ginard, A., Jordi, C., Luri, X., et al. 2020, A&A, 635, A45
Castro-Ginard, A., McMillan, P. J., Luri, X., et al. 2021, A&A, 652, A162
Castro-Ginard, A., Jordi, C., Luri, X., et al. 2022, A&A, 661, A118
Chen, Y., Bressan, A., Girardi, L., et al. 2015, MNRAS, 452, 1068
Dias, W. S., Monteiro, H., Moitinho, A., et al. 2021, MNRAS, 504, 356
Ekström, S., Georgy, C., Eggenberger, P., et al. 2012, A&A, 537, A146
Ferreira, F. A., Corradi, W. J. B., Maia, F. F. S., Angelo, M. S., & Santos, J. F. C. Jr. 2020, MNRAS, 496, 2021
Flagey, N., Noriega-Crespo, A., Petric, A., & Geballe, T. R. 2014, AJ, 148, 34
Gaia Collaboration (Brown, A. G. A., et al.) 2018, A&A, 616, A1
Gaia Collaboration (Brown, A. G. A., et al.) 2021, A&A, 649, A1
Galadi-Enriquez, D., Jordi, C., & Trullols, E. 1998, A&A, 337, 125
Hao, C., Xu, Y., Wu, Z., He, Z., & Bian, S. 2020, PASP, 132
Hao, C. J., Xu, Y., Wu, Z. Y., et al. 2022, A&A, 660, A4
Harmanec, P. 1988, Bull. astr. Inst. Czechosl., 39, 329
He, Z.-H., Xu, Y., Hao, C.-J., Wu, Z.-Y., & Li, J.-J. 2021, Res. Astron. Astrophys., 21, 093
He, Z., Liu, X., Luo, Y., Wang, K., & Jiang, Q. 2023, ApJS, 264, 8
Kharchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., & Scholz, R.-D. 2013, A&A, 558, A53
Kraemer, K. E., Hora, J. L., Egan, M. P., et al. 2010, AJ, 139, 2319
Kuhn, M. A., Benjamin, R. A., Zucker, C., et al. 2021, A&A, 651, L10
Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, A&A, 616, A2

Lindegren, L., Bastian, U., Biermann, M., et al. 2021, A&A, 649, A4
Liu, L., & Pang, X. 2019, ApJS, 245, 32
Maíz Apellániz, J., & Weiler, M. 2018, A&A, 619, A180
Maíz Apellániz, J., Crespo Bellido, P., Barbá, R. H., Fernández Aranda, R., & Sota, A. 2020a, A&A, 643, A138
Maíz Apellániz, J., Pantaleoni González, M., Barbá, R. H., García-Lario, P., & Nogueras-Lara, F. 2020b, MNRAS, 496, 4951
Maíz Apellániz, J., Pantaleoni González, M., & Barbá, R. H. 2021, A&A, 649, A13
Negueruela, I. 2017, in The Lives and Death-Throes of Massive Stars, eds. J. J. Eldridge, J. C. Bray, L. A. S. McClelland, & L. Xiao, 329, 271

Negueruela, I., & Schurch, M. P. E. 2007, A&A, 461, 631
Negueruela, I., Monguió, M., Marco, A., et al. 2018, MNRAS, 477, 2976
Negueruela, I., Chené, A. N., Tabernero, H. M., et al. 2021, MNRAS, 505, 1618
Rieke, G. H., & Lebofsky, M. J. 1985, ApJ, 288, 618
Roslund, C. 1963, Ark. Astron., 3, 97
Rygl, K. L. J., Brunthaler, A., Sanna, A., et al. 2012, A&A, 539, A79
Sanders, W. L. 1971, A&A, 14, 226
Taylor, M. B. 2005, ASP Conf. Ser., 347, 29
Winkler, H. 1997, MNRAS, 287, 481
Xu, Y., Li, J. J., Reid, M. J., et al. 2013, ApJ, 769, 15
Yilmaz, F. 1966, ZAp, 64, 61

Appendix A: Parameters of NGC 6683

As discussed in Sect. 4.1, a search for outlying members of UBC 103 finds a strong concentration of objects with compatible astrometric parameters in the vicinity of the nominal position for NGC 6683. This is a poorly studied cluster that was first characterised by Yilmaz (1966), who obtained *RGU* photometry from photographic plates, finding a distance modulus *DM* = 10.5 and an estimated earlier spectral type around b4. This is the only dedicated study that we find in the literature. Kharchenko et al. (2013) estimated a distance *d* = 1.4 kpc and an age of only 5.6 Ma. By fitting isochrones to 2MASS photometry, Chen et al. (2015) derived a much longer *DM* ≈ 11.6 (*d* ≈ 2.1 kpc) and an age of ∼ 160 Ma.

To characterise the cluster, we performed a Bayesian analysis of the field by using the Virtual Observatory tool Clusterix 2.0 (Balaguer-Núñez et al. 2020), an interactive web-based application that calculates the grouping probability of a list of objects by using proper motions and the non-parametric method proposed by Cabrera-Cano & Alfaro (1990) and described in Galadi-Enriquez et al. (1998). We took EDR3 objects with *RUWE* < 1.4 and errors in proper motion below 0.5 mas a^{-1}, with magnitudes in the $10 \leq G \leq 18$ range. The cluster central concentration seems to lie about 1 arcmin south of the nominal position in SIMBAD, close to the location considered by Yilmaz (1966), whose stars 1 and 3-6 all seem to be members. We therefore centred the circles used by Clusterix on RA = 18:42:15, Dec:−06:13:40, taking a radius of 10 arcmin for the whole sample and an initial cluster radius of 4′.8.

Based on an empirical determination of the frequency functions in the vector point diagram (Sanders 1971), Clusterix assigns to each object a probability of belonging to a distinct population. The cluster is clearly noticeable in the vector point diagram and stands out very strongly in the probability distribution. We made a cut in probability that leaves out most objects not related to the cluster and then calculated weighted averages for parallax, μ_α and μ_δ. Finally, we performed a sigma-clipping iterative procedure to clean the sample of outliers. For this, we calculated the weighted average (weighted by their errors) of all astrometric parameters and iteratively removed objects more than 2 σ away until convergence was reached.

For all three astrometric parameters, median values are indistinguishable from weighted averages, confirming that this is a very well defined population. The average values are μ_α = −0.34 ± 0.07 mas a^{-1}, μ_δ = −2.33 ± 0.08 mas a^{-1} and ϖ = −0.32 ± 0.05 mas a^{-1} (after zero-point correction). These values are all less than one σ away from the average values for
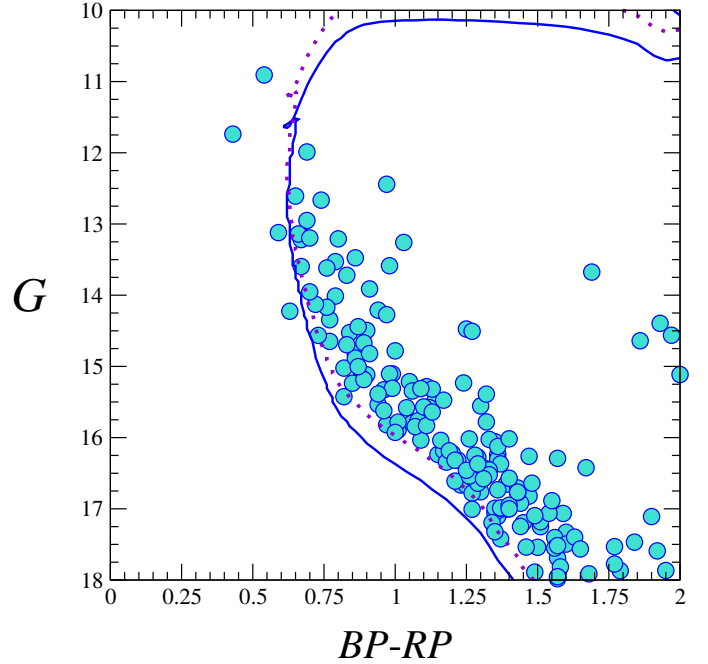
Fig. A.1. *Gaia* EDR3 CMD for stars selected as members of NGC 6883. The solid (blue) line is the same Padova isochrone that gives the best fit to UBC 103, i.e. 70 Ma, displaced to *DM* = 12.4, but affected by A_V = 1.8. Dotted (purple) line is the same isochrone displaced to *DM* = 12.0.

UBC 103, with the parallax being indistinguishable. UBC 103 and NGC 6683 are therefore twin clusters. Lacking any evolved star, it is difficult to fix the age of NGC 6683, but it is certainly not very different from that of UBC 103. Figure A.1 shows the *Gaia* photometry together with the same isochrone that fits UBC 103, namely, a solar metallicity 70 Ma isochrone. The extinction required is slightly higher, A_V = 1.8. As in the case of UBC 103, the fit is not bad, but a shorter distance modulus *DM* = 12.0 would be favoured from the photometry, if we lacked the *Gaia* parallax.

In any event, based on *Gaia* data, NGC 6683 seems significantly more distant and younger than assumed by previous works. The membership of CK Sct in NGC 6683, proposed by Chen et al. (2015), is not confirmed by *Gaia* astrometry, as this object has a parallax ϖ = 0.55±0.08 in DR2 and ϖ = 0.45±0.02 in EDR3, indicating a shorter distance, and very different proper motions.