

# A massive open cluster hiding in full sight

I. Negueruela,<sup>1,2\*</sup> A.-N. Chené,<sup>3,4</sup> H. M. Taberero,<sup>2,5</sup> R. Dorda,<sup>2,6,7</sup>

J. Borissova,<sup>3,8</sup> A. Marco,<sup>2</sup> and R. Kurtev<sup>3,8</sup>

<sup>1</sup>*Departamento de Física Aplicada, Facultad de Ciencias, Universidad de Alicante, Carretera de San Vicente s/n, E03690, San Vicente del Raspeig, Spain*

<sup>2</sup>*Departamento de Física, Ingeniería de Sistemas y Teoría de la Señal, Universidad de Alicante, Carretera de San Vicente s/n, E03690, San Vicente del Raspeig, Spain*

<sup>3</sup>*Instituto de Física y Astronomía, Universidad de Valparaíso, Av. Gran Bretaña 1111, Playa Ancha, Casilla 5030, Chile.*

<sup>4</sup>*Gemini Observatory, Northern Operations Center, 670 A'ohoku Place, Hilo, HI 96720, USA*

<sup>5</sup>*Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, 4150-762 Porto, Portugal*

<sup>6</sup>*Instituto de Astrofísica de Canarias, Vía Láctea s/n, E38200, La Laguna, Tenerife, Spain*

<sup>7</sup>*Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain*

<sup>8</sup>*Instituto Milenio de Astrofísica, Nuncio Monseñor Sotero Sanz 100, Of. 104, Providencia, Santiago, Chile*

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

Obscuration and confusion conspire to limit our knowledge of the inner Milky Way. Even at moderate distances, the identification of stellar systems becomes compounded by the extremely high density of background sources. Here we provide a very revealing example of these complications by unveiling a large, massive, young cluster in the Sagittarius arm that has escaped detection until now despite containing more than 30 stars brighter than  $G = 13$ . By combining *Gaia* DR2 astrometry, *Gaia* and 2MASS photometry and optical spectroscopy, we find that the new cluster, which we name Valparaiso 1, located at  $\sim 2.3$  kpc, is about 75 Ma old and includes a large complement of evolved stars, among which we highlight the 4 d classical Cepheid CM Sct and an M-type giant that probably represents the first detection of an AGB star in a Galactic young open cluster. Although strong differential reddening renders accurate parameter determination unfeasible with the current dataset, direct comparison to clusters of similar age suggests that Valparaiso 1 was born as one of the most massive clusters in the Solar Neighbourhood, with an initial mass close to  $10^4 M_{\odot}$ .

**Key words:** stars: evolution – supergiants – Hertzsprung-Russell and colour-magnitude diagrams – stars: variables: Cepheids – open clusters and associations: individual: Valparaiso 1

## 1 INTRODUCTION

Star formation is a complex hierarchical process. Its outcome depends on many factors that we do not know yet. Giant molecular clouds can give rise to dispersed OB associations that contain no populous clusters, as in Cyg OB2 (Wright et al. 2016), but also produce massive, compact clusters, with masses  $M_{\text{cl}} \gtrsim 10^4 M_{\odot}$ , such as NGC 3603 (e.g. Rochau et al. 2010). Despite this diversity, the mass distribution of open clusters in non-interacting galaxies seems to be quite similar and independent of their star formation rates (Fall & Chandar 2012), especially among spirals (e.g. Larsen 2009). Based on the expectations of such a standard distribution, the Milky way should produce one bound cluster more massive than  $10^4 M_{\odot}$  about every  $\sim 500\,000$  years (Larsen 2009). Indeed, about a dozen star clusters with ages  $\lesssim 20$  Ma and masses above this limit have now been found in

our Galaxy (Portegies Zwart et al. 2010; Negueruela 2014). There is, however, a general lack of clusters with comparable masses at slightly older ages. In fact, only M11 (Santos et al. 2005; Cantat-Gaudin et al. 2014) has been confirmed to be above the  $10^4 M_{\odot}$  landmark among moderately young clusters. Bavarsad et al. (2016), based on the properties of a eclipsing binary in the cluster, give a turn-off mass of  $3.6 M_{\odot}$  and thus an age around 250 Ma, which implies an initial mass well above this value (cf. Cantat-Gaudin et al. 2014). The heavily obscured cluster Mercer 13 (Messineo et al. 2009) may have a similar age and an equivalent population, but its parameters are still uncertain.

Similar clusters must have been forming throughout the history of the Milky Way. In fact, any old open cluster that has managed to survive to this day in the Galactic gravitational potential must have started as a rather massive object. For instance, Hurley et al. (2005) estimate that M 67, which is  $\sim 4$  Ga old, started its life with almost  $20\,000 M_{\odot}$  in order to retain  $\sim 1\,500 M_{\odot}$  at present. Several authors have proposed

\* E-mail: ignacio.negueruela@ua.es

the heavily reddened compact cluster GLIMPSE-C01 to be the intermediate-age descendent of a very massive young open cluster. Data are compatible with a 2 Ga cluster still having  $\sim 10^5 M_{\odot}$ , but also with an old globular cluster that happens to be crossing the Galactic Plane (Hare et al. 2018). At intermediate ages, the outer Galaxy contains several clusters that seem to have had very high initial masses. Among them, we can count NGC 2477, for which Eigenbrod et al. (2004) estimate a current mass of  $\sim 5400 M_{\odot}$  in stars more massive than the Sun at an age of 1.0 Ga, which seems confirmed by more recent works. NGC 1817 is similar (Mermilliod et al. 2003), while NGC 2099 is about half this age and at least as massive (Kalirai et al. 2001).

Towards the inner Galaxy, there are very few equivalent clusters. NGC 4337, which is some distance above the Plane ( $b = +4^{\text{d}}6$ ), has an estimated mass  $\gtrsim 2 \times 10^3 M_{\odot}$  at 1.6 Ga and may be the remnant of a rather massive cluster (Seleznov et al. 2017), but very few similar objects have been found and studied. Considering that most of the Galactic mass lies towards the inside and that most young massive clusters known are within the Solar circle, there are two immediate explanations<sup>1</sup> for this apparent difference between the inner and outer regions of the Milky Way:

- The stronger gravitational potential in the inner Galaxy must be much more efficient at dissolving initially massive clusters over the course of 1 Ga.

- High obscuration and crowding, prevalent towards the inner Galaxy, render many other existing massive cluster remnants essentially invisible to observers situated in our location. In support of this interpretation, the few populous moderately young open clusters known, such as M11 or NGC 6067 (Alonso-Santiago et al. 2017) are relatively nearby ( $d \lesssim 2$  kpc) and in the middle of low-extinction windows. Despite this, they do not stand out strongly from their background. While younger populous clusters, such as Stephenson 2 (Negueruela et al. 2012) or vdBH 222 (Marco et al. 2014), are easily identified as a collection of extremely bright infrared stars (their red supergiants), older clusters contain giants whose intrinsic brightness is comparable to the field red giant branch (RGB) or asymptotic giant branch (AGB) stars.

In this paper, we present evidence in the sense that many moderately young and intermediate-age clusters must remain unidentified because of background confusion by presenting a new, large cluster containing bright stars (about 30 members are brighter than  $G = 13$ ) that appears to be rather massive and has escaped detection until now. The new cluster, which we call Valparaiso 1, lies on the Galactic plane at  $l = 27^{\text{d}}0$  and is thus projected on top of a line of sight that goes across the Sagittarius and Scutum-Crux arms and reaches the base of Scutum-Crux arm, close to the near end of the Bar (Negueruela et al. 2012). During a search for background red luminous stars (reported in Negueruela et al. 2012), we found a large number of candidate early type stars in the vicinity of the catalogued emission-line star THA 14-29 (= TYC

5121-662-1) and the catalogued luminous star LSIV  $-05^{\circ}12$ . Some of these stars were observed spectroscopically, revealing a population of mid-B giants, highly indicative of a single-age population. *Gaia* DR2 astrometric data confirm the existence of this population, while revealing a much larger extent and a high number of associated red giants. The cluster that we study is included in the list of cluster candidates presented by Castro-Ginard et al. (2020), under the name UBC 106. In this work, we use intermediate-resolution spectra of these red giants to determine their stellar parameters and *Gaia* and 2MASS photometry to estimate cluster parameters.

## 2 OBSERVATIONS AND DATA

Classification spectroscopy of blue stars in the area was obtained with the Boller & Chivens Spectrograph on the 2.5-meter Irénée du Pont telescope at Las Campanas Observatory (Chile) during a run on 2012 June 25–27 June. The instrument was equipped with a Marconi CCD mated to a Bowen Schmidt camera as detector. We used the blue 1200 l/mm grating, which gives a nominal dispersion of  $0.8 \text{ \AA}/\text{pixel}$  over a wavelength range of 163 nm. Together with a 2 arcsec slit, this grating results in a resolving power of  $R = 2450$ .

Intermediate-resolution spectra of red giant candidates was taken on 2018 July 5–7 with the Intermediate Dispersion Spectrograph (IDS) mounted on the 2.5-m Isaac Newton Telescope (INT) at El Roque de los Muchachos Observatory, in La Palma (Spain). We used the 235-mm camera with the *Red+2* CCD and grating H1800V. This grating provides intermediate-high resolution over the whole optical range. In combination with a 1 arcsec slit, it gives an oversampled 2-pixel resolution element of  $0.66 \text{ \AA}$  over an unvignetted range of  $\sim 640 \text{ \AA}$ . We used it on July 5 and 6 centred on  $6700 \text{ \AA}$ , achieving a resolving power, measured on arc lines, of  $R \approx 10000$ . On July 7th, we centred the observed region on  $8600 \text{ \AA}$ , obtaining  $R \approx 13000$  spectroscopy over the Ca II triplet range.

Finally, on July 8th 2018, we used the H1800V grating centred on  $4400 \text{ \AA}$  to take classification spectra of two further blue cluster members. By using a 1.4 arcsec slit, we obtained a resolving power of  $R \approx 6500$ .

All the spectra have been reduced according to standard procedure with IRAF<sup>2</sup>. Bias removal, flat-fielding and wavelength calibration were done with frames taken on the same nights as the observations.

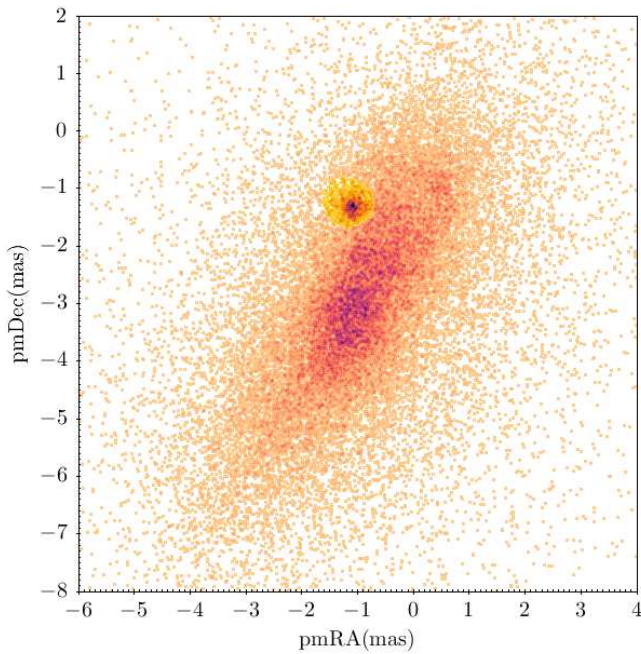
## 3 RESULTS

### 3.1 Cluster definition

*Gaia* DR2 data (Gaia Collaboration et al. 2018) provide precise positions, proper motions and parallaxes for most of the stars brighter than  $G = 20$ . Although there are systematic errors affecting all astrometric parameters (Luri et al. 2018),

<sup>1</sup> Many works indicate a very rapid decline in the number of clusters with age. However, these references use the term “cluster” in a wide sense, mostly referring to OB associations (e.g. Pfalzner et al. 2015); see the discussion on terminology in Krumholz (2014).

<sup>2</sup> IRAF is distributed by the National Optical Astronomy Observatories, which are operated by the Association of Universities for Research in Astronomy, Inc., under cooperative agreement with the National Science Foundation



**Figure 1.** Density map (linear scale) of the input catalogue for Clusterix in the proper motion plane. The yellow circle surrounding the densest clump, around  $(-1.1, -1.3) \text{ mas a}^{-1}$ , indicates the region selected for the initial sample by the Clusterix analysis.

this is a high precision dataset that allows a very good definition of cluster membership. Since the cluster seems to be extended, we downloaded all the DR2 data within a circle of radius 50 arcmin around a nominal centre placed at RA: 18 41 52 Dec:  $-05\ 27\ 00$ , in the midst of the blue stars detected. After removing objects with no astrometric solution or very large errors in the proper motions (error greater than  $1 \text{ mas a}^{-1}$  in any of the two ppms), we examined the proper motion plane, where the cluster is an obvious over-density around ( $\text{pmRA} \equiv \mu_{\alpha} \cos \delta$ ,  $\text{pmDec} \equiv \mu_{\delta}$ )  $\sim (-1.1, -1.3) \text{ mas a}^{-1}$  for any radius taken (cf. Fig. 1).

As this is a very crowded field with differential extinction, defining the extent of the cluster in the proper motion plane is not straightforward. Contamination by field stars is high, as can be seen in Fig. 1<sup>3</sup>. To obtain an initial estimate, we used the Virtual Observatory tool Clusterix 2.0 (Balaguer-Núñez et al. 2020). Clusterix is an interactive web-based application to calculate 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). In its current version, Clusterix works only in the proper motion plane, ignoring all other information. We run the tool interactively by defining concentric circles of changing radius around the nominal centre. For ease of computation, we restricted ourselves to DR2 objects with  $G \leq 16$  and errors in proper

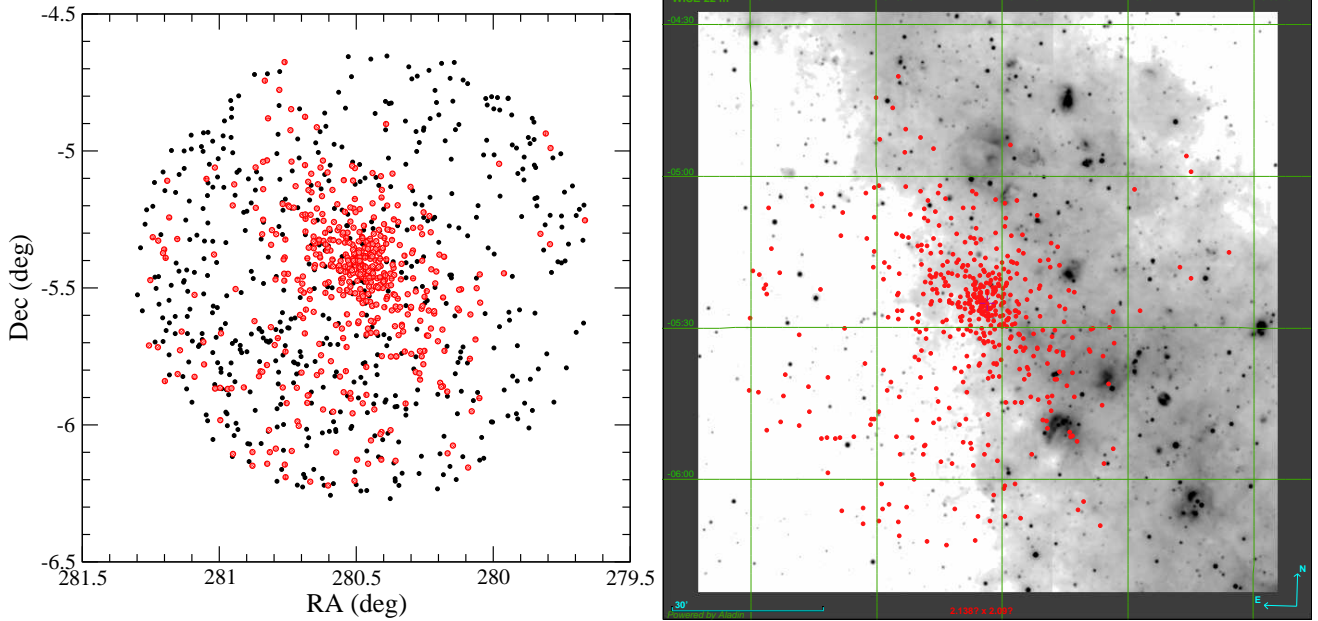
<sup>3</sup> The density of sources in the area is so high that the open cluster Teutsch 145, which is included in our circle, does not stand out at all in Fig. 1. This is a faint and distant cluster for which Cantat-Gaudin et al. (2018) identify around 100 likely members based on DR2 data and give ( $\text{pmRA}$ ,  $\text{pmDec}$ )  $\sim (-1.64, -3.43) \text{ mas a}^{-1}$ .

motion below  $1 \text{ mas a}^{-1}$ . The best results were obtained for a cluster of radius 9 arcmin and a field located between 36 arcmin from the centre and the edge of the dataset. Based on an empirical determination of the frequency functions in the vector point diagram (Sanders 1971), Clusterix assigns each object a probability of belonging to a distinct population, identified as the cluster.

In fact, Clusterix identifies two distinct populations within the area. The first one, representing a weak over-density around ( $\text{pmRA}$ ,  $\text{pmDec}$ )  $\sim (+0.5, -1.0) \text{ mas a}^{-1}$  is not strongly concentrated towards the cluster position, but rather extends over the whole field, with a higher frequency around RA: 18:41:15, Dec:  $-05:26:20$ , a region that seems to show a gap in obscuration. Our data do not allow the identification of this population as a cluster. Nevertheless, Castro-Ginard et al. (2020), by using automated non-parametric techniques, have recently identified a high-confidence cluster candidate in this area (UBC 105) with astrometric parameters  $\text{pmRA} = 0.46 \pm 0.11 \text{ mas a}^{-1}$ ,  $\text{pmDec} = -0.99 \pm 0.09 \text{ mas a}^{-1}$ , which seems to correspond exactly to this population. Therefore we assume that this is a real cluster for which we will use the designation UBC 105 throughout.

The second, much better defined, over-density – easily detected by eye in Fig. 1 – is strongly concentrated towards the area where we had previously identified a group of early-type stars. After analysing the correlation between probability assigned by Clusterix and distribution in the proper motion plane, we made a cut in probability (which is given in arbitrary units) that separates the region shown in Fig. 1. The population selected consists of 1047 objects (out of a total population of over 35 500 objects). The spatial distribution of this selection can be seen in Fig. 2 (left panel), where the objects are marked as black dots, and very clearly demonstrates its clustered nature. However, given the high density of field stars over the whole extent of Fig. 1, the sample must still contain substantial contamination. We performed an initial cleaning by making use of *Gaia* parallaxes. The sample has a mean and median parallax of  $\pi = 0.39 \text{ mas}$  and a mode  $\pi = 0.40 \text{ mas}$ , with standard deviations of 0.05 mas. This excellent agreement confirms the existence of a well-defined population. We analysed the distribution of objects in the  $\text{pmRA}/\text{plx}$  and  $\text{pmDE}/\text{plx}$  plane to identify the locus of the population, and then proceeded to clean iteratively the sample of outliers. Next, we plotted the *Gaia* CMD for the remaining sample (678 stars), which is shown in Fig. 3. The CMD confirms that most of the stars selected are part of a single population, with a typical stellar sequence around  $BP - RP \approx 1.2$ , although considerably broadened by differential reddening, and an apparent clump of bright, red objects. The reality of this clump of evolved stars is substantiated by *Gaia* radial velocities (RVs; Katz et al. 2019). There are five stars brighter than  $G \approx 11$  that have *Gaia* RVs, and they are all consistent with a given value (see Table 2). Nevertheless, substantial contamination can be seen at fainter magnitudes for  $BP - RP > 1.6$ . In addition, there is a hint of a second population around  $G \approx 13.5$  with  $BP - RP < 1.0$ . Closer examination, however, shows that all these objects with bluer colours are located in the Southeast quadrant, where extinction seems to be noticeably lower. They are then in all likelihood less reddened cluster members.

In view of this, we clean our sample by removing all stars with  $G > 11.5$  and  $BP - RP > 1.5$ . The median astrometric



**Figure 2.** *Left panel:* Spatial distribution of candidate cluster members selected from *Gaia* data. The black dots represent the initial selection from the probabilities assigned by Clusterix based on proper motions only. The small red circles are the result of the cleaning procedure described in the text, by using parallaxes and *Gaia* photometry (i.e. position in the CMD). *Right panel:* The same clean sample is shown on top of the WISE 22  $\mu\text{m}$  map of the region. The anticorrelation between likely member density and dust emission is strong.

parameters of the remaining sample are  $(\text{pmRA}, \text{pmDec}) = (-1.08 \pm 0.11, -1.32 \pm 0.11) \text{ mas a}^{-1}$ ,  $\pi = 0.40 \pm 0.05 \text{ mas}$ , where the errors represent the standard deviations of the sample. After removal of a handful of outliers, we are left with a sample of 479 objects down to  $G = 16$ , compatible within their errors at  $2\sigma$  with these average values. Their spatial distribution, also showed in Fig. 2, is much more concentrated than that of the initial sample. The almost complete absence of objects in the Western third of the circle correlates strongly with the presence of dark clouds in the DSS images of the area (cf. Fig. 2, right panel).

We used the Automated Stellar Cluster Analysis (ASteCA) code (Perren et al. 2015) to explore the cluster’s geometry. Running ASteCA on both the initial and cleaned samples displayed in Fig. 2 (left panel), we fit a King profile. The centre of the cluster is found at  $\text{RA} = 280^{\circ}.47$  (18 41 53),  $\text{Dec} = -5^{\circ}.42$  (−05 25 12), about 2 arcmin to the North of our initial guess. The radius is given as  $r = 0.15 \pm 0.02 \text{ deg}$ , but this is clearly an underestimation, as the “background” in these samples is made of the same population.

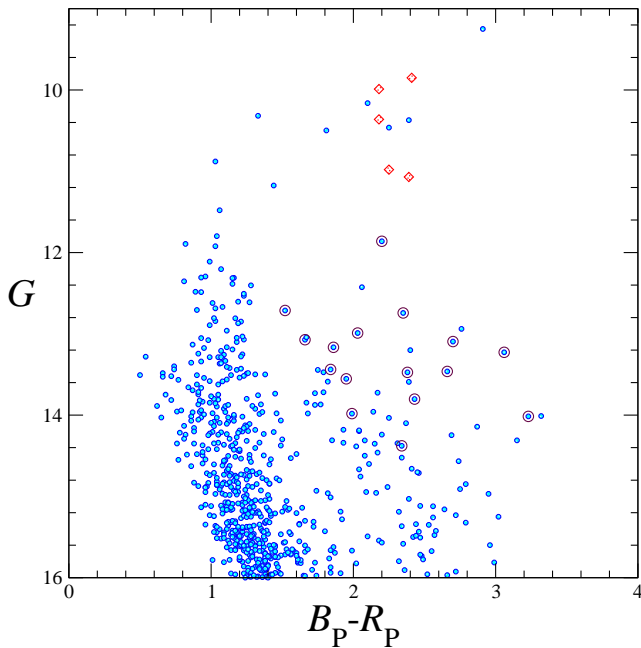
For an independent confirmation, we took 2MASS data within 20 arcmin of this centre. We retained only stars with “good” quality flags (A or E) in all three filters and removed all stars with high errors by taking 0.1 mag in the  $(J - K_S)$  colour as a limit. We selected early type stars by making use of the infrared  $Q$  parameter (see Negueruela & Schurch 2007). We considered a range  $-0.14 < Q_{\text{IR}} < 0.08$  to include the region inhabited by early-type (OBA) stars and take into account the typical photometric errors. We rejected stars brighter than  $K = 9$ , as these will almost certainly be foreground objects. This procedure leaves 1864 candidate early-type stars. Again ASteCA detects a strong overdensity, this time centred on  $\text{RA} = 280^{\circ}.49$ ,  $\text{Dec} = -5^{\circ}.41$ , about one arcmin to the northeast of the centre determined from *Gaia*

candidates. Given that the sample and method are completely independent from the *Gaia* data, this confirms the definition of the cluster. The King profile, however, does not result in a good fit, as the spatial distribution is very asymmetric. Again, the reason for this is evident in Fig. 2 (right panel). There is a number of dark regions immediately to the West of the cluster core that significantly reduce stars counts to that side. In particular, the density maps of both *Gaia* sources and 2MASS sources show a strip starting close to the cluster centre and extending to the southwest for about half a degree with a much lower source count than neighbouring areas. Optical images show that this strip corresponds to a dark lane, which must be foreground to the cluster. In view of these complications, we do not attempt to define a real cluster centre, although we note that members are very strongly concentrated towards the position used as nominal centre.

Finally, we note that Castro-Ginard et al. (2020) identify a high-confidence cluster candidate in this area, UBC 106, with astrometric parameters  $\text{pmRA} = -1.10 \pm 0.10 \text{ mas a}^{-1}$ ,  $\text{pmDec} = 1.32 \pm 0.11 \text{ mas a}^{-1}$ ,  $\pi = 0.40 \pm 0.04 \text{ mas}$ . Given the total correspondence in all parameters, this cluster candidate UBC 106 is identical to our Valparaiso 1, offering independent confirmation of its existence and characteristics.

### 3.2 Stellar content

Classification spectra of stars observed with the DuPont telescope are shown in Fig. 4. Classification was performed using classical criteria by comparison to a set of spectra of standard stars (as discussed in Negueruela et al. 2019) degraded to the same resolution. Since the target selection was made at the time by using only 2MASS CMDs, there were a few F-type interlopers (not shown). The spectral classification of the B-type stars is shown in Table 1, together with their *Gaia* DR2



**Figure 3.** *Gaia* CMD for the sample selected using astrometric data. The diamonds indicate the five bright stars with compatible RVs. Large circles show other stars with *Gaia* RVs within the sample. Only one of them has a RV compatible with the brighter group.

astrometric parameters. The majority of the stars selected turn out to be B5–7 giants and subgiants, as is typical of a moderately young cluster. LS IV  $-05^{\circ}12$  is brighter, and at B9.5 II falls a bit short of supergiant luminosity. TYC 5121-570-1 has a spectral type B2 IV and different proper motions, so we consider it a clear non-member. Star 9 has proper motions compatible with membership, but its parallax ( $\pi = 0.68$ ) and spectral type B9.5 III suggest it is a foreground object. All the other stars have very similar spectra and hence spectral classifications. TYC 5121-533-1, however, is far from the central concentration and has very different proper motions. In fact, according to Castro-Ginard et al. (2020), TYC 5121-533-1 is the second brightest likely member of UBC 105.

All the other stars have astrometric parameters compatible with cluster average to within three sigma, and so we do not discard their membership. TYC 5121-381-1 has a high parallax ( $\pi = 0.65$ ), but this is a Be star, and so this value could be affected by the presence of an envelope. In fact, its EDR3 parallax (Lindegren et al. 2020) is a much shorter  $0.56 \pm 0.03$ . We keep it as a possible member until further *Gaia* releases. Two other stars in the sample are also clearly Be stars. We note that Be stars with late-B types tend to have weak emission features (cf. Steele et al. 1999) and so we could be missing some other Be stars that are only detectable in  $H\alpha$  (as we have not observed this spectral range). The brightest star in our sample, LS IV  $-05^{\circ}12$ , is not picked as a member by Castro-Ginard et al. (2020) because of its poor DR2 astrometric parameters, but its EDR3 values bring it close to the cluster average.

To check the validity of the astrometric selection adopted to define the cluster, we observed two of the brightest blue stars chosen as astrometric members outside the central concentration with the INT + IDS. These two stars (101 and 109 in Table 1) have the same spectral types as the members

of similar brightness in the central concentration, confirming that the cluster extends well beyond the central 5 arcmin. Star 109 may be a mild Be star, but an  $H\alpha$  spectrum will be needed to confirm this.

The nine objects occupying positions compatible with being luminous cool stars in Fig. 3 were observed in the Ca II triplet (CaT) range for classification purposes. Criteria to classify late-type stars in this range are discussed in Negueruela et al. (2011) and Dorda et al. (2016). Their spectra, displayed in Fig. 5, confirm that they are all luminous cool stars, straddling the limit between bright giants and supergiants, as seen in moderately young (i.e.  $\sim 50$ – $100$  Ma) clusters (e.g. Negueruela & Marco 2012; Alonso-Santiago et al. 2017; Negueruela et al. 2018).

### 3.3 Spectroscopic analysis

The spectra of cool luminous stars can be used to determine stellar properties. For the ten objects listed in Table 2 we carried out such an analysis. Because of the poor weather during the night and low response of the CCD in the Ca II region, these triplet spectra have a low signal-to-noise ratio and we used them for spectral classification only. We used the spectra in the  $H\alpha$  region to compute effective temperature ( $T_{\text{eff}}$ ), surface gravity ( $\log g$ ), and metallicity ( $[M/H]$ ) by comparing them to a grid of synthetic spectra (as in Lohr et al. 2018). We employed the new version of the code STEPARSYN (see Tabernero et al. 2018, 2021, Tabernero et al., in prep.), which uses a Markov-chain Monte Carlo algorithm (emcee, see Foreman-Mackey et al. 2013) for optimisation. We explored the parameter space by using 12 Markov-Chains of 1000 points.

We generated a grid of synthetic spectra by means of the MARCS spherical atmospheric models (Gustafsson et al. 2008) and the radiative transfer code *spectrum* (Gray & Corbally 1994). As line list, we employed a selection from the VALD3 database (Piskunov et al. 1995; Kupka et al. 2000; Ryabchikova et al. 2015), which takes into account all the relevant atomic and molecular features (dominated by TiO and CN) that can appear in cool luminous stars earlier than mid-M types. We employed the Anstee, Barklem, and O’Mara theory (ABO) as Van der Waals damping prescription, when available in VALD3 (see Barklem et al. 2000). The grid spans  $T_{\text{eff}}$  from 3500 K to 8000 K with a step of 250 K above 4000 K and 100 K otherwise. The metallicity ranges from  $[M/H] = -1.0$  dex to  $[M/H] = 1.0$  dex in 0.25 dex steps, while surface gravity goes from  $-0.5$  to 2.0 dex in 0.5 dex steps, when available in the MARCS grid. The microturbulence ( $\xi$ ) was adjusted according to the 3D model based calibration described in Dutra-Ferreira et al. (2016). We convolved our grid of synthetic spectra with a Gaussian kernel to account for the line-spread function of the instrument. The results of the analysis are listed in Table 2.

The metallicities derived for all the cool stars are roughly compatible with a solar metallicity. The only exception is the Cepheid variable CM Sct, which gives a decidedly supersolar metallicity. This is in open contrast to the slightly supersolar values reported by Luck (2018). A similarly large discrepancy is seen between this authors’ value for  $\log g$  (2.1) and ours (1.0). Since metallicity and effective gravity are highly degenerate, we excluded CM Sct from the calculation of the

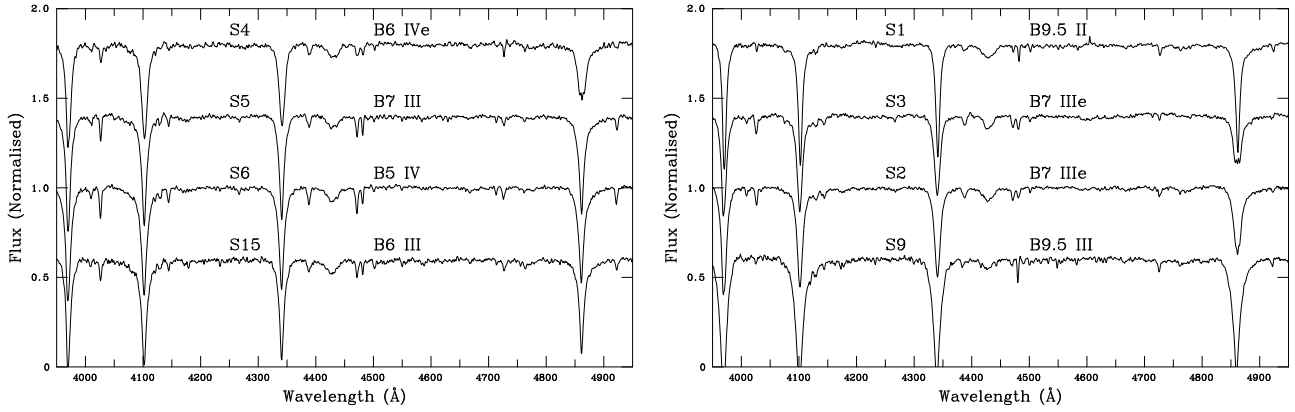


Figure 4. Intermediate-resolution spectra of blue stars in the field of Valparaiso 1.

Table 1. Observed parameters for blue stars with classification spectra.

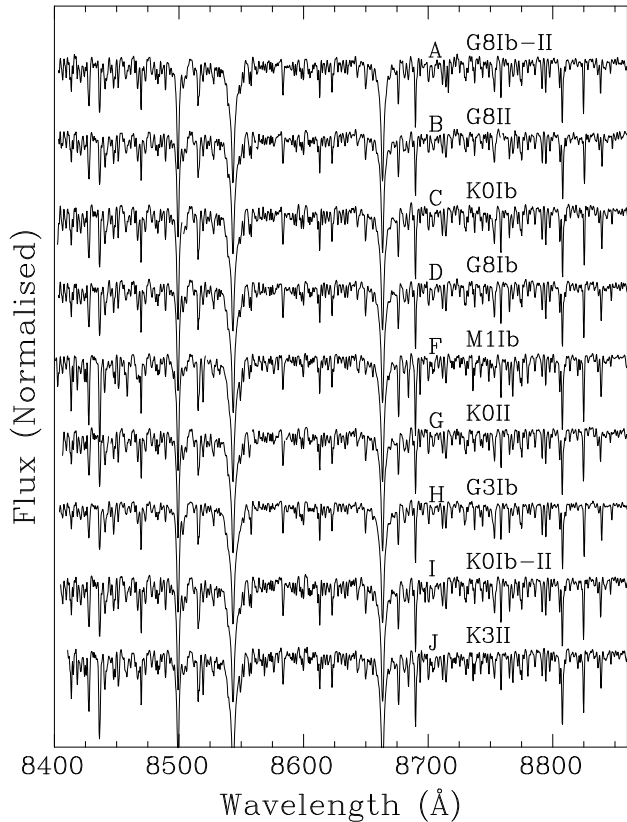
<i>Gaia</i>	RA	Dec	Other name	Spectral type	pm (RA) (mas)	pm (Dec) (mas)	$\pi$ (mas)	<i>G</i> (mag)	<i>BP-RP</i> (mag)	Member <sup>†</sup>
1	18:41:52.26	-05:23:22.1	LS IV -05°12	B9.5 II	$-0.75 \pm 0.14$	$-1.12 \pm 0.13$	$0.43 \pm 0.07$	10.32	1.32	L
2	18:41:52.13	-05:26:58.1	TYC 5121-381-1	B7 IIIe	$-1.06 \pm 0.09$	$-1.53 \pm 0.09$	$0.64 \pm 0.05$	11.92	1.11	P
3	18:41:47.13	-05:27:55.2	TYC 5121-662-1	B7 IIIe	$-0.86 \pm 0.08$	$-1.66 \pm 0.08$	$0.35 \pm 0.05$	11.85	0.98	L
4	18:41:55.90	-05:26:58.5		B6 IVe	$-1.10 \pm 0.08$	$-1.30 \pm 0.08$	$0.41 \pm 0.05$	12.53	1.23	C
5	18:41:52.27	-05:26:08.6		B7 III	$-1.16 \pm 0.07$	$-1.35 \pm 0.07$	$0.37 \pm 0.04$	12.60	1.19	C
6	18:41:52.50	-05:25:15.1		B5 IV	$-0.97 \pm 0.07$	$-1.39 \pm 0.07$	$0.44 \pm 0.04$	12.51	1.23	C
8	18:41:48.20	-05:26:46.2		B5 IV	$-0.42 \pm 0.08$	$-0.53 \pm 0.07$	$0.48 \pm 0.05$	12.42	0.90	P
9	18:41:47.66	-05:25:53.1		B9.5 III	$-0.82 \pm 0.08$	$-1.28 \pm 0.08$	$0.68 \pm 0.05$	12.23	1.18	U
11	18:41:32.80	-05:25:15.5	TYC 5121-570-1	B2 IV	$-0.91 \pm 0.09$	$-3.69 \pm 0.08$	$0.45 \pm 0.05$	11.70	0.86	N
13	18:41:26.06	-05:24:54.9	TYC 5121-533-1	B5 III	$+0.42 \pm 0.10$	$-0.97 \pm 0.09$	$0.49 \pm 0.05$	11.97	1.00	N
15	18:41:20.28	-05:24:43.1	TYC 5121-63-1	B6 III	$-1.04 \pm 0.07$	$-1.28 \pm 0.07$	$0.42 \pm 0.04$	11.92	1.03	C
101	18:41:27.37	-05:31:23.1		B7 III	$-1.22 \pm 0.07$	$-1.31 \pm 0.07$	$0.43 \pm 0.04$	11.80	1.04	C
109	18:41:51.14	-05:19:55.6		B6 IIIe?	$-1.13 \pm 0.09$	$-1.27 \pm 0.08$	$0.40 \pm 0.05$	12.31	1.15	C

<sup>†</sup> Membership is considered C(ertain) for stars with astrometric parameters compatible with the cluster average and appropriate spectral types; (L)ikely for stars with appropriate spectral types whose astrometric parameters are compatible at  $2\sigma$  within their respective errors; (P)ossible for a star with appropriate spectral type and the right parallax, but divergent proper motion; (U)nlikely for a star with divergent parallax and a spectral type B9.5 III, which does not agree with the rest of members. The B2 IV star and the member of UBC 105, Star 13, are considered certain (N)on-members.

Table 2. Stellar parameters for cool stars in the field of Valparaiso 1.

Star	Other name	<i>Gaia</i> DR2	Spectral type	$T_{\text{eff}}$ (K)	$\log g$	[M/H] (dex)	$v_{\text{hel}}$ (km s <sup>-1</sup> )	RV ( <i>Gaia</i> ) (km s <sup>-1</sup> )
A	TYC 5121-543-1	4256511915482900608	G8 Ib-II	$4693 \pm 46$	$1.1 \pm 0.11$	$-0.05 \pm 0.06$	$40.6 \pm 0.2$	—
B	GSC 05121-00622	4256512843232515840	G8 II	$4620 \pm 51$	$1.42 \pm 0.12$	$+0.01 \pm 0.07$	$42.4 \pm 0.2$	$40.1 \pm 0.4$
C	TYC 5121-819-1	4253508943153458048	K0 Ib	$4639 \pm 40$	$0.88 \pm 0.11$	$+0.02 \pm 0.06$	$39.8 \pm 0.2$	—
D	TYC 5121-218-1	4253508702635208832	G8 Ib	$4640 \pm 48$	$1.02 \pm 0.11$	$-0.07 \pm 0.06$	$41.1 \pm 0.2$	$40.8 \pm 0.5$
E	CM Sct	4253603501158148736	- <sup>†</sup>	$5431 \pm 36$	$1.03 \pm 0.09$	$-0.15 \pm 0.04$	$47.4 \pm 0.3$	—
F	TYC 5121-758-1	4253603501158148736	M1 Ib	$3840 \pm 20$	$0.33 \pm 0.09$	$-0.10 \pm 0.05$	$41.6 \pm 0.2$	—
G		4256511468842481408	K0 II	$4725 \pm 44$	$1.33 \pm 0.09$	$+0.12 \pm 0.05$	$41.5 \pm 0.2$	$41.6 \pm 0.4$
H	TYC 5121-684-1	4253603501158148736	G3 Ib	$5105 \pm 27$	$0.72 \pm 0.08$	$-0.07 \pm 0.04$	$41.1 \pm 0.2$	$41.8 \pm 0.2$
I	TYC 5125-1531-1	4253499219346450432	K0 Ib-II	$4755 \pm 22$	$0.91 \pm 0.06$	$+0.08 \pm 0.03$	$43.5 \pm 0.6$	$42.1 \pm 0.3$
J		4253597556923196672	K3 II	$4137 \pm 40$	$0.65 \pm 0.1$	$-0.06 \pm 0.06$	$43.4 \pm 0.2$	—

<sup>†</sup> CM Sct is a classical Cepheid and thus spectral variable. Therefore we did not observe it in the CaT range.



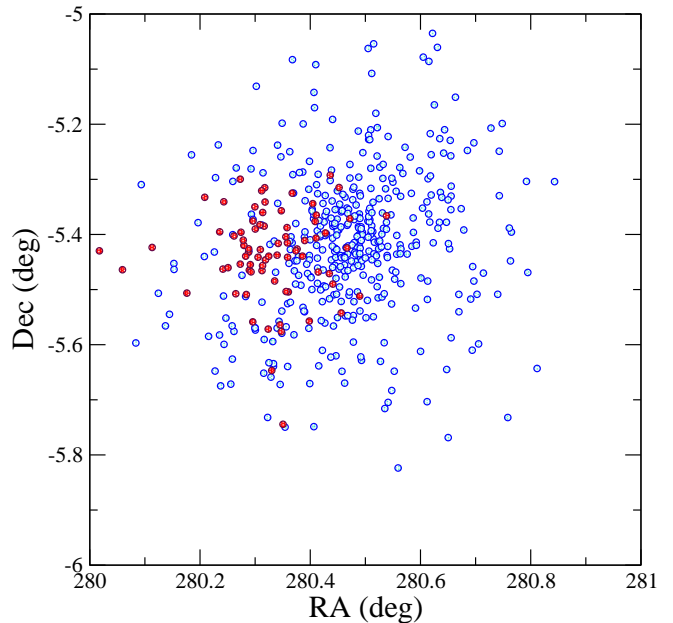
**Figure 5.** Spectra of nine cool luminous members of Valparaiso 1, used for spectral classification.

average metallicity, which, after weighing each value with its error, is exactly solar. In view of this, we use solar-metallicity isochrones in the subsequent analysis.

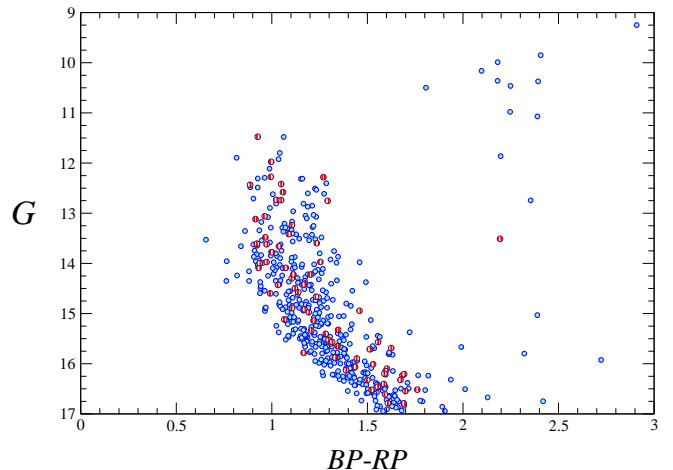
RVs were measured by cross-correlating each individual observation against the Arcturus atlas (Hinkle et al. 2000) by means of the iSPEC code (Blanco-Cuaresma et al. 2014). Comparison of our RVs to *Gaia* DR2 values for five stars in common shows excellent agreement. Our formal errors are clearly an underestimation, as systematic effects preclude errors better than  $\sim 3 \text{ km s}^{-1}$  for an instrument with such large flexures as IDS. All ten stars display consistent values of RV. Only CM Sct appears slightly deviant, but this well-studied Cepheid is known to display RVs ranging from  $\sim +23$  to  $\sim +55 \text{ km s}^{-1}$  (Metzger et al. 1991), with a systemic velocity of  $+40.8 \text{ km s}^{-1}$  (Pont et al. 1994), fully consistent with the others. Removing the variable, our average RV value is  $+41.7 \text{ km s}^{-1}$  ( $\sigma = 1.2 \text{ km s}^{-1}$  for nine stars), while the *Gaia* DR2 unweighted average is  $+41.3 \text{ km s}^{-1}$  ( $\sigma = 0.8 \text{ km s}^{-1}$  for five stars).

#### 4 CLUSTER PARAMETERS

The derivation of cluster parameters is complicated. The cluster is large on the sky and extinction varies strongly across its surface. The photometric diagrams are complicated by the presence of other populations. In particular, UBC 105 is another young cluster at about the same dis-



**Figure 6.** Spatial distribution of likely members (according to Castro-Ginard et al. 2020) of UBC 106 = Valparaiso 1 (blue circles) and UBC 105 (dashed red circles). Although the two clusters overlap, Valparaiso 1 is much more populous and extends over a vast area that encompasses UBC 105.

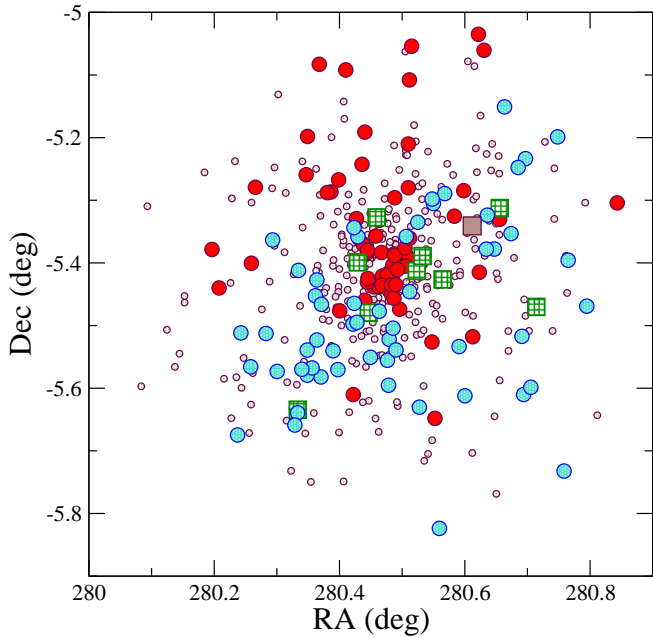


**Figure 7.** Colour magnitude diagram, using *Gaia* DR2 photometry of likely members (Castro-Ginard et al. 2020) of Valparaiso 1 (blue circles) and UBC 105 (dashed red circles). The two populations occupy indistinguishable positions in this diagram.

tance ( $\pi = 0.47 \pm 0.03 \text{ mas}$ ; Castro-Ginard et al. 2020). Although the two clusters are very well separated in the proper motion plane, they are spatially coincident, as seen in Fig. 6.

Moreover, the two clusters have similar reddening, and their populations blend in all photometric diagrams. As seen in Fig. 7, they must have about the same age, as also suggested by the fact that TYC 5121-533-1, one of the brightest members of UBC 105, has the same spectral type as the brightest members of Valparaiso 1.

The effect of differential extinction is obvious in Fig. 7. Astrometric cluster members of a given brightness have



**Figure 8.** Spatial distribution of highly-reddened (red circles) and less reddened (blue circles) early-type members of Valparaiso 1. Fainter members are shown as smaller purple circles. Cool luminous stars appear as green squares, while the Cepheid CM Sct is the filled brown square.

( $BP - RP$ ) colours ranging over more than half a magnitude. In fact, it is possible to identify two almost separate branches in the upper part of the CMD. To investigate their meaning, we divided blue members of Valparaiso 1 brighter than  $G \approx 14.5$  into two groups, according to their position in the CMD. Stars with  $BP - RP > 1.05$  were assigned to the high-reddening group, while stars with lower values were tagged as low reddening. Fig. 8 shows the spatial distribution of the two groups. Although there is a fair amount of blending, the spatial segregation is strong. All the stars in the core region belong to the high-reddening group, together with the population to the northwest. This distribution clearly shows that the clouds visible in the optical images contribute large amounts of obscuration to the cluster core.

#### 4.1 *Gaia* photometry

In view of these difficulties, we have decided to carry out the analysis by utilising only the list of likely members determined by [Castro-Ginard et al. \(2020\)](#), as their criteria are stricter than ours – we identify 479 objects down to  $G = 16$  as high-probability members, while they only list 431 members down to  $G = 17$ . Of the B-type stars listed in Table 1, only those that we chose as certain members are included in their list. None of our likely or possible members is. As seen in Fig. 7, less than ten of the 431 astrometric members occupy positions incompatible with membership in the *Gaia* CMD. However, the width of the main sequence noticeably decreases for  $G > 15$ . Given that reddening is expected to affect equally stars of all magnitudes, this effect is very likely suggesting that many faint stars are being rejected as high probability members by the procedure employed by [Castro-Ginard et al.](#)

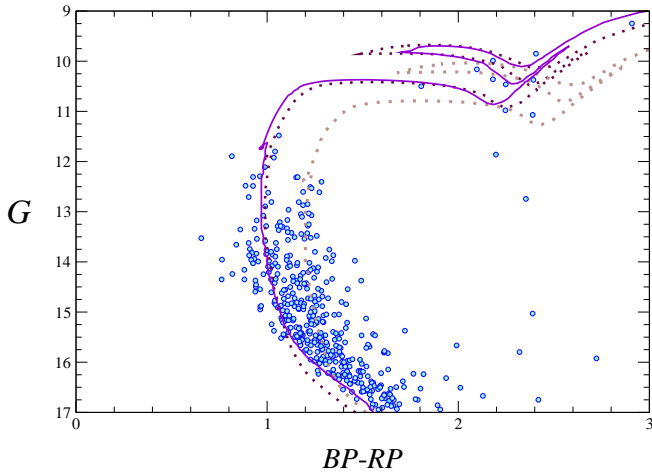
(2020) because of large astrometric errors (note that this effect is not present in the diagram presented in Fig. 3, where we select likely members simply by taking astrometric parameters within  $2\sigma$  of the cluster averages). In consequence, we cannot consider the list of members that we are using complete, at least for faint magnitudes. The reddening to the evolved members is more difficult to measure, but their spatial distribution (shown in Fig. 8) suggests that we can expect to find them at both high and low reddening.

With such strong differential reddening, isochrone fitting becomes very difficult. We can use the *Gaia* DR2 parallax to provide an estimate of the distance, although being aware of the possibility of large systematic errors ([Luri et al. 2018](#)). A  $\pi = 0.40$ , together with a global systematic error of  $-0.03$ , implies a distance of 2.3 kpc, i.e., a distance modulus  $DM = 11.8$ . In Fig. 9, we show a number of isochrones that can fit part of the cluster population. We downloaded isochrones from the Padova PARSEC server ([Bressan et al. 2012](#)), selecting a Kroupa IMF corrected for binaries and the *Gaia* passbands from [Maíz Apellániz & Weiler \(2018\)](#). A 70 Ma isochrone reddened with  $A_V = 2.7$  can fit the upper main sequence and the cool luminous giants for  $DM = 12.0$ . However, it fails to fit the fainter part of the main sequence. Given that there seems to be a substantial lack of members in this region, the importance of this failure is difficult to assess. A 90 Ma isochrone reddened with  $A_V = 2.6$  gives a very good fit to the less reddened main sequence, and provides a decent fit to the evolved stars. Much higher extinction is needed to fit the more heavily reddened population. The 90 Ma isochrone requires  $A_V = 3.1$  to fit the more reddened blue members. But then a  $DM$  no longer than 11.2 mag is required to reach the position of the evolved stars. This would be equivalent to  $\pi = 0.59$  mas, a value incompatible with the *Gaia* DR2 determination, even if all sources of uncertainty are taken into account. On the other hand, the 70 Ma isochrone requires  $A_V = 3.2$  to fit the stars from the core region, and reaches the position of the evolved stars for  $DM$  in the 11.6 to 11.8 mag range, as illustrated in Fig. 9. Therefore, if we assume a single age and give value to the *Gaia* DR2 distance, the older age is strongly disfavoured and we are left with an age in the range 70 to 80 Ma, and a distance modulus in the 11.6–12.0 mag range ( $2.1 \text{ kpc} \lesssim d \lesssim 2.5 \text{ kpc}$ ).

Along this line of sight, the Sagittarius arm is believed to run almost perpendicularly at around 2 kpc (see [Marco & Negueruela 2011](#), for a detailed discussion). Therefore the shorter distance is favoured in terms of Galactic structure. We note the presence of the similarly aged open clusters NGC 6664 and NGC 6649 at about the same distance ([Alonso-Santiago et al. 2020](#)) in the same region of the sky. The older clusters M11 and NGC 6704, which are very close in the sky to Valparaiso 1, are also located at about the same distance ([Cantat-Gaudin et al. 2018](#)). Interestingly, within 2 degrees of Valparaiso 1, [Castro-Ginard et al. \(2020\)](#) find five cluster candidates with very similar parallaxes. Apart from UBC 105, discussed above, UBC 109 and UBC 349 are also well populated clusters, whose *Gaia* DR2 CMD suggest somewhat older ages than Valparaiso 1, while UBC 353 and UBC 355 are poorly populated candidates.

Finally, we note that [Cantat-Gaudin et al. \(2020\)](#), by using a machine-learning automated procedure, derive a distance modulus  $DM = 11.9$  and age of 160 Ma for UBC 106, which we identify with Valparaiso 1. However, this fit is obtained





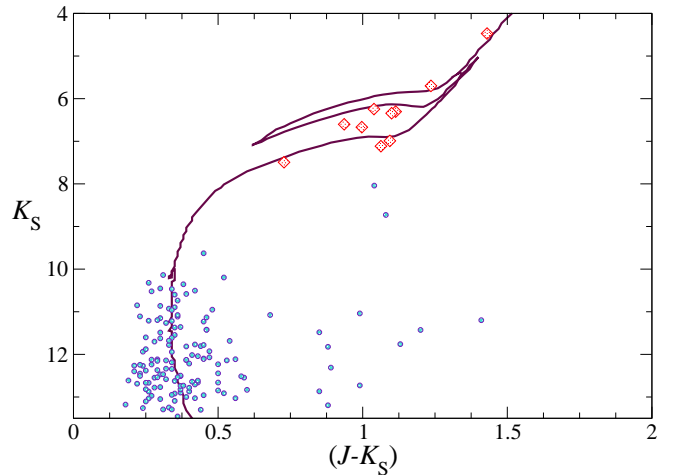
**Figure 9.** Possible PARSEC isochrone fits for the *Gaia* DR2 CMD of Valparaiso 1. The solid line is a 90 Ma isochrone reddened with  $A_V = 2.6$  and displaced to  $DM = 11.6$ . The dashed line is a 70 Ma isochrone with  $A_V = 2.7$  displaced to  $DM = 12.0$ . Finally, the dotted line is the same 70 Ma isochrone with  $A_V = 3.2$  displaced to  $DM = 11.8$ . No 90 Ma isochrone can fit the high-reddening branch for a distance compatible with the *Gaia* parallax.

for an extinction  $A_V$  of only 1.9 mag, suggesting that the automated procedure has only taken into account the small subset of less reddened members of the cluster.

#### 4.2 2MASS photometry

For a complementary analysis, we used near-infrared photometry. We obtained  $JHK_S$  photometry from the 2MASS catalogue (Skrutskie et al. 2006). The completeness limit of this catalogue is set at  $K_S = 14.2$ . Nevertheless, in this crowded region of the Galactic Plane, a very large fraction of objects are affected by crowding, resulting in bad quality flags and measurements marked as upper limits (flag U) even for bright objects. In consequence, we did not attempt to use the whole 2MASS dataset for the region, but rather only the photometry available for the objects selected from *Gaia* DR2 data. We performed a cross-match over the whole region of interest between the members selected by Castro-Ginard et al. (2020) and 2MASS, using a match distance of only 0.2 arcsec, which seems appropriate given the high density of 2MASS sources and the accuracy of astrometry in both catalogues. Even so, the resulting selection contains a very large number of faint objects close to the detection limit of 2MASS. In consequence, the list was cleaned by removing all objects with poor photometry (upper limits or errors higher than 0.1 mag) in any of the three filters. Although almost two thirds of the sources are rejected, we are still left with 150 objects with good photometry.

The corresponding  $(J - K_S) - K_S$  diagram is displayed in Fig. 10. Although differential reddening should be less important in the near infrared, photometric accuracy is worse in the 2MASS catalogue than in *Gaia* DR2. As a consequence, there is still a 0.3 mag spread in  $(J - K_S)$  at a given magnitude. No object fainter than  $K_S = 13.5$  survives our quality cuts, and a few of the matches are very clearly incorrect. Even so, the cluster sequence is well defined. An isochrone



**Figure 10.** Near-IR CMD for Valparaiso 1. The blue circles represent 2MASS cross-matches for members according to Castro-Ginard et al. (2020) that have well-defined photometric errors. The diamonds indicate the ten astrometric and RV cool luminous members. A representative isochrone is shown, with the preferred parameters from the optical-data fit (75 Ma and  $DM = 11.8$ ) and a global extinction  $A_V = 2.7$  mag.

with the preferred values from the optical photometry analysis, i.e.  $\tau = 75$  Ma,  $\mu = 11.8$  is shown. The fit is not bad, although we must point out that a value  $A_V = 2.7$  was used to redden the isochrone. This is somewhat lower than the values used in the fit to *Gaia* data, but the difference is small. It could be pointing to a small deviation from the standard reddening law, but it could also be simply reflecting the inability of the simple polynomial fit of the Cardelli et al. (1989) law to reproduce the effect of high extinction values. A much better characterisation of the extinction law will be possible with the DR3 spectrophotometric data.

## 5 DISCUSSION

We have identified a large and populous open cluster in the constellation of Scutum. Analysis of *Gaia* DR2 data confirms its reality and extent<sup>4</sup>. Its basic parameters are corroborated by an independent analysis by Castro-Ginard et al. (2020). Although very strong differential reddening prevents an accurate isochrone fit, its age is well constrained by the *Gaia* DR2 parallax. Ages younger than 70 Ma or decidedly older than 80 Ma seem incompatible with the parallax, even after accounting for the possible systematic errors. At a nominal age of 75 Ma, a solar metallicity cluster must have its main-sequence turn-off around B5 V, in good agreement with the spectral types that we find for the brightest members (Table 1), although we do not reach deep enough to find any dwarf stars. At this age, the initial mass of the evolved stars is  $\approx 6M_{\odot}$ , according to the isochrones.

<sup>4</sup> After the analysis was completed, *Gaia* EDR3 was published. We have checked that the average cluster values for all astrometric parameters in EDR3 are perfectly compatible within the dispersion given with the DR2 data and preferred to retain the analysis based on the likely members of Castro-Ginard et al. (2020) only

### 5.1 Cluster size

Valparaiso 1 is a very large cluster on the sky (more than 40 arcmin across, if we consider the region over which both our analysis and that of [Castro-Ginard et al. 2020](#) find a non-negligible density of probable members). Given its age, angular size and evolved population, a direct comparison can be made to NGC 6067, one of the most massive clusters of a similar age, with a present day mass of  $\approx 5000 M_{\odot}$  ([Alonso-Santiago et al. 2017](#)). NGC 6067 has a similar size in the sky, with a tidal radius of  $\approx 15$  arcmin at a slightly shorter  $DM = 11.3$  ([Alonso-Santiago et al. 2017](#)). Even with the strict selection criteria of [Castro-Ginard et al. \(2020\)](#), Valparaiso 1 has 190 members brighter than  $G = 15$ . With the parameters derived for the cluster, this limit corresponds to an absolute magnitude  $\approx +0.2$ , with the exact value depending on the reddening. This magnitude is typical of a B9 V star ( $\approx 3 M_{\odot}$ ). Leaving aside the evolved stars, we thus have about 180 mid- and late-B cluster members. This number can be compared to the  $\approx 100$  B-type stars in NGC 6067 ([Alonso-Santiago et al. 2017](#)). NGC 6067 is slightly older and has a higher metallicity than Valparaiso 1. Its main-sequence turn-off is thus found at B6 V. Therefore the main sequence in Valparaiso 1 should extend a bit more towards early types. Even so, the number of B-type members indicates that Valparaiso 1 is almost certainly more massive than NGC 6067. This is even more evident if we resort to the  $2\text{-}\sigma$  membership criterion used in Section 3.1, which gives 257 likely members down to  $G = 15$ . For NGC 6067, [Alonso-Santiago et al. \(2017\)](#) estimate an initial mass above  $7000 M_{\odot}$ .

Such a high number of members (and hence mass) is borne out by the very large extent of Valparaiso 1. Even accepting the cluster radius from the King profile fit of only 9 arcmin, this implies that the cluster is 12 pc across. In fact, we detect members at distances of 20 arcmin from the nominal centre, equivalent to about 13 pc, as also do [Castro-Ginard et al. \(2020\)](#).

### 5.2 CM Sct

CM Sct is a well studied 3.9 d Cepheid, used as part of several period/luminosity calibrations (e.g. [Tammann et al. 2003](#); [Fouqué et al. 2007](#)). It has been generally assumed to be a fundamental mode pulsator, and this is confirmed by the recent reassessment by [Ripepi et al. \(2019\)](#), using DR2 data. For a fundamental mode pulsator, such a pulsation period is expected at an age around 80 Ma, according to the classical calibration of [Bono et al. \(2005\)](#). The accuracy of such aging is debatable, as the models used by [Bono et al. \(2005\)](#) assume a solar metallicity  $Z_{\odot} = 0.02$ , while the PARSEC isochrones that we are using take  $Z_{\odot} = 0.015$ , but the agreement with the cluster age derived from isochrone fitting is very encouraging.

CM Sct is a halo member of Valparaiso 1, lying about 9.3 arcmin from the centre, close to the edge of the cluster determined by the King profile (which we know to underestimate the true cluster size), but substantially closer than many other members. About half of the classical Cepheids confirmed as cluster members are halo members of their respective clusters ([Anderson et al. 2013](#)). At 2.3 kpc, this angular distance from the centre corresponds to 6.2 pc, comparable to the distances of V Cen, EV Sct or QZ Nor to their

respective cluster centres ([Anderson et al. 2013](#)). CM Sct lies to the Northeast of the cluster, a region which, as seen in Fig. 8, seems to have relatively low extinction. In fact, published photometry for the Cepheid indicates a rather lower reddening than the cluster average. [Fernie \(1990\)](#) quotes  $E(B - V) = 0.75$ , while [Laney & Caldwell \(2007\)](#) give only  $E(B - V) = 0.7$ . For a standard reddening law, these values correspond to  $A_V \approx 2.2 - 2.3$ , similar to the less reddened main sequence members.

Thanks to *Gaia* DR2 astrometric parameters, CM Sct joins the growing number of classical Cepheids that belong to open clusters. DR2 data have also allowed the identification of the parental cluster of SV Vul ([Negueruela et al. 2020](#)) and associated two previously unknown Cepheids to the open clusters Berkeley 51 and Berkeley 55 ([Lohr et al. 2018](#)).

### 5.3 The nature of star F

Among the large complement of evolved stars in Valparaiso 1, star F (TYC 5121-758-1) is remarkable for being much redder and brighter than any other member. In fact, its position on the preferred isochrones suggests it is an early AGB star. As such, it would be the first identification of an AGB star in a young open cluster. Recently, [Fragkou et al. \(2019\)](#) found a planetary nebula associated to NGC 6067, as discussed above another massive cluster of about the same age. These connections provide for the first time empirical evidence of the evolutionary paths for  $\sim 6 M_{\odot}$  stars. Until now, it had been assumed that the brightest (O-rich) Galactic AGB stars were descended from moderately massive stars ( $4 - 8 M_{\odot}$ ) because of their population characteristics (e.g. [García-Hernández et al. 2007](#)), but direct comparison to models could only be performed for objects in the Magellanic Clouds. The accurate identification of members allowed by *Gaia* for spatially extended clusters, such as NGC 6067 or Valparaiso 1, is starting to give us valuable examples in the Milky Way itself.

Morphologically, star F looks like a normal low-luminosity M-type supergiant. We find no obvious anomalies within the limited spectral range observed. Formally, we classify it as a supergiant (M1 Ib), because the strength of the luminosity indicators, such as the equivalent width of the Ca II triplet or the blend at 8648 Å (see [Dorda et al. 2016](#), and references therein), put it within this luminosity class. Nevertheless, this classification does not imply any assumption about its mass or internal structure. In fact, its luminosity,  $\log(L_*/L_{\odot}) \approx 3.8$  according to its position on the isochrone, is too low for a real supergiant (i.e. an object with an initial mass above  $8 M_{\odot}$ ).

The presence of strong lithium lines, namely at the location of the Li I 6708 Å doublet, is believed a telltale sign of a massive AGB star ([García-Hernández et al. 2007](#)). According to modern models (e.g. [Mazzitelli et al. 1999](#); [van Raaij et al. 2012](#)), lithium production happens during the early AGB phase and Li surface abundance reaches a maximum during the early thermal pulse phase. We do not detect a strong line at the position of the Li I doublet. We observe a weak feature, typical of stars of the spectral type, which may include a contribution of the Li line, but may also be due to a number of neighbouring metallic lines. Unfortunately, our spectrum lacks the resolution needed for spectral synthesis and an estimate of the Li abundance. In any event, the lack of a strong Li feature does not rule out the early-AGB na-

ture of star F, as its position on the isochrone suggests it has only recently entered the asymptotic branch. An alternative explanation as a red supergiant straggler formed by binary interaction seems ruled out by its low luminosity.

## 6 CONCLUSIONS

*Gaia* astrometric data reveals that an inconspicuous grouping of mid-B giants is in reality the core of a very large open cluster, which we call Valparaíso 1, comparable in size and mass to the most massive open clusters in the solar neighbourhood. Ten luminous bright giants or supergiants are members of the cluster, including the well-known 4-d Cepheid CM Sct. Analysis of their intermediate-resolution spectra reveals a solar-like metallicity. Although strong differential reddening renders accurate determination of cluster parameters unfeasible with the current dataset of *Gaia* and 2MASS photometry, the *Gaia* parallax – implying a distance of  $2.3^{+0.7}_{-0.4}$  kpc, where the errors take into account possible systematic effects – is only compatible with ages no younger than  $\sim 70$  Ma and not much older than  $\sim 80$  Ma. Such an age is compatible with the complement of B-type giants observed (in the range B5–B7 III) and implies a mass of  $\approx 6M_{\odot}$  for the more evolved stars.

The similarly aged clusters NGC 6649 and NGC 6664 are pretty close in the sky, lying at about the same distance and also hosting Cepheids. Both clusters contain about 60 B-type stars, implying initial masses somewhat below  $3000M_{\odot}$  (Alonso-Santiago et al. 2020). By comparison, the  $> 180$  B-type stars in Valparaíso 1 suggest an initial mass approaching  $10000M_{\odot}$ . This number is also borne out by direct comparison to the similarly large NGC 6067 (Alonso-Santiago et al. 2017). Both NGC 6649 and NGC 6644 have been studied for decades, while the much larger Valparaíso 1 remained unknown, despite the attention drawn by its Cepheid member, CM Sct. This suggests that confusion due to high foreground and background contamination has been a major issue preventing the identification of clusters without a strong central concentration. In fact, many of the new open clusters found by Castro-Ginard et al. (2020) are characterised by low stellar densities and a dispersed structure. In view of this, it is highly likely that substantial numbers of open clusters remain hidden in areas of high stellar density towards the inner Galaxy. As discussed in the Introduction, the age distribution of known massive clusters suggests that some of them may be very massive.

Unlike in younger clusters (e.g. vdBH 222; Marco et al. 2014), the bright evolved members of Valparaíso 1 do not stand out at all over their neighbours. Although they are among the brightest stars in the field, there are several dozen stars of comparable (or higher) brightness in the  $K_S$  band with similar colours. In view of this, it is highly likely that future searches for clusters based on DR3 and successive releases will find more moderately-young and intermediate-age populous clusters.

Meanwhile, Valparaíso 1 offers us a precious laboratory to study the evolution of intermediate-mass stars, with star F very likely representing a  $6M_{\odot}$  star entering the AGB. Broadband photometry and DR3 spectrophotometry will allow a much more accurate determination of the extinction and hence cluster parameters.

## DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

## ACKNOWLEDGEMENTS

We thank the anonymous referee for helpful comments. This research is partially supported by the Spanish Government under grants AYA2015-68012-C2-2-P and PGC2018-093741-B-C21 (MICIU/AEI/FEDER, UE). HMT is also supported by FCT - Fundação para a Ciência e a Tecnologia through national funds and by FEDER through COMPETE2020 - Programa Operacional Competitividade e Internacionalização by these grants: UID/FIS/04434/2019, UIDB/04434/2020; UIDP/04434/2020, PTDC/FIS-AST/28953/2017, and POCI-01-0145-FEDER-028953. ANC is supported by the international Gemini Observatory, a program of NSF's NOIRLab, which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation, on behalf of the Gemini partnership of Argentina, Brazil, Canada, Chile, the Republic of Korea, and the United States of America. JB and RK are partially funded by ANID – Millennium Science Initiative Program – ICN12\_009 awarded to the Millennium Institute of Astrophysics MAS

The INT is operated on the island of La Palma by the Isaac Newton Group, installed in the Spanish Observatorio del Roque de Los Muchachos of the Instituto de Astrofísica de Canarias. The Starlink software (Currie et al. 2014) is currently supported by the East Asian Observatory. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement.

This research has made use of the Simbad, Vizier and Aladin services developed at the Centre de Données Astronomiques de Strasbourg, France. It also makes use of data products from the Two Micron All Sky Survey, which is a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

## REFERENCES

- Alonso-Santiago J., Negueruela I., Marco A., Tabernero H. M., González-Fernández C., Castro N., 2017, *MNRAS*, **469**, 1330
- Alonso-Santiago J., Negueruela I., Marco A., Tabernero H. M., Castro N., 2020, arXiv e-prints, p. [arXiv:2009.12418](https://arxiv.org/abs/2009.12418)
- Anderson R. I., Eyer L., Mowlavi N., 2013, *MNRAS*, **434**, 2238
- Balaguer-Núñez L., et al., 2020, *MNRAS*, **492**, 5811
- Barklem P. S., Piskunov N., O'Mara B. J., 2000, *A&AS*, **142**, 467
- Bavarsad E. A., Sandquist E. L., Shetrone M. D., Orosz J. A., 2016, *ApJ*, **831**, 48
- Blanco-Cuaresma S., Soubiran C., Jofré P., Heiter U., 2014, in *Astronomical Society of India Conference Series*. pp 85–91 ([arXiv:1312.4545](https://arxiv.org/abs/1312.4545))

- Bono G., Marconi M., Cassisi S., Caputo F., Gieren W., Pietrzynski G., 2005, *ApJ*, **621**, 966
- Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, *MNRAS*, **427**, 127
- Cabrera-Cano J., Alfaro E. J., 1990, *A&A*, **235**, 94
- Cantat-Gaudin T., et al., 2014, *A&A*, **569**, A17
- Cantat-Gaudin T., et al., 2018, *A&A*, **618**, A93
- Cantat-Gaudin T., et al., 2020, *A&A*, **640**, A1
- Cardelli J. A., Clayton G. C., Mathis J. S., 1989, *ApJ*, **345**, 245
- Castro-Ginard A., et al., 2020, *A&A*, **635**, A45
- Dorda R., González-Fernández C., Negueruela I., 2016, *A&A*, **595**, A105
- Dutra-Ferreira L., Pasquini L., Smiljanic R., Porto de Mello G. F., Steffen M., 2016, *A&A*, **585**, A75
- Eigenbrod A., Mermilliod J. C., Clariá J. J., Andersen J., Mayor M., 2004, *A&A*, **423**, 189
- Fall S. M., Chandar R., 2012, *ApJ*, **752**, 96
- Fernie J. D., 1990, *ApJS*, **72**, 153
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, **125**, 306
- Fouqué P., et al., 2007, *A&A*, **476**, 73
- Fragkou V., Parker Q. A., Zijlstra A. A., Crause L., Barker H., 2019, *Nature Astronomy*, **3**, 851
- Gaia Collaboration et al., 2018, *A&A*, **616**, A1
- Galadí-Enriquez D., Jordi C., Trullols E., 1998, *A&A*, **337**, 125
- García-Hernández D. A., García-Lario P., Plez B., Manchado A., D'Antona F., Lub J., Habing H., 2007, *A&A*, **462**, 711
- Gray R. O., Corbally C. J., 1994, *AJ*, **107**, 742
- Gustafsson B., Edvardsson B., Eriksson K., Jørgensen U. G., Nordlund Å., Plez B., 2008, *A&A*, **486**, 951
- Hare J., Kargaltsev O., Rangelov B., 2018, *ApJ*, **865**, 33
- Hinkle K., Wallace L., Valenti J., Harmer D., 2000, Visible and Near Infrared Atlas of the Arcturus Spectrum 3727-9300 A
- Hurley J. R., Pols O. R., Aarseth S. J., Tout C. A., 2005, *MNRAS*, **363**, 293
- Kalirai J. S., Ventura P., Richer H. B., Fahlman G. G., Durrell P. R., D'Antona F., Marconi G., 2001, *AJ*, **122**, 3239
- Katz D., et al., 2019, *A&A*, **622**, A205
- Krumholz M. R., 2014, *Phys. Rep.*, **539**, 49
- Kupka F. G., Ryabchikova T. A., Piskunov N. E., Stempels H. C., Weiss W. W., 2000, *Baltic Astronomy*, **9**, 590
- Laney C. D., Caldwell J. A. R., 2007, *MNRAS*, **377**, 147
- Larsen S. S., 2009, *A&A*, **494**, 539
- Lindegren L., et al., 2020, arXiv e-prints, p. [arXiv:2012.03380](https://arxiv.org/abs/2012.03380)
- Lohr M. E., Negueruela I., Tabernero H. M., Clark J. S., Lewis F., Roche P., 2018, *MNRAS*, **478**, 3825
- Luck R. E., 2018, *AJ*, **156**, 171
- Luri X., et al., 2018, *A&A*, **616**, A9
- Maíz Apellániz J., Weiler M., 2018, *A&A*, **619**, A180
- Marco A., Negueruela I., 2011, *A&A*, **534**, A114
- Marco A., Negueruela I., González-Fernández C., Maíz Apellániz J., Dorda R., Clark J. S., 2014, *A&A*, **567**, A73
- Mazzitelli I., D'Antona F., Ventura P., 1999, *A&A*, **348**, 846
- Mermilliod J. C., Latham D. W., Glushkova E. V., Ibrahimov M. A., Batirshinova V. M., Stefanik R. P., James D. J., 2003, *A&A*, **399**, 105
- Messineo M., Davies B., Ivanov V. D., Figer D. F., Schuller F., Habing H. J., Menten K. M., Petr-Gotzens M. G., 2009, *ApJ*, **697**, 701
- Metzger M. R., Caldwell J. A. R., McCarthy J. K., Schechter P. J., 1991, *ApJS*, **76**, 803
- Negueruela I., 2014, in *Massive Young Star Clusters Near and Far: From the Milky Way to Reionization*. pp 9–15 ([arXiv:1402.3328](https://arxiv.org/abs/1402.3328))
- Negueruela I., Marco A., 2012, *AJ*, **143**, 46
- Negueruela I., Schurch M. P. E., 2007, *A&A*, **461**, 631
- Negueruela I., González-Fernández C., Marco A., Clark J. S., 2011, *A&A*, **528**, A59
- Negueruela I., Marco A., González-Fernández C., Jiménez-Esteban F., Clark J. S., Garcia M., Solano E., 2012, *A&A*, **547**, A15
- Negueruela I., Monguió M., Marco A., Tabernero H. M., González-Fernández C., Dorda R., 2018, *MNRAS*, **477**, 2976
- Negueruela I., Simón S., Dorda R., González-Fernández C., Holgado G., Maíz Apellániz J., Monguió M., Berlanas S. R., 2019, in *Highlights on Spanish Astrophysics X*. pp 425–425
- Negueruela I., Dorda R., Marco A., 2020, *MNRAS*, **494**, 3028
- Perren G. I., Vázquez R. A., Piatti A. E., 2015, *A&A*, **576**, A6
- Pfalzner S., Vincke K., Xiang M., 2015, *A&A*, **576**, A28
- Piskunov N. E., Kupka F., Ryabchikova T. A., Weiss W. W., Jeffery C. S., 1995, *A&AS*, **112**, 525
- Pont F., Mayor M., Burki G., 1994, *A&A*, **285**, 415
- Portegies Zwart S. F., McMillan S. L. W., Gieles M., 2010, *ARA&A*, **48**, 431
- Ripepi V., Molinaro R., Musella I., Marconi M., Leccia S., Eyer L., 2019, *A&A*, **625**, A14
- Rochau B., Brandner W., Stolte A., Gennaro M., Gouliermis D., Da Rio N., Dzyurkevich N., Henning T., 2010, *ApJ*, **716**, L90
- Ryabchikova T., Piskunov N., Kurucz R. L., Stempels H. C., Heiter U., Pakhomov Y., Barklem P. S., 2015, *Phys. Scr.*, **90**, 054005
- Sanders W. L., 1971, *A&A*, **14**, 226
- Santos Jr. J. F. C., Bonatto C., Bica E., 2005, *A&A*, **442**, 201
- Seleznev A. F., Carraro G., Capuzzo-Dolcetta R., Monaco L., Baume G., 2017, *MNRAS*, **467**, 2517
- Skrutskie M. F., et al., 2006, *AJ*, **131**, 1163
- Steele I. A., Negueruela I., Clark J. S., 1999, *A&AS*, **137**, 147
- Tabernero H. M., Dorda R., Negueruela I., González-Fernández C., 2018, *MNRAS*, **476**, 3106
- Tabernero H. M., Dorda R., Negueruela I., Marfil E., 2021, *A&A*, **646**, A98
- Tammann G. A., Sandage A., Reindl B., 2003, *A&A*, **404**, 423
- Wright N. J., Bouy H., Drew J. E., Sarro L. M., Bertin E., Cuillandre J.-C., Barrado D., 2016, *MNRAS*, **460**, 2593
- van Raai M. A., Lugaro M., Karakas A. I., García-Hernández D. A., Yong D., 2012, *A&A*, **540**, A44

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.