LETTER TO THE EDITOR

The Milky Way’s cluster age function in light of Gaia DR2

Friedrich Anders¹, Tristan Cantat-Gaudin¹, Irene Quadrino-Lodoso¹, Carme Jordi¹, Mark Gieles¹,², Alfred Castro-Ginard¹, Lola Balaguer-Núñez¹

¹ Institut de Ciències del Cosmos, Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, E-08028 Barcelona, Spain
² ICREA, Passeig Lluís Companys 23, E-08010 Barcelona, Spain

Received June 3, 2020; accepted ???, ???

ABSTRACT

We perform a systematic reanalysis of the age distribution of Galactic open star clusters. Using a catalogue of homogeneously determined ages for 834 open clusters contained in a 2 kpc cylinder around the Sun and characterised with astrometric and photometric data from the Gaia satellite, we find that it is necessary to revise earlier works that relied on data from the Milky Way Star Cluster Survey. After establishing age-dependent completeness limits for our sample, we find that the cluster age function in the range 6.5 < log t < 10 is compatible with a Schechter-type or broken power-law function, whose parameters we determine by MCMC fitting. Our best-fit values indicate an earlier drop of the age function (by a factor of 2 – 3) with respect to the results obtained in the last five years, and are instead more compatible with results obtained in the early 2000s and radio observations of inner-disc clusters. Furthermore, we find excellent agreement with the dynamical cluster formation and destruction models of Lamers et al. (2005), indicating a typical destruction time-scale of ~ 1.5 Gyr for a 10⁶ M⊙ cluster and a present-day cluster-formation rate of ~ 0.6±0.1 Myr⁻¹ kpc⁻², suggesting that only 8 – 15 % of all stars born in the solar neighbourhood form in bound clusters. Accurate cluster-mass measurements are now needed to place more precise constraints on open-cluster formation and evolution models.

Key words. Galaxy: open clusters, Galaxy: evolution, Galaxy: solar neighbourhood, methods: data analysis, statistical

1. Introduction

It is becoming increasingly difficult to understand the formation of galaxies without taking into account several levels of baryonic hierarchical structure formation. To unravel the formation history of the Milky Way disc, however, it is often useful to study open star clusters (OCs): groups of stars of the same age and abundance pattern, held together by mutual gravitation.

The physical processes governing the formation and evolution of OCs are encoded in the distribution of their properties: mass, age, size, etc. (for a recent review, see Krumholz et al. 2019). Since it is relatively easy to estimate at least differential ages for OCs, one of the key observables of the local OC population is the completeness-corrected age function (e.g. Wielen 1971; Janes et al. 1988; Battinelli & Capuzzo-Dolcetta 1991; Lamers et al. 2005; Piskunov et al. 2006; Morales et al. 2013; Piskunov et al. 2018; Krumholz et al. 2019). This cluster age function (CAF) can be thought of as an integral of the cluster distribution function over several other parameters that are much more difficult to determine (such as present-day mass, initial mass, internal rotation, binary fraction, etc.).

In the Milky Way, the census of OCs is highly incomplete, at least beyond a local volume of ~ 1 – 2 kpc (Kharchenko et al. 2013). Thanks to the unprecedented quality of the astrometric and photometric data released with Gaia (DR2: Gaia Collaboration et al. 2018), hundreds of new clusters have recently been detected even at smaller distances (e.g. Cantat-Gaudin et al. 2018, 2019; Castro-Ginard et al. 2019, 2020; Liu & Pang 2019; Sim et al. 2019). In addition, some analyses have shown that previous catalogues also contained large numbers of false positives and asterisms (Cantat-Gaudin et al. 2018; Cantat-Gaudin & Anders 2020). The impact of Gaia on the field of Galactic cluster studies can thus hardly be overstated.

The main remaining challenges for obtaining a clean and unbiased CAF for the Milky Way (or at least for the local solar neighbourhood of a few kpc) are a) the irregular dust distribution in the Galactic disc; b) the intrinsically patchy distribution of star clusters and other young disc tracers (Becker 1963; Becker & Fenkart 1970; Efremov 2010; Moitinho 2010; Cantat-Gaudin et al. 2018; Reid et al. 2019; Skowron et al. 2019), rendering completeness estimates difficult; c) the smooth transition between moving groups, associations, and physically bound OCs (Krumholz et al. 2019; Kounkel & Covey 2019; Cantat-Gaudin & Anders 2020; Kounkel et al. 2020); and d) the availability of homogeneously derived cluster parameters.

The last problem has recently been addressed by Cantat-Gaudin et al. (2020, hereafter CGA20) who published a catalogue of homogeneous age estimates for 1,867 Galactic OCs confirmed by Gaia DR2. In this Letter we use this catalogue to reevaluate the Milky Way’s CAF. Our analysis is reproducible via the python code provided here: https://github.com/fjaellet/gaidr2-caf.

2. The Gaia DR2 open-cluster census

The precise Gaia DR2 astrometry (positions, proper motions and parallaxes) allows for detections of OC members (including their tidal tails; Röser et al. 2019; Röser & Schilbach 2019) and the discovery of thousands of new clusters and moving groups almost entirely from the proper-motion measurements (e.g. Gaia Collaboration et al. 2018; Cantat-Gaudin et al. 2018; Kounkel & Covey 2019; Meingast et al. 2019). The Gaia photometry (Evans...
A proofs: manuscript no. clusteragefunction_letter

Fig. 1. Galactic distribution of the OC samples studied in this Letter, sliced into logarithmic age bins. Left: pre-Gaia census using the MWSC catalogue (Kharchenko et al. 2013). Right: Post-Gaia DR2 census, using the catalogue of Cantat-Gaudin et al. (2020). In each panel we show a 2D kernel density estimate with a fixed bandwidth of 0.05 kpc. For the MWSC, the dashed circle corresponds to the completeness limit of 1.8 kpc used in the literature (e.g. Piskunov et al. 2018), while for the Gaia DR2 census we use a sample limit of 2 kpc, together with age-dependent completeness fractions indicated in each panel (see Sect. 2, last paragraph).

et al. 2018) allows to characterise these objects in detail through their colour-magnitude diagrams.

Here we use the homogeneously derived parameters for 1,867 Gaia-detected clusters recently published by CGa20. For that catalogue, the main cluster parameters age, distance modulus, and extinction were computed from the observed Gaia DR2 parallaxes and $G$ vs. $(G_{BP} - G_{RP})$ colour-magnitude diagrams by a multi-layer-perceptron neural network trained on a set of 347 OCs with well-determined parameters (mostly taken from Bossini et al. 2019). The cluster membership lists were mostly taken from Cantat-Gaudin & Anders (2020) and Castro-Ginard et al. (2020). The typical log $t$ uncertainties derived by the neural network amount to 0.15-0.25 for clusters younger than 300 Myr, and 0.1-0.15 for clusters older than that. For details of the method, we refer to CGa20.

Recent studies of the CAF (Joshi et al. 2016; Piskunov et al. 2018; Krumholz et al. 2019) have relied on the cluster data compiled in the latest version of the Milky Way Star Cluster survey (Kharchenko et al. 2016). A substantial fraction of objects contained in this catalogue (among them all the putatively old, high-latitude inner-galaxy OC candidates, as well as many dubious NGC objects and about 50% of the old nearby FSR cluster candidates of Froebrich et al. 2007), however, could not be confirmed with Gaia DR2 (Cantat-Gaudin et al. 2018; Cantat-Gaudin & Anders 2020). In the following analysis, we therefore compare both to the Kharchenko et al. (2013) version of the MWSC as well as to its latest version.

To illustrate the transformative power of Gaia DR2 on the field, Fig. 1 compares the distribution of OCs in heliocentric Cartesian coordinates derived from the MWSC catalogue with the distribution obtained from the new catalogue of Gaia-detected OCs of CGa20. For a deeper discussion of the structures emerging from this figure, we refer to the latter paper. Here our main objective is to estimate the (age-dependent) completeness of the new catalogue in order to determine the CAF.

In order to correct for our incomplete view of the Galactic OC population, we need to quantify how selection biases affect our samples. The different aspect of the OC distributions in the right column of Fig. 1 already suggests that the present Gaia DR2 census is unlikely to be complete to a fixed limit, as was frequently assumed for the MWSC catalogue (the dashed grey circle in the left-column panels denotes the 1.8 kpc completeness limit used by Kharchenko et al. 2016; Joshi et al. 2016, and Piskunov et al. 2018).

In this work, we estimate the age-dependent completeness of the Gaia DR2 cluster census within a cylinder of radius $d_{\text{cy}} = 2$ kpc (right column of Fig. 1). The analysis can be retraced in an accompanying jupyter notebook. In a nutshell, we take into account two effects: 1. Undetected clusters: we use the OC recovery experiment performed for the latest Galactic-plane OC search of Castro-Ginard et al. (2020) to estimate the detection efficiency of their conservative method as a function of distance, sky region, and age. 2. Uncharacterised clusters: Not for all Gaia-detected OCs it was possible to infer physical parameters in CGa20. Within the 2 kpc cylinder, however, this effect is minor: only 32 non-characterised clusters have Bayesian parallax distances smaller than 2 kpc. Estimating their age distribution using the values of Kharchenko et al. (2013), we find that they are mostly younger than log $t = 7.5$. The combined completeness fractions for each age bin are given in Fig. 1.

https://github.com/fjaellet/gaidr2-caf/blob/master/gaia_dr2_clusters_completeness.ipynb
3. The post-\textit{Gaia} DR2 cluster age function

Having established the completeness limits of the \textit{Gaia} DR2 cluster sample, we can now determine the age distribution and the CAF. The top panel of Fig. 2 shows the histogram and a kernel-density estimate of the logarithmic age distribution for the \textit{Gaia} DR2 census within a 2 kpc cylinder around the Sun. For comparison we also show the results obtained with the original MWSC catalogue (Kharchenko et al. 2013) and the latest version used by Piskunov et al. (2018) and Krumholz et al. (2019). From Fig. 2 we can already appreciate some important differences with respect to these pre-\textit{Gaia} works: the peak of the distribution lies around \( \log t \sim 8.2 \), and despite the fact that the \textit{Gaia} census is much more complete for old OCs, we see a lot less of those objects.

The typical metric for the cluster age distribution, used both by the Galactic and the extragalactic community, is the cluster age function (CAF); the number of clusters per unit of age, in logarithmic age bins. Following the method of Piskunov et al. (2018), we derived the CAF for the MWSC and the \textit{Gaia} samples. Our results are shown in the middle panel of Fig. 2. In this panel, we also show some of the literature results compiled by Piskunov et al. (2018), namely Wielen (1971); Pandey & Mahra (1986); Lamers et al. (2005); Piskunov et al. (2006), and Morales et al. (2013).

In the bottom panel of Fig. 2, we compare our data to a number of models. In particular, these are two of the cluster destruction models presented by Lamers et al. (2005), a fit to the Lamers & Gieles (2006) model, and the results of our fits to two simple analytical functions, all performed with the Markov-chain Monte-Carlo sampler \texttt{emcee} (Foreman-Mackey et al. 2013).

We confirm the conclusion of Krumholz et al. (2019) that the Milky Way CAF is well fitted by a Schechter function or, slightly worse, by a broken power law. The fit parameters for those functions, however, have to be revised. In particular, we obtain best-fit values of \( \alpha_1 = -0.66^{+0.10}_{-0.11}, \log t^\star = 9.30^{+0.07}_{-0.05} \) for the Schechter case (Krumholz et al. 2019; \( \alpha_T = -0.55, \log t^\star = 9.59 \)), and \( \alpha_1 = -0.60^{+0.16}_{-0.11}, \alpha_2 = -2.30^{+0.30}_{-0.31}, \log t^\text{break} = 8.53^{+0.25}_{-0.14} \) for the case of a broken power law (Krumholz et al. 2019; \( \alpha_1 = -0.61, \alpha_2 = -1.67, \log t^\text{break} = 8.89 \)). Our basic conclusion is that the downturn in the CAF occurs at lower ages (by a factor 2 – 3) and the slope beyond the break is steeper. Single power-law models such as the \( \propto t^{-1} \) model inspired by observations of the Antennae galaxies (Whitmore et al. 2007) do not provide a good approximation for the Milky Way.

![Fig. 2. Age distribution for Galactic open clusters in the solar vicinity. Top panel: normalised histograms and kernel-density estimates. The cyan and blue distributions show the results from the MWSC survey (Kharchenko et al. 2013 and Piskunov et al. 2018, respectively); the red distribution shows our \textit{Gaia} DR2-derived results. Middle and bottom panels: Observational CAF determinations for the extended solar neighbourhood, from Wielen (1971) to our completeness-corrected \textit{Gaia} DR2-based census (red). Errorbars correspond to Poissonian uncertainties in the number of clusters per bin. Bottom panel: CAF comparison to models, as indicated in the legend.](image-url)
This also implies that some cluster formation and destruction models from the pre-MWSC era are still compatible with our measurements. In particular, this is the case for the model of Lamers et al. (2005) with a typical destruction time-scale for a $10^5 M_\odot$ OC of $t_\text{d} \sim 1.5$ Gyr, which we show in Fig. 2 (dashed black line). Those authors modelled the cluster destruction as $t_\text{d} \propto (M/10^5 M_\odot)^{\gamma}$, with $\gamma \approx 0.62$ and a star-formation rate in clusters of $\sim 300 M_\odot$ Myr$^{-1}$ kpc$^{-2}$. Surprisingly, almost all of our CAF data points (except for the lowest age bin) are consistent with the Lamers et al. (2005) model within 1$\sigma$. In addition, we find - in accordance with Morales et al. (2013) - a hint of a short bump in the cluster-formation rate at very young ages, around 6–20 Myr (dash-dotted curve in Fig. 2). The proximity of our data to the CAF obtained by Morales et al. (2013) from ATLASGAL radio data (Schuller et al. 2009) of mostly embedded clusters towards the inner Galaxy also suggests little change in the cluster destruction rate within a few kpc from the Sun.

Lamers & Gieles (2006) parameterised the destruction time of initially bound OCs in the solar neighbourhood, taking into account four processes in the life of OCs: stellar evolution, tidal destruction, cluster destruction, and parallaxes, thus making our measurement a new benchmark for extragalactic studies as well.

References

Adamo, A., Zeidler, P., Kruisjes, J. M. D., et al. 2020, Science Reviews, accepted, arXiv:2005.06188
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Battinelli, P. & Capuzzo-Dolcetta, R. 1991, MNras, 249, 76
Becker, W. 1963, ZAp, 57, 117
Borisso, D. & Kroupa, P. 2001, in IAU Symposium, Vol. 193, The Structure of our Galaxy, ed. W. Becker & G. I. Kontopoulos, 205
Bossini, D., Vallarini, A., Bragaglia, A., et al. 2019, A&A, 623, A108
Cantat-Gaudin, T. & Anderegg, M. 2020, A&A, 633, A99
Cantat-Gaudin, T., Anders, F., Castro-Ginard, A., et al. 2020, A&A, accepted, arXiv:2004.07274
Cantat-Gaudin, T., Jordi, C., Vallarini, A., et al. 2018, A&A, 618, A93
Castro-Ginard, A., Krone-Martins, A., Sedaghat, N., et al. 2019, A&A, 624, A126
Castro-Ginard, A., Jordi, C., Luri, X., et al. 2020, A&A, 635, A45
Castro-Ginard, A., Jordi, C., Luri, X., Cantat-Gaudin, T., & Balaguer-Núñez, L. 2019, A&A, 627, A35
Ermefors, Y. N. 2010, MNras, 405, 1531
Evans, D. W., Riebel, M., DeAngelis, F., et al. 2018, A&A, 616, A4
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Frisch, D., Scholz, A., & Raftery, C. I. 2007, MNras, 374, 399
Gaia Collaboration, Brown, A. G. A., Vallarini, A., et al. 2018, A&A, 616, A1
Hunter, J. D. 2007, Computing in Science and Engineering, 9, 90
James, K. A., Tilley, C., & Lynga, G. 1988, AJ, 95, 771
Joshchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., & Scholz, R. D. 2013, A&A, 558, A53
Joshchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., & Scholz, R. D. 2016, A&A, 585, A101
Kounkel, M. & Covey, K. 2019, AJ, 158, 122
Kounkel, M., Covey, K., & Stassun, K. G. 2020, AJ, submitted, arXiv:2004.07261
Krumholz, M. R., McKee, C. F., & Bland-Hawthorn, J. 2019, ARA&A, 57, 227
Lamers, H. J. G. L. M. & Gieles, M. 2006, A&A, 455, L17
Lamers, H. J. G. L. M., Gieles, M., Bastian, N., et al. 2005, A&A, 441, 117
Liu, L. & Fang, X. 2019, ApJS, 245, 3
Meingast, S., Alves, J., & Pflamm-Altenburg, J. 2019, A&A, 622, L13
Mottin, R. 2010, in IAU Symposium, Vol. 266, Star Clusters: Basic Galactic Building Blocks Throughout Time and Space, ed. R. de Grijs & J. R. D. Lépine, 106–116
Morales, E. F. E., Wyrowski, F., Schuller, F., & Menten, K. M. 2013, A&A, 560, A76
Piskunov, A. E., & Mahra, H. S. 1986, Ap&SS, 126, 167
Piskunov, A. E., Just, A., Karchenko, N. V., et al. 2018, A&A, 614, A22
Piskunov, A. E., Karchenko, N. V., Röser, S., Schilbach, E., & Scholz, R. D. 2016, A&A, 593, A45
Reid, M. J., Menten, K. M., Brünninger, A., et al., 2019, ApJ, 885, 131
Röser, S. & Schilbach, E. 2019, A&A, 627, A4
Röser, S., Schilbach, E., & Goldman, B. 2019, A&A, 621, L2
Schuller, F., Menten, K. M., Contreras, Y., et al. 2009, A&A, 504, 415
Sim, G., Lee, S. H., Ann, H. B., & Kim, S. 2019, Journal of Korean Astronomical Society, 52, 145
Skowron, D. M., Skowron, J., Mróz, P., et al. 2019, Science, 365, 478
Taylor, M. B. 2005, Astronomical Society of the Pacific Conference Series, Vol. 346, PC2005 & STSC2, Mariinko Link/TableVOTable Processing Software, ed. P. Shopbell, M. Britton, & R. Ebert, 29
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, Computing in Science and Engineering, 13, 22
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2019, arXiv e-prints, arXiv:1907.10121
Ward, J. L., Kruisjes, J. M. D., & Rix, H.-W. 2020, MNRAS, 495, 663
Whitmore, B. C., Chandar, R., & Fall, S. M. 2007, AJ, 133, 1067
Wielen, R. 1971, A&A, 13, 309
Acknowledgements. We thank Mercè Romero-Gómez and Cesca Figuera for their comments. This work has made use of data from the European Space Agency (ESA) mission Gaia (www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC, 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 work was supported by the MINECO (Spanish Ministry of Economy) through grant RTI2018-095076-B-C21 (MINECO/FEDER, UE). MCG acknowledges support from Juan de la Ciera - Formación 2015 grant, MINECO (FEDER/UE). GA is grateful for funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 800502. The preparation of this work has made use of TOPCAT (Taylor 2005), NASA’s Astro- physics Data System Bibliographic Services, as well as the open-source Python packages Astropy (Collaboration et al. 2018), NumPy (van der Walt et al. 2011), scikit-learn (Virtanen et al. 2019), and en-core (Foreman-Mackey et al. 2013). The figures in this paper were produced with Matplotlib (Hunter 2007).