

The birth of the Milky Way as uncovered by accurate stellar ages with Gaia

Carme Gallart^{1,2,*}, Edouard J. Bernard³, Chris B. Brook^{1,2}, Tomás Ruiz-Lara^{1,2}, Santi Cassisi^{4,5}, Vanessa Hill³, and Matteo Monelli^{1,2}

¹Instituto de Astrofísica de Canarias, E-38200 La Laguna, Tenerife, Spain

²Departamento de Astrofísica, Universidad de La Laguna, E-38205 La Laguna, Tenerife, Spain

³Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, France

⁴INAF – Astronomical Observatory of Abruzzo, via M. Maggini, sn, 64100 Teramo, Italy

⁵INFN, Sezione di Pisa, Largo Pontecorvo 3, 56127 Pisa, Italy

⁶Laboratoire Lagrange, Université de Nice Sophia-Antipolis, Observatoire de la Côte d'Azur, CNRS, Bd de l'Observatoire, CS 34229, F-06304 Nice cedex 4, France

*carme@iac.es

Knowledge of ages for stars formed over a galaxy's lifetime is fundamental to understand its formation and evolution. However, stellar ages are difficult to obtain since they cannot be measured from observations, being comparison with stellar models(1) required. Alternatively, age distributions can be derived applying the robust technique of colour-magnitude diagram fitting(2), till now mainly employed to study nearby galaxies. The new distances to individual Milky Way stars from the Gaia mission(3) have allowed us to use this technique to derive ages from a thick disk colour-magnitude diagram, and from the enigmatic, two-sequenced colour-magnitude diagram of the kinematically hot local halo(4), which blue-sequence has been linked to a major accretion event(5; 6). Because accurate ages were lacking, the time of the merger and its role on our Galaxy's early evolution remained unclear. We show that the stars in both halo sequences share identical age distributions, and are older than the bulk of thick disc stars. The sharp halo age cut 10 Gyr ago can be identified with the accretion of Gaia-Enceladus. Along with state-of-the-art cosmological simulations of galaxy formation(7), these robust ages allow us to order the early sequence of events that shaped our Galaxy, identifying the red-sequence as the first stars formed within the Milky Way progenitor which, because of their kinematics, can be described as its long sought in-situ halo.

The new accurate parallaxes and luminosities provided in the second data release (DR2) of the Gaia mission(3) have allowed us to construct, for the first time, colour-magnitude diagrams (CMD) in the *absolute* plane for stars located in a large volume of the Milky Way, encompassing different Galactic structural components. These CMDs, in units of absolute magnitudes and colours, are what is required to derive star formation histories(2) and stellar age distributions, by comparing them with theoretical CMDs derived from stellar evolution models(8).

The top panels of Figure 1 show the CMD of two sub-populations of Milky Way stars taken from a parent population that lies within a sphere of $\simeq 2$ Kpc around the Sun, as observed by Gaia. In this volume, accurate distances and absolute magnitudes can be derived directly from parallaxes. The CMD in the top left panel contains about sixty thousand stars from this spherical region with large tangential velocities relative to the Sun (greater than 200 km/s). Stars with such high velocities are classified in this study as belonging to a kinematically defined stellar halo(4). The CMD in the top right panel is of some

half million stars from the same spherical region but selected to be at least 1.1 Kpc above or below the Galactic plane. At this distance from the plane, the majority of stars are expected to belong to the thick disc(9), rather than to the young thin disc component. We have excluded from this sample the stars with high velocity (greater than 200 km/s) that have been included in the halo CMD. Note that our thick disk definition is morphological rather than chemical or kinematical: some stars with thin disk kinematics (4) would be included in our sample. For simplicity, in this paper we will refer to the two samples we have just defined as *halo* and *thick disk* (see Methods for further information on the sample selection criteria). As is readily apparent, the CMD of the kinematically hot halo population contains two distinct sequences: a "blue" sequence on the left, and a "red" sequence on the right. This peculiar morphology indicates the presence of two distinct sub-populations within the halo. The blue sequence has been associated to a very significant merging event, Gaia-Enceladus, occurred early during the Milky Way formation(6), and of which the first hints date back to at least 15 years ago(10; 11; 12; 13). The nature of the red sequence has been less clear. It has been associated with the Milky Way thick disc (5), with support for this association coming from the characteristics of its chemical composition (12).

We have compared the distribution of stars in these observed CMDs with that in CMDs computed with state-of-the-art stellar evolution models (8) and found the combination of ages and chemical composition of the stars (or metallicities, that is, amount of chemical elements other than Hydrogen and Helium) that best reproduce the observed CMDs (see Methods). These best-fit CMDs are displayed in the bottom two panels of Figure 1. The model CMDs clearly succeed at reproducing the various features observed in the empirical CMDs.

Figure 2 displays the distribution of ages and metallicities of these best-fit model CMDs, divided in the blue and red sequences in the case of the halo CMD. The age distributions clearly demonstrate that *the two sequences in the halo CMD are composed by stars that are coeval and formed at the earliest possible times in the life of the Universe* (peak age 13.4 Gyr, 50% of stars formed by 12.3 Gyr ago). The difference in colour is a result of the red and blue sequence stars having different metallicities. The models also clearly show that the thick disc population spans a wider age range than the two halo populations: its stars closely follow the age distribution of the halo for the first $\simeq 2$ Gyr, but the thick disk keeps forming stars after this initial period and until $\simeq 6$ Gyr ago, with its age distribution reaching a second peak later on ($\simeq 9.5$ Gyr ago, with 50% of stars formed by 10.5 Gyr ago).

These age results contrast with literature values for the ages (12; 14; 15) assigned to stars associated with these components on which previous conclusions on their nature relied(6). These former age determinations indicated that the two halo populations had different ages, with blue-sequence stars about 1-2 Gyr younger than stars in the red-sequence. However, they were based either on i) a relatively small number of stars, and used spectroscopically determined stellar parameters which allow less robust age estimates than absolute luminosities and colours or ii) indirect age estimates based on low resolution spectroscopy. They contrasted with ages inferred with simple isochrone fitting taking into account average metallicities for stars in the two sequences(4; 6). Our robust age distributions for the thick disk and halo components provide key insights into the first events that shaped the Milky Way structure, providing crucial new evidence that complements previous studies which have largely dealt with chemical and kinematic information (5; 6; 16; 17; 18; 19; 20; 21).

The fact that the two halo populations are coeval means not only that the red and blue populations formed stars at similar times, but that they also stopped forming at similar times. Additionally, the well established relation (22) between a galaxy's mass and the amount of metals (defined as chemical elements other than Hydrogen and Helium) contained in their stars means that the stars in the red sequence of the halo, being more metal rich, must have formed in a galaxy that was more massive than the one where the stars in the blue sequence were formed. Both these findings are precisely what is expected if these

two populations were involved in a merger event, with the red sequence stars belonging to the main progenitor of the Milky Way, and the blue sequence belonging to a smaller accreted galaxy, the one dubbed Gaia-Enceladus. This is shown in Figure 3. The left panel shows the age-metallicity distribution derived from the observed CMDs, with the red and blue sequence halo stars indicated as red and blue contours. The right panel shows the signature in the age-metallicity plane left by a merger event in a simulated Milky-Way analogue galaxy (see Methods), where the main progenitor stars are shown as red contours and those from the accreted galaxy as blue contours. The merger can be identified as two tracks in this plot, with the higher metallicity stars belonging to the main progenitor and the lower metallicity stars coming from the less massive merged galaxy.

The difference in the amount of metals in the red and blue sequences indicates that the accreted Gaia-Enceladus had about 30% (23) of the mass of stars in the main progenitor, although we stress that this ratio remains quite uncertain. This would indicate a total mass ratio of around 4:1 between the two galaxies, given the relation between stellar mass and total mass (24). This estimated mass ratio is similar to that derived in previous works (6) using different arguments. Regardless of the exact mass ratio, this encounter heated some of the main progenitor stars that had been forming in a disc-like structure, to the extreme kinematics that lead them to be classified as halo stars. The existence of such halo stars has been predicted by cosmological simulations of Milky Way-type galaxies (17; 25) and has also been postulated from observational evidence (5; 11; 12). *Our age determination discloses that the stars heated to halo-like kinematics were among the first formed in the Milky Way, during the first $\simeq 3$ Gyr of its evolution, right before the merger with Gaia-Enceladus took place. We can date the merger as occurring about 10 Gyr ago.*

The similar star formation timescales for the two halo sequences also has important implications for the origin of their detailed chemical enrichment. It is known (4; 5; 6; 12) that for stars with similar amounts of Fe, the blue sequence stars (Gaia-Enceladus) have significantly less α -elements than the red sequence stars (formed in the main progenitor). The run of $[\alpha/\text{Fe}]$ vs. $[\text{Fe}]$ in different stellar systems is commonly used to trace the star formation timescale of a galaxy, since α elements are produced in Supernova Type II on shorter timescales than iron, which is mostly produced on a longer timescale by Supernovae Type Ia. The fact that we find similar star formation timescales for the two components indicates that other mechanisms should be at play. In our simulations, the progenitor and satellite have different star formation histories leading up to the merger (19), with the progenitor star formation rate increasing whilst the satellite is small enough to have a basically constant star formation rate, regulated by feedback.

A further interesting result is the relation between the red sequence, or 'in-situ' halo, and the thick disc. The CMDs of these two components overlap, as seen in Figure 1, as do their ages, chemical abundances and kinematics as seen in Figures 2 and 4 as well as in previous studies(4; 5). The implication is that the in-situ halo and thick disc have the same origin, linked to the main progenitor of the present day Milky Way: in this self consistent scenario, the old age tail of the thick disc population shown in Figure 2 is populated by stars formed in the main progenitor, which were not heated to halo-like velocities. We may speculate that the metal weak thick disk(26) would be part of this population. Therefore, stars that would be currently associated to the thick disc were forming in the main progenitor before, during and after the merger with Gaia-Enceladus, whilst a fraction of main progenitor stars formed very early, and were heated to halo kinematics by the accretion event. The infalling gas associated to the merger likely contributed to maintain and even stoke star formation in the early disk, which stellar age distribution shows a second peak of high star formation rate $\simeq 9.5$ Gyr ago. Such a formation scenario of the thick disc has been foreshadowed by cosmological galaxy formation simulations(27): the thick disc forms at high redshift during a period characterised by gas rich mergers.

A final implication of our findings regards the transition (28; 29) between thick and thin disc formation. The derived ages for thick disc stars indicate that this transition occurred around 8-6 Gyr ago, in agreement

with previous studies (30), and that this transition was not associated with the earlier Gaia-Enceladus merger.

In conclusion, the reported new stellar age distributions, aided by state-of-the-art cosmological simulations of disc galaxy formation, delineate a clear picture of the formation of our Galaxy. In this picture, a primitive Milky Way had been forming stars during some 3 Gyr when a smaller galaxy, which had been forming stars on a similar timescale but was less chemically enriched owing to its lower mass, was accreted to it. This merger heated part of the existing stars in the main progenitor to a stellar halo-like configuration. A ready supply of infalling gas during the merger ensures the maintenance of a disc-like configuration, with the thick disc continuing to form stars at a substantial rate. Subsequently, around 8-6 Gyr ago, the gas settled into a thin disc that has continued to form stars until the present day.

Correspondence and requests for materials should be addressed to CG (carme.gallart@iac.es).

Acknowledgements CG, TRL and MM acknowledge support by the Spanish Ministry of Economy and Competitiveness (MINECO) under the grants AYA2014-56795-P and AYA2017-89076-P as well as AYA2016-77237-C3-1-P (TRL). SC acknowledges support from Premiale INAF "MITIC" and has been supported by INFN (Iniziativa specifica TAsP). We used data from the European Space Agency mission Gaia (<http://www.cosmos.esa.int/gaia>), processed by the Gaia Data Processing and Analysis Consortium (DPAC; see <http://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement. We also used data from the LAMOST and GALAH surveys. Guoshoujing Telescope (the Large Sky Area Multi-Object Fiber Spectroscopic Telescope LAMOST) is a National Major Scientific Project built by the Chinese Academy of Sciences. Funding for the project has been provided by the National Development and Reform Commission. LAMOST is operated and managed by the National Astronomical Observatories, Chinese Academy of Sciences. The GALAH survey is based on observations made at the Australian Astronomical Observatory, under programmes A/2013B/13, A/2014A/25, A/2015A/19, A/2017A/18. We acknowledge the traditional owners of the land on which the AAT stands, the Gamilaraay people, and pay our respects to elders past and present.

Author contributions All authors have critically contributed to different aspects of the data analysis and model calculation, and to the interpretation of the results. The writing of the manuscript to which all authors contributed, was led by CG and CBB. Regarding specific aspects, CG selected the Gaia samples and performed the CMD fitting. EJB helped with the data selection, wrote the CMD fitting software (*TheStorm*) based on earlier work with MM and CG, and performed the calculation of the 3D interstellar reddening. CBB contributed the galaxy formation models which were key in the interpretation of the results. TRL participated in the CMD fitting, contributed key software for various steps such as the error simulation in the synthetic CMDs, and created the figures. SC contributed the software to calculate the synthetic CMD, including all the necessary libraries of stellar models and bolometric correction tables for the Gaia photometric passbands. VH selected the spectroscopic samples, which were analyzed together with TRL.

Competing interests The authors declare no competing financial interests

Correspondence and requests for materials should be addressed to CG (carme.gallart@iac.es).

Methods

1 Data selection

A detailed and reliable determination of the age and metallicity distributions of a stellar system can be obtained from a CMD that is essentially complete down to the oldest main sequence turn-offs (2). For a relatively metal-rich stellar population such as that of the Milky Way, the old turnoff occurs at an absolute magnitude $M_G=4$ in the Gaia G band, but it is necessary to reach about one magnitude deeper to accurately sample this critical point in the CMD. At a distance of 2 Kpc from the Sun, $M_G=5$ corresponds to an apparent magnitude $G=16.5$, down to which Gaia DR2 is essentially complete.

From Gaia DR2(3) we selected all the stars inside a sphere of radius $\simeq 2$ Kpc centered in the Sun down to $M_{G,R}=7$ (where $M_{G,R}$ stands for M_G before applying any correction for interstellar extinction. $M_{G,R}$ was calculated directly from the apparent magnitude using the Gaia parallaxes, as $M_{G,R}=G+5+5 \times \log_{10}(\omega/1000.0)$, with ω being the parallax in milliarcseconds (mas). A parallax zero-point offset of 0.054 (34) mas (in the sense that the corrected parallaxes are larger), has been applied. This value has been preferred to the Gaia DR2 recommended value of 0.03 mas (35) since several subsequent works have found a larger value (36; 37; 38; 39; 40; 41). The direct calculation of distance from the parallax value is valid because the great majority (over 95%) of stars in our sample have small relative parallax error ($\text{parallax_over_error} > 5$ (42)). The stars excluded by this cut are more scattered in the CMD than the thick disk and halo sequences. Even if they would correspond to actual stars in these populations, their numbers are so low that the CMD fit or the derived age distribution would not be affected. Finally, following (4; 35), it was also required that $\text{astrometric_chi2_al}/(\text{astrometric_n_good_obs_al} - 5) < 1.44 \times \max(1, \exp(-0.4 \times (\text{phot_g_mean_mag} - 19.5)))$ (only 2% of stars were removed by this cut) and that the stars had $(G_{BP}-G_{RP})$ measured colour.

From this parent population, we selected as thick disc sample all the stars in the North and South caps of the sphere, with distances from the Galactic plane, $|Z| > 1.1$ Kpc, and tangential velocities, calculated as $V_T=(4.74/\omega) \times \sqrt{p_{mra}^2 + p_{mdec}^2}$, lower than 200 km/s. For the halo sample, we selected all the relatively high Galactic latitude stars ($b > 30$) with $V_T > 200$ km/s. The latitude cut was performed to avoid substantial contamination from the thin disc at low galactic latitude. Note that, since distances are actually obtained directly from parallax inversion, as $d(\text{pc})=1000./\text{parallax}$, a small bias is introduced, such that true distances would tend to be smaller (43) by $\simeq 100$ pc on average. This bias affects very mildly the exact spatial shape and size of the actual volume covered, but does not change the fact that we are considering a thick disk sample (defined morphologically) and a halo sample (defined kinematically), and thus does not affect the conclusions of the paper.

Corrections for interstellar extinction were calculated on a star by star basis by interpolating the colour excess from a 3-D map (44). This map is strictly valid only up to Heliocentric distances of 2.6 Kpc and $|Z| < 0.6$ Kpc, and thus, our thick disk sample and part of the halo sample are outside its validity range. However, most of the dust is confined in a layer thinner than 0.6 Kpc and consequently, the extinction to the thick disk stars integrated up to this height is a very good approximation of the total extinction. We used published extinction coefficients (45), together with the effective temperature of each star (we neglected the dependence on metallicity). The temperature was estimated using a correlation between $G_{BP}-G_{RP}$ colour and temperature T_{eff} determined from a large sample of Gaia DR2 stars with measured T_{eff} . Relatively small extinction values are inferred for the bulk of stars in both samples: 96% [93%] of the stars have $A_G < 0.2$ and 79% [70%] have $A_G < 0.1$ for the thick disk [halo] sample.

After keeping only stars brighter than $M_G=5$, the thick disc sample contains 518,968 stars and the halo sample contains 61,427 stars. The CMD of these stellar samples are displayed in the top panels of Figure 1.

2 Modelling of the CMD

In order to derive age and metallicity distributions of each population, the observed CMDs have been modelled through the comparison with synthetic CMDs following well established techniques (46; 47; 48; 49; 50; 51; 52; 53; 54) that have been used for over 20 years to determine detailed star formation histories for Local Group galaxies (2; 55). These galaxies are sufficiently close to observe individual stars, yet far enough away that all their stars can be considered to be at the same distance, which can be obtained accurately using various well calibrated distance indicators (56; 57).

However, applying this technique to populations of Milky Way stars had only been possible within an extremely small volume around the Sun (58; 59; 60; 61) owing to the difficulty of obtaining accurate distances for large stellar samples. This is one of the many aspects in which the Gaia mission is providing a breakthrough: in its second data release (DR2)(3), Gaia provided distances for 1.3 billion stars in the Milky Way. This distance information has allowed us, for the first time, to apply the CMD-fitting technique to the different morphological components of our own Galaxy and therefore to robustly derive their age distributions. This provides an unprecedented window to look at the early sequence of events that shaped the formation and structure of the Milky Way.

In the process of fitting an observed CMD, the combination of simple stellar populations (that is, synthetic populations with small ranges of ages, $\Delta T \leq 1$ Gyr, and metallicities, $\Delta [\text{Fe}/\text{H}] \sim 0.1$ dex) which provide the best fit is obtained. The minimization has been performed with the code *TheStorm* (62; 63). This code uses a *mother* synthetic CMD from which the simple stellar populations are extracted. In order to check for possible systematic effects derived from the chemical properties of the stellar evolution models, these synthetic CMDs have been calculated using the BaSTI stellar evolution library in both its solar-scaled (64) and α -enhanced heavy element distribution versions (8). For the comparison itself, we make use of most of the area within the CMD populated by stars, excluding the Horizontal Branch (HB) which is still difficult to reproduce quantitatively in stellar evolution models. Although this portion of the CMD is not included in the fit, we tried to obtain a fine match of this evolutionary stage in the observed CMDs of both the halo -where it is quite prominent- and the thick disc -where the blue extension of the HB is basically missing. Since, for any given metallicity, the HB morphology is mainly driven by the mass loss efficiency during the red-giant branch stage, we used stellar model sets accounting for different efficiency of this process appropriate for each component: we used $\eta=0.2$ for simulating the thick disc CMD and $\eta=0.4$ for the halo CMD, where η is the free parameter used in combination with the Reimers' mass loss law in stellar modelling (8). Although different η values have been chosen for the different components under analysis, we should state that the results presented in this paper are not contingent upon this particular choice. Common to all *mother* synthetic CMDs (with 10^8 stars) are a constant star formation rate with flat age and metallicity distributions (14 Gyr to 30 Myr, and $-2.6 < [\text{Fe}/\text{H}] < 0.1$) and the Kroupa initial mass function (65). A binary fraction of 0.7 with mass ratio q uniformly distributed between 0.1 and 1 has been adopted. Finally, stellar evolutionary model predictions have been transferred to the Gaia photometric system by using appropriate bolometric correction tables using the prescriptions for the DR2 catalogue (66).

The results displayed in Figures 1, 2, and 3 are obtained with the α -enhanced models ($\alpha/[\text{Fe}]=0.4$ (8), which are the most adequate for thick disc and halo stellar populations. According to the analysis(5) of the metallicity and α abundances(11) of a fraction of stars in the blue and red sequences of the halo CMD, red sequence stars have $\alpha/[\text{Fe}] \simeq 0.3$ (with a mild dependence on metallicity), a value also typical for thick disc stars(67). Stars in the blue sequence have $\alpha/[\text{Fe}]$ that range between $\alpha/[\text{Fe}] \simeq 0.4$ at the lower metallicity extreme and $\alpha/[\text{Fe}] \simeq 0.15$ at the higher metallicity extreme ($[\text{Fe}/\text{H}] \simeq -0.8$ dex). However, we note that our conclusions, and in particular the age distributions, which are the most robust output of

our CMD fitting process, are not affected by the choice of a scaled-solar heavy element distribution; we expect, therefore, that they would be even less affected by the small differences in the $\alpha/[\text{Fe}]$ that would be required to match exactly the chemical characteristics of the different populations. A figure equivalent to 2 displaying the results obtained adopting a solar scaled heavy element distribution is provided as supplementary material to demonstrate this point.

Prior to the comparison between the observed Gaia CMDs and the *mother* synthetic CMDs, it is necessary to perform a simulation of the observational errors in the synthetic CMD, in such a way that it can be consistently compared with the observations. We have considered two error sources, namely: the error in the absolute magnitudes introduced by the formal error in the parallax, and the error in the Gaia photometry. To model the first error source, which dominates the error in M_G , we have statistically reproduced the distributions of `parallax` and `parallax_over_error` of the observed sample in the synthetic CMD (the `parallax_over_error` has been updated from the values provided in the Gaia DR2 archive taking into account the parallax offset). Then, the theoretical M_G values are transformed to account for observational errors by adding a correction given by a Gaussian distribution centred at 0 and with a standard deviation given by $\Delta M_G = 2.17 / \text{parallax_over_error}$. The maximum parallax errors in the samples used in this paper are of the order of 0.1 mas. Similarly, the errors in the Gaia photometry have been implemented by adding a dispersion of 0.004 magnitudes to the $(G_{BP} - G_{RP})$ of all synthetic stars at all magnitudes (66). We have not simulated uncertainties related to the extinction correction. These are expected to slightly smooth the CMD, and therefore the details in the star formation history. These uncertainties are of the order of 0.02-0.03 mag in $E(B-V)$ over most of the sky, and so one order or magnitude smaller than the dispersion of the CMD due to the parallax uncertainty (which vary between 0.0 and 0.3, with a typical value of 0.15).

The comparison between the observed and synthetic CMDs is performed using the number of stars in small -magnitude boxes. No a-priori constraint on the age-metallicity relation is adopted: *TheStorm* code solves simultaneously for age and metallicity within the range covered by the simple stellar populations. For more details on the fitting procedure and error calculation, the reader is referred to published references (52; 63; 68).

The results of this CMD fitting process are the star formation rate as a function of time and the age-metallicity distribution. From these, it is possible to compute a best-fit CMD (displayed in the lower panels of Figure 1) which is then used to derive the age and metallicity distributions displayed in Figures 2 and 3 (panel a). In the case of the halo CMD, the fitting process has been performed in the CMD including both the blue and the red sequences. The age and metallicity distributions presented in Figure 2 separately for each sequence have been derived by drawing a polygon enclosing each sequence in the observed CMD (avoiding the red clump where both sequences naturally overlap) and using this polygon to assign synthetic stars in the best fit CMD to either the blue or the red sequence. The boundary between the blue and red sequences corresponds to the fiducial line shown in Figure 1. We have tested that this particular delimitation has a negligible impact on the results presented here.

3 Cross-match with spectroscopic surveys

We used a homogeneous catalogue(33) of distances, 3-D velocities, and ages for the cross-match between major Milky-Way spectroscopic surveys and Gaia DR2. We have used this catalogue, together with the original Gaia DR2 information and the original $[\text{Fe}/\text{H}]$ information from LAMOST third data release A, F, G, K catalogue <http://dr3.lamost.org/> (`feb` value) (31) and GALAH first data release (`fe_h` value) (32), to select stars matching exactly the same spatial and kinematical selection as our CMDs. The cross-identifications were performed requiring exact match of the corresponding identification numbers

in the different catalogues (Gaia `source_id`, LAMOST `obsid` and GALAH `gaia_dr2_id`). We selected these two surveys because i) LAMOST includes a significant number of stars on the main-sequence turnoff region of the CMD overlapping with our halo sample (2,588 stars after quality cuts, see below), and with our thick disk sample (31,810 stars). SEGUE also has a significant overlap in the main sequence region, but the survey target selection function somehow selected mostly stars on the blue main sequence. The other two high resolution surveys in the catalog (GES and APOGEE) contained too few stars in the main sequence turnoff region to be used here. While GALAH offers only a small sample of stars overlapping with our halo sample (282 stars after quality cut), it include main sequence stars covering both blue and red sequences (133 and 125 stars, respectively), and it provides a larger overlap with the thick disc sample (1460 stars), made essentially of giant stars. The collected spectroscopic samples were used in Figure 2 to build a metallicity distribution for the blue and red sequence halo, as well as the thick disc, and in Figure 4 to explore the 3D kinematics of the samples in cylindrical coordinates.

Some quality flags were used to clean the spectroscopic samples: both were required to have a quality flag=0 in the compilation catalogue (33), as advised by the authors. Additionally, LAMOST stars were considered only when `snrg` ≥ 30 and GALAH stars were considered only when `flag_cannon` = 0.

4 Galaxy simulation

Cosmological simulations of the Milky Way’s halo have generally been done by simulating halos of similar mass to the Milky Way ($\sim 10^{12} M_{\odot}$), often imposing a relatively quiet merger history for the Milky Way since redshift ~ 1 . A range of increasingly detailed explorations of the origins of high kinematic stars have been made using such simulations (69; 70; 71; 72; 73; 74; 75; 76; 77).

For this study, we had a suite of 18 Milky Way mass galaxies from the NIHAO (78) and MaGICC (7) programs, and selected the one which had a merger history and properties most like the Milky Way. The simulation used for Figures 3 and 4 comes from a fully cosmological, hydrodynamical simulation of a Milky Way mass analogue galaxy from the MaGICC program. Importantly, these simulated galaxies have been shown to match a wide variety of galaxy scaling relations over a wide mass range, as well as reproducing a range of other galaxy properties such as enrichment of the Circum-Galactic Medium (79). The particular realisation, g15784, has total mass of 1.4×10^{12} solar masses and has been shown to be in good agreement with the Milky Way in terms of its detailed chemical and kinematic properties of the thick and thin discs (80). In the present study we show that the stars in the region around the Sun match the most important properties of the stars observed in our own Milky Way, and use this to interpret the red and blue sequences of the observed halo CMD as belonging to the main Milky Way progenitor and Gaia-Enceladus respectively, with the merger of these two proto-galaxies resulting in the in-situ and vast majority of the inner stellar halo.

In the simulation, stars form when gas has become sufficiently cool ($T < 10,000$ K) and sufficiently dense ($n_{\text{th}} > 9.3 \text{ cm}^{-3}$). When gas particles meet these criteria, stars form according to: $dM_{*}/dt = c_{*} M_{\text{gas}}/t_{\text{dyn}}$, where M_{*} is the mass of stars formed in time dt , M_{gas} is the mass of a gas particle, t_{dyn} is the dynamical time of gas particles, and c_{*} is the star formation efficiency – i.e., the fraction of gas that will be converted into stars. Supernova feedback follows the blastwave model (81) with thermal energy (10^{51} erg) deposited to the surrounding ISM from each supernova. The simulations also include radiation feedback from massive stars (82). Cooling is disabled in the blast region (~ 100 pc) for ~ 10 Myr. The simulation includes heating from a uniform UV ionising background radiation field (83). Cooling takes into account both primordial gas and metals (84). The metal cooling grid is derived using CLOUDY (85).

Along with the fact that the same physical prescriptions used in this simulation resulted in matching scaling relations at a range of masses, meaning that low mass accreted satellites contain the correct amount

of stars and gas for their halo mass, the correspondence with the observations of the thick disc gives us confidence in using this Milky Way analogue as an interpretive tool, as we do in this paper.

5 Data availability

All data analyzed in this paper are publicly available. They can be retrieved from the DR2 Gaia archive (<http://gea.esac.esa.int/archive/>) and from the (33) compiled and supplemented GALAH and LAMOST catalogues, cross-matched as indicated in Methods. The datasets which contain information that is not available in public catalogues, and that is necessary to produce the figures can be retrieved from <https://owncloud.iac.es/index.php/s/I50sjdlzvbSsMlu>, where an explanatory README file is included (a different repository can be chosen for the final version of the paper). The extended dataset consists of: i) (**a** and **b** panels of figure 1): two tables directly retrieved from the Gaia archive as described in Methods, supplemented by extinction information on a star-by-star basis; the code used to interpolate the 3-D extinction maps (44) can be retrieved from https://github.com/edober/dust_maps_3d. ii) (**c** and **d** panels of Figure 1, Figure 2 and panel **a** of Figure 3): two tables with the derived solution CMDs. Two files with the necessary data to define the boxes used to select stars in the blue and red sequences of the halo CMD are also included. iii) (panel **b** of Figure 3): necessary tables with the age, metallicity and velocity data for the main progenitor and accreted satellite for realisation g15784 of the MaGICC program (7). The complete information on the final timestep of that simulation, together with scripts to read and plot the data are also included in the extended dataset.

References

1. Soderblom, D. R. The Ages of Stars. *Annual Rev. Astron. Astrophys.* **48**, 581–629 (2010).
2. Gallart, C., Zoccali, M. & Aparicio, A. The Adequacy of Stellar Evolution Models for the Interpretation of the Color-Magnitude Diagrams of Resolved Stellar Populations. *Annual Rev. Astron. Astrophys.* **43**, 387–434 (2005).
3. Gaia Collaboration *et al.* Gaia Data Release 2. Summary of the contents and survey properties. *Astron. Astrophys.* **616**, A1 (2018).
4. Gaia Collaboration *et al.* Gaia Data Release 2. Observational Hertzsprung–Russell diagrams. *Astron. Astrophys.* **616**, A10 (2018).
5. Haywood, M. *et al.* In Disguise or Out of Reach: First Clues about In Situ and Accreted Stars in the Stellar Halo of the Milky Way from Gaia DR2. *Astrophys. J.* **863**, 113 (2018).
6. Helmi, A. *et al.* The merger that led to the formation of the Milky Way’s inner stellar halo and thick disk. *Nature* **563**, 85–88 (2018).
7. Brook, C. B., Stinson, G., Gibson, B. K., Wadsley, J. & Quinn, T. MaGICC discs: matching observed galaxy relationships over a wide stellar mass range. *Mon. Not. R. Astron. Soc.* **424**, 1275–1283 (2012).
8. Pietrinferni, A., Cassisi, S., Salaris, M. & Castelli, F. A Large Stellar Evolution Database for Population Synthesis Studies. II. Stellar Models and Isochrones for an α -enhanced Metal Distribution. *Astrophys. J.* **642**, 797–812 (2006).
9. Gilmore, G. & Reid, N. New light on faint stars. III - Galactic structure towards the South Pole and the Galactic thick disc. *Mon. Not. R. Astron. Soc.* **202**, 1025–1047 (1983).

10. Brook, C. B., Kawata, D., Gibson, B. K. & Flynn, C. Galactic Halo Stars in Phase Space: A Hint of Satellite Accretion? *Astrophys. J.* **585**, L125–L129 (2003).
11. Nissen, P. E. & Schuster, W. J. Two distinct halo populations in the solar neighborhood. Evidence from stellar abundance ratios and kinematics. *Astron. Astrophys.* **511**, L10 (2010).
12. Schuster, W. J., Moreno, E., Nissen, P. E. & Pichardo, B. Two distinct halo populations in the solar neighborhood. III. Evidence from stellar ages and orbital parameters. *Astron. Astrophys.* **538**, A21 (2012).
13. Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E. & Deason, A. J. Co-formation of the disc and the stellar halo. *Mon. Not. R. Astron. Soc.* **478**, 611–619 (2018).
14. Hawkins, K., Jofré, P., Gilmore, G. & Masseron, T. On the relative ages of the α -rich and α -poor stellar populations in the Galactic halo. *Mon. Not. R. Astron. Soc.* **445**, 2575–2588 (2014).
15. Ge, Z. S. *et al.* Ages of 70 Dwarfs of Three Populations in the Solar Neighborhood: Considering O and C Abundances in Stellar Models. *Astrophys. J.* **833**, 161 (2016).
16. Hawkins, K., Jofré, P., Masseron, T. & Gilmore, G. Using chemical tagging to redefine the interface of the Galactic disc and halo. *Mon. Not. R. Astron. Soc.* **453**, 758–774 (2015).
17. Bonaca, A., Conroy, C., Wetzel, A., Hopkins, P. F. & Kereš, D. Gaia Reveals a Metal-rich, in situ Component of the Local Stellar Halo. *Astrophys. J.* **845**, 101 (2017).
18. Hayes, C. R. *et al.* Disentangling the Galactic Halo with APOGEE. I. Chemical and Kinematical Investigation of Distinct Metal-poor Populations. *Astrophys. J.* **852**, 49 (2018).
19. Fernández-Alvar, E. *et al.* Disentangling the Galactic Halo with APOGEE. II. Chemical and Star Formation Histories for the Two Distinct Populations. *Astrophys. J.* **852**, 50 (2018).
20. Mackereth, J. T. *et al.* The origin of accreted stellar halo populations in the Milky Way using APOGEE, Gaia, and the EAGLE simulations. *Mon. Not. R. Astron. Soc.* **482**, 3426–3442 (2019).
21. Di Matteo, P. *et al.* The Milky Way has no in-situ halo but it has a thick disc. Composition of the stellar halo and age-dating the last significant merger with Gaia DR2 and APOGEE. *arXiv e-prints* (2018).
22. Erb, D. K. *et al.* The Mass-Metallicity Relation at $z > \sim 2$. *Astrophys. J.* **644**, 813–828 (2006).
23. Ma, X. *et al.* The origin and evolution of the galaxy mass-metallicity relation. *Mon. Not. R. Astron. Soc.* **456**, 2140–2156 (2016).
24. Behroozi, P. S., Wechsler, R. H. & Conroy, C. The Average Star Formation Histories of Galaxies in Dark Matter Halos from $z = 0$ -8. *Astrophys. J.* **770**, 57 (2013).
25. Zolotov, A. *et al.* The Dual Origin of Stellar Halos. *Astrophys. J.* **702**, 1058–1067 (2009).
26. Beers, T. C. *et al.* Population Studies. XIII. A New Analysis of the Bidelman-MacConnell “Weak-metal” Stars Confirmation of Metal-poor Stars in the Thick Disk of the Galaxy. *Astrophys. J.* **794**, 58 (2014).

27. Brook, C. B., Kawata, D., Gibson, B. K. & Freeman, K. C. The Emergence of the Thick Disk in a Cold Dark Matter Universe. *Astrophys. J.* **612**, 894–899 (2004).
28. Fuhrmann, K. Nearby stars of the Galactic disc and halo - V. *Mon. Not. R. Astron. Soc.* **414**, 2893–2922 (2011).
29. Recio-Blanco, A. *et al.* The Gaia-ESO Survey: the Galactic thick to thin disc transition. *Astron. Astrophys.* **567**, A5 (2014).
30. Haywood, M. *et al.* When the Milky Way turned off the lights: APOGEE provides evidence of star formation quenching in our Galaxy. *Astron. Astrophys.* **589**, A66 (2016).
31. Zhao, G., Zhao, Y.-H., Chu, Y.-Q., Jing, Y.-P. & Deng, L.-C. LAMOST spectral survey – An overview. *Research in Astronomy and Astrophysics* **12**, 723–734 (2012).
32. Buder, S. *et al.* The GALAH Survey: second data release. *Mon. Not. R. Astron. Soc.* **478**, 4513–4552 (2018).
33. Sanders, J. L. & Das, P. Isochrone ages for 3 million stars with the second Gaia data release. *Mon. Not. R. Astron. Soc.* **481**, 4093–4110 (2018).
34. Schönrich, R., McMillan, P. & Eyer, L. Distances and parallax bias in Gaia DR2. *arXiv:1902.02355* (2019).
35. Lindegren, L. *et al.* Gaia Data Release 2. The astrometric solution. *Astron. Astrophys.* **616**, A2 (2018).
36. Graczyk, D. *et al.* Testing Systematics of Gaia DR2 Parallaxes with Empirical Surface Brightness: Color Relations Applied to Eclipsing Binaries. *Astrophys. J.* **872**, 85 (2019).
37. Hall, O. J. *et al.* Testing asteroseismology with Gaia DR2: hierarchical models of the Red Clump. *Mon. Not. R. Astron. Soc.* **486**, 3569–3585 (2019).
38. Khan, S. *et al.* New light on the Gaia DR2 parallax zero-point: influence of the asteroseismic approach, in and beyond the Kepler field. *arXiv:1904.05676* (2019).
39. Leung, H. W. & Bovy, J. Simultaneous calibration of spectro-photometric distances and the Gaia DR2 parallax zero-point offset with deep learning. *arXiv:1902.08634* (2019).
40. Muraveva, T., Delgado, H. E., Clementini, G., Sarro, L. M. & Garofalo, A. RR Lyrae stars as standard candles in the Gaia Data Release 2 Era. *Mon. Not. R. Astron. Soc.* **481**, 1195–1211 (2018).
41. Zinn, J. C., Pinsonneault, M. H., Huber, D. & Stello, D. Confirmation of the *Gaia* DR2 parallax zero-point offset using asteroseismology and spectroscopy in the *Kepler* field. *arXiv:1805.02650* (2018).
42. Luri, X. *et al.* Gaia Data Release 2. Using Gaia parallaxes. *Astron. Astrophys.* **616**, A9 (2018).
43. Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G. & Andrae, R. Estimating Distance from Parallaxes. IV. Distances to 1.33 Billion Stars in Gaia Data Release 2. *Astron. J.* **156**, 58 (2018).
44. Lallement, R. *et al.* Three-dimensional maps of interstellar dust in the Local Arm: using Gaia, 2MASS, and APOGEE-DR14. *Astron. Astrophys.* **616**, A132 (2018).

45. Casagrande, L. & Vandenberg, D. A. On the use of Gaia magnitudes and new tables of bolometric corrections. *Mon. Not. R. Astron. Soc.* **479**, L102–L107 (2018).
46. Gallart, C., Freedman, W. L., Aparicio, A., Bertelli, G. & Chiosi, C. The Star Formation History of the Local Group Dwarf Galaxy Leo I. *Astron. J.* **118**, 2245–2261 (1999).
47. Hernandez, X., Valls-Gabaud, D. & Gilmore, G. Deriving star formation histories: inverting Hertzsprung-Russell diagrams through a variational calculus maximum likelihood method. *Mon. Not. R. Astron. Soc.* **304**, 705–719 (1999).
48. Holtzman, J. A. *et al.* Observations and Implications of the Star Formation History of the Large Magellanic Cloud. *Astron. J.* **118**, 2262–2279 (1999).
49. Dolphin, A. E. Numerical methods of star formation history measurement and applications to seven dwarf spheroidals. *Mon. Not. R. Astron. Soc.* **332**, 91–108 (2002).
50. Aparicio, A. & Gallart, C. IAC-STAR: A Code for Synthetic Color-Magnitude Diagram Computation. *Astron. J.* **128**, 1465–1477 (2004).
51. Aparicio, A. & Hidalgo, S. L. IAC-pop: Finding the Star Formation History of Resolved Galaxies. *Astron. J.* **138**, 558–567 (2009).
52. Monelli, M. *et al.* The ACS LCID Project. III. The Star Formation History of the Cetus dSph Galaxy: A Post-reionization Fossil. *Astrophys. J.* **720**, 1225–1245 (2010).
53. Cignoni, M. & Tosi, M. Star Formation Histories of Dwarf Galaxies from the Colour-Magnitude Diagrams of Their Resolved Stellar Populations. *Advances in Astronomy* **2010**, 158568 (2010).
54. Ruiz-Lara, T. *et al.* Integrated-light analyses vs. colour-magnitude diagrams. II. Leo A: an extremely young dwarf in the Local Group. *Astron. Astrophys.* **617**, A18 (2018).
55. Tolstoy, E., Hill, V. & Tosi, M. Star-Formation Histories, Abundances, and Kinematics of Dwarf Galaxies in the Local Group. *Annual Rev. Astron. Astrophys.* **47**, 371–425 (2009).
56. Benedict, G. F. *et al.* Hubble Space Telescope Fine Guidance Sensor Parallaxes of Galactic Cepheid Variable Stars: Period-Luminosity Relations. *Astron. J.* **133**, 1810–1827 (2007).
57. Beaton, R. L. *et al.* The Carnegie-Chicago Hubble Program. I. An Independent Approach to the Extragalactic Distance Scale Using Only Population II Distance Indicators. *Astrophys. J.* **832**, 210 (2016).
58. Hernandez, X., Valls-Gabaud, D. & Gilmore, G. The recent star formation history of the Hipparcos solar neighbourhood. *Mon. Not. R. Astron. Soc.* **316**, 605–612 (2000).
59. Bertelli, G. & Nasi, E. Star Formation History in the Solar Vicinity. *Astron. J.* **121**, 1013–1023 (2001).
60. Cignoni, M., Degl’Innocenti, S., Prada Moroni, P. G. & Shore, S. N. Recovering the star formation rate in the solar neighborhood. *Astron. Astrophys.* **459**, 783–796 (2006).

61. Bernard, E. J. Reconstructing the star formation history of the Solar neighbourhood with Gaia. In Chiappini, C., Minchev, I., Starkenburg, E. & Valentini, M. (eds.) *Rediscovering Our Galaxy*, vol. 334 of *IAU Symposium*, 158–161 (2018).
62. Bernard, E. J. *et al.* The spatially-resolved star formation history of the M31 outer disc. *Mon. Not. R. Astron. Soc.* **453**, L113–L117 (2015).
63. Bernard, E. J. *et al.* Star formation history of the Galactic bulge from deep HST imaging of low reddening windows. *Mon. Not. R. Astron. Soc.* **477**, 3507–3519 (2018).
64. Pietrinferni, A., Cassisi, S., Salaris, M. & Castelli, F. A Large Stellar Evolution Database for Population Synthesis Studies. I. Scaled Solar Models and Isochrones. *Astrophys. J.* **612**, 168–190 (2004).
65. Kroupa, P., Tout, C. A. & Gilmore, G. The distribution of low-mass stars in the Galactic disc. *Mon. Not. R. Astron. Soc.* **262**, 545–587 (1993).
66. Evans, D. W. *et al.* Gaia Data Release 2. Photometric content and validation. *Astron. Astrophys.* **616**, A4 (2018).
67. Hayden, M. R. *et al.* Chemical Cartography with APOGEE: Metallicity Distribution Functions and the Chemical Structure of the Milky Way Disk. *Astrophys. J.* **808**, 132 (2015).
68. Hidalgo, S. L. *et al.* The ACS LCID Project. V. The Star Formation History of the Dwarf Galaxy LGS-3: Clues to Cosmic Reionization and Feedback. *Astrophys. J.* **730**, 14 (2011).
69. Brook, C. B., Kawata, D., Gibson, B. K. & Flynn, C. Stellar halo constraints on simulated late-type galaxies. *Mon. Not. R. Astron. Soc.* **349**, 52–56 (2004).
70. Okamoto, T., Eke, V. R., Frenk, C. S. & Jenkins, A. Effects of feedback on the morphology of galaxy discs. *Mon. Not. R. Astron. Soc.* **363**, 1299–1314 (2005).
71. Font, A. S. *et al.* Cosmological simulations of the formation of the stellar haloes around disc galaxies. *Mon. Not. R. Astron. Soc.* **416**, 2802–2820 (2011).
72. Pillepich, A. *et al.* Halo mass and assembly history exposed in the faint outskirts: the stellar and dark matter haloes of Illustris galaxies. *Mon. Not. R. Astron. Soc.* **444**, 237–249 (2014).
73. Rodriguez-Gomez, V. *et al.* The stellar mass assembly of galaxies in the Illustris simulation: growth by mergers and the spatial distribution of accreted stars. *Mon. Not. R. Astron. Soc.* **458**, 2371–2390 (2016).
74. Tissera, P. B. *et al.* The central spheroids of Milky Way mass-sized galaxies. *Mon. Not. R. Astron. Soc.* **473**, 1656–1666 (2018).
75. Obreja, A. *et al.* Introducing galactic structure finder: the multiple stellar kinematic structures of a simulated Milky Way mass galaxy. *Mon. Not. R. Astron. Soc.* **477**, 4915–4930 (2018).
76. Monachesi, A. *et al.* The Auriga stellar haloes: connecting stellar population properties with accretion and merging history. *Mon. Not. R. Astron. Soc.* **485**, 2589–2616 (2019).

77. Fattahi, A. *et al.* The origin of galactic metal-rich stellar halo components with highly eccentric orbits. *Mon. Not. R. Astron. Soc.* **484**, 4471–4483 (2019).
78. Wang, L. *et al.* NIHAO project - I. Reproducing the inefficiency of galaxy formation across cosmic time with a large sample of cosmological hydrodynamical simulations. *Mon. Not. R. Astron. Soc.* **454**, 83–94 (2015).
79. Stinson, G. S. *et al.* MAGICC haloes: confronting simulations with observations of the circumgalactic medium at $z=0$. *Mon. Not. R. Astron. Soc.* **425**, 1270–1277 (2012).
80. Miranda, M. S. *et al.* Origin of the metallicity distribution in the thick disc. *Astron. Astrophys.* **587**, A10 (2016).
81. Stinson, G. *et al.* Star formation and feedback in smoothed particle hydrodynamic simulations - I. Isolated galaxies. *Mon. Not. R. Astron. Soc.* **373**, 1074–1090 (2006).
82. Stinson, G. S. *et al.* Making Galaxies In a Cosmological Context: the need for early stellar feedback. *Mon. Not. R. Astron. Soc.* **428**, 129–140 (2013).
83. Haardt, F. & Madau, P. Radiative Transfer in a Clumpy Universe. II. The Ultraviolet Extragalactic Background. *Astrophys. J.* **461**, 20 (1996).
84. Shen, S., Wadsley, J. & Stinson, G. The enrichment of the intergalactic medium with adiabatic feedback - I. Metal cooling and metal diffusion. *Mon. Not. R. Astron. Soc.* **407**, 1581–1596 (2010).
85. Ferland, G. J. *et al.* CLOUDY 90: Numerical Simulation of Plasmas and Their Spectra. *Pubbl. Astron. Soc. Pacific* **110**, 761–778 (1998).

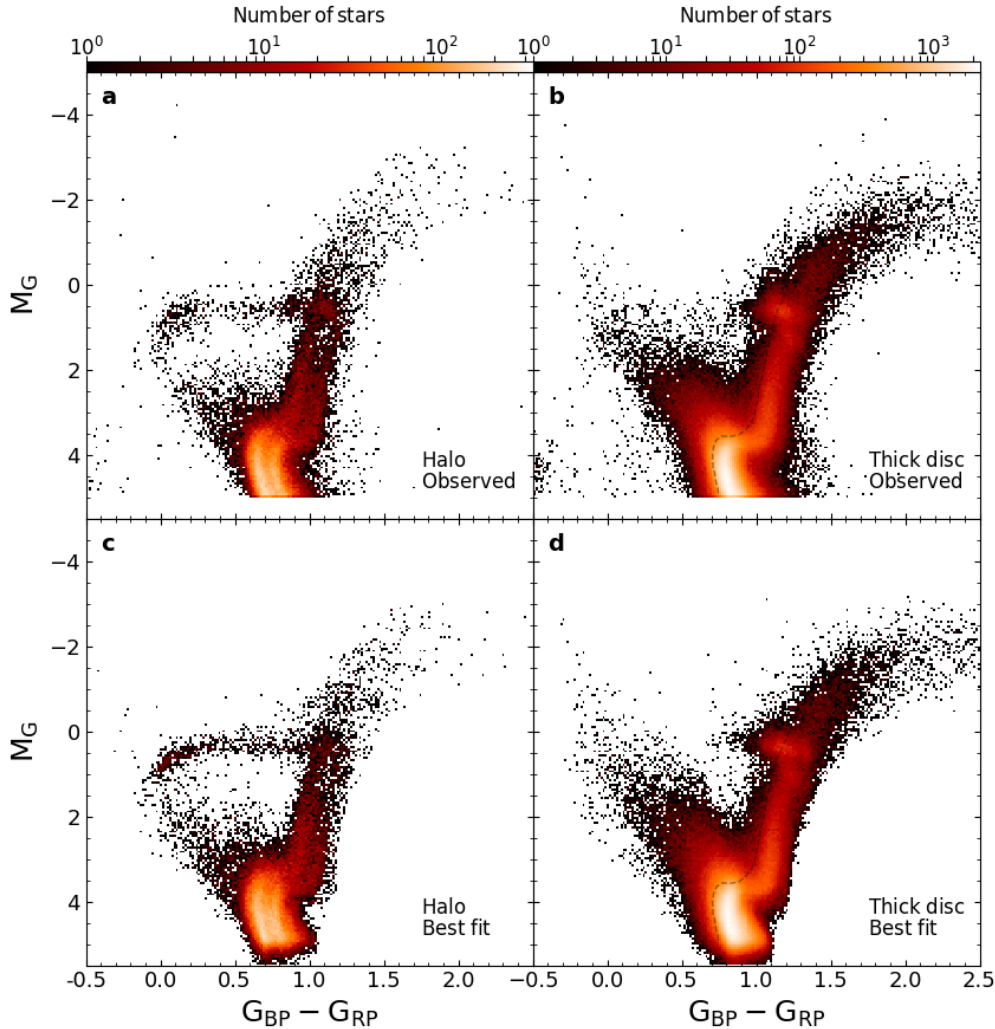


Figure 1. Milky Way halo and thick disc observed and modelled colour-Magnitude Diagrams (CMDs). Panels **a** and **b** display the observed Gaia CMDs for the halo and thick disc sub-populations, respectively. They have been represented as Hess diagrams with a logarithmic normalization of the number of stars in the colour-magnitude bins. Two sequences are clear in panel **a**: the blue sequence on the left and the red sequence on the right. Panels **c** and **d** display the best-fit models to the halo and thick disc CMDs, respectively. The solutions for the fit using α -enhanced models have been represented in this Figure, see Methods. The solutions obtained using the solar scaled models are not noticeably different. The dashed line in panels **b** and **d** indicates the position of the gap between the blue and red halo sequences. See Methods for more information on the sample selection and the CMD fitting technique.

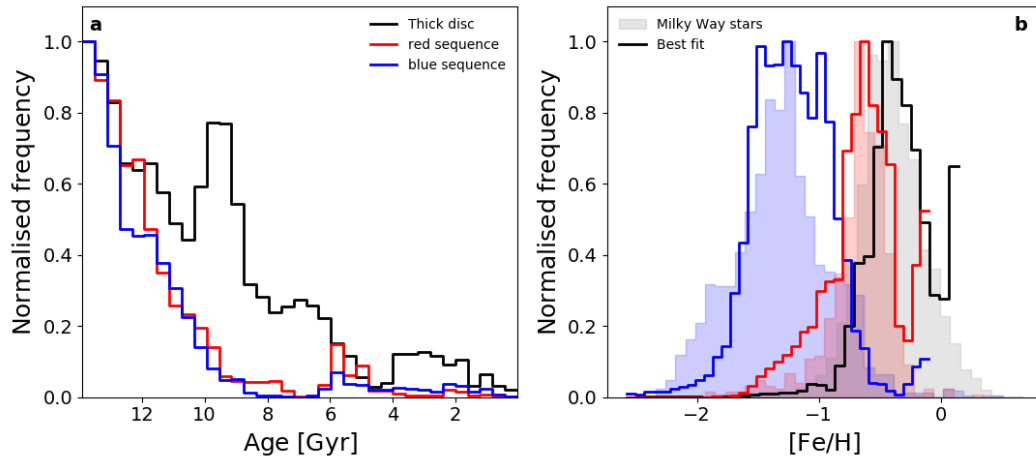


Figure 2. Milky Way halo and thick disc stellar age and iron [Fe/H] distributions. Panel **a** shows the distribution of stellar ages of the thick disc (grey histogram), blue (blue histogram) and red (red histogram) halo sequences according to their best-fit model CMDs adopting the α -enhanced stellar evolution models. The solutions obtained using the solar scaled models show similar features. Panel **b** shows the distribution of [Fe/H] of the best-fitting models for the blue and red halo sequences as well as the thick disc stars (solid lines), compared to [Fe/H] distributions for the stars in each sample that had measured [Fe/H] in the LAMOST or GALAH spectroscopic surveys(31; 32; 33) (filled histograms). The fair agreement between the [Fe/H] distributions inferred with the CMD fitting technique and those measured spectroscopically provides an external check of the reliability of the results.

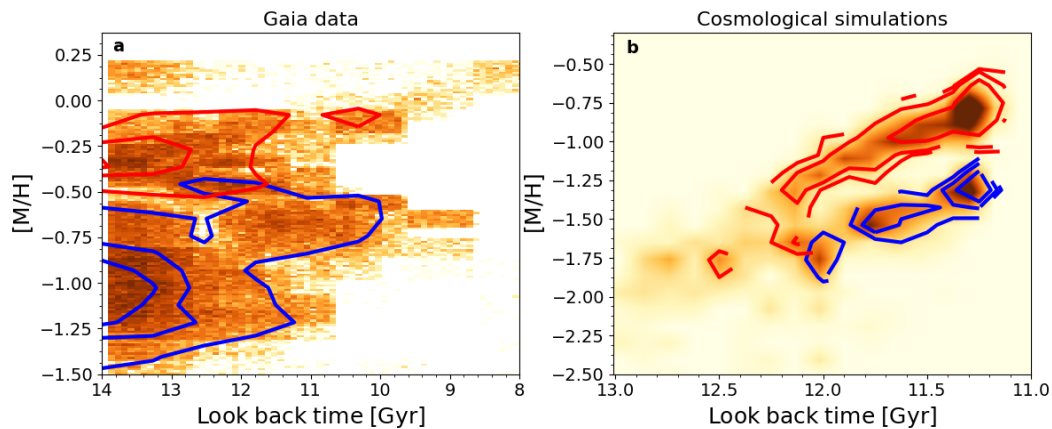


Figure 3. Age-metallicity relations for the two halo populations as compared to a simulated Milky Way analogue. Panel **a**: age-metallicity relation for the best-fit model of the halo CMD. The blue and red contours outline the distribution of stars in the blue and red halo sequences, respectively, overlaid to the general distribution of halo stars. Panel **b**, same as **a** but for a simulated Milky Way analogue (see Methods). Blue and red contours depict the age-metallicity distribution of simulated stellar particles belonging to an accreted galaxy (analogue to Gaia-Enceladus) and the Galaxy progenitor, respectively. The last major merger can be identified as two tracks in the age-metallicity plot, with the higher metallicity stars belonging to the main progenitor and the lower metallicity stars coming from the merged satellite. The agreement, only expected to be qualitative, is remarkable. The time and mass of the last major merger, which can be seen to differ between the simulation and the real Milky Way, determine the details of the ages and metallicities of the red and blue halo populations.

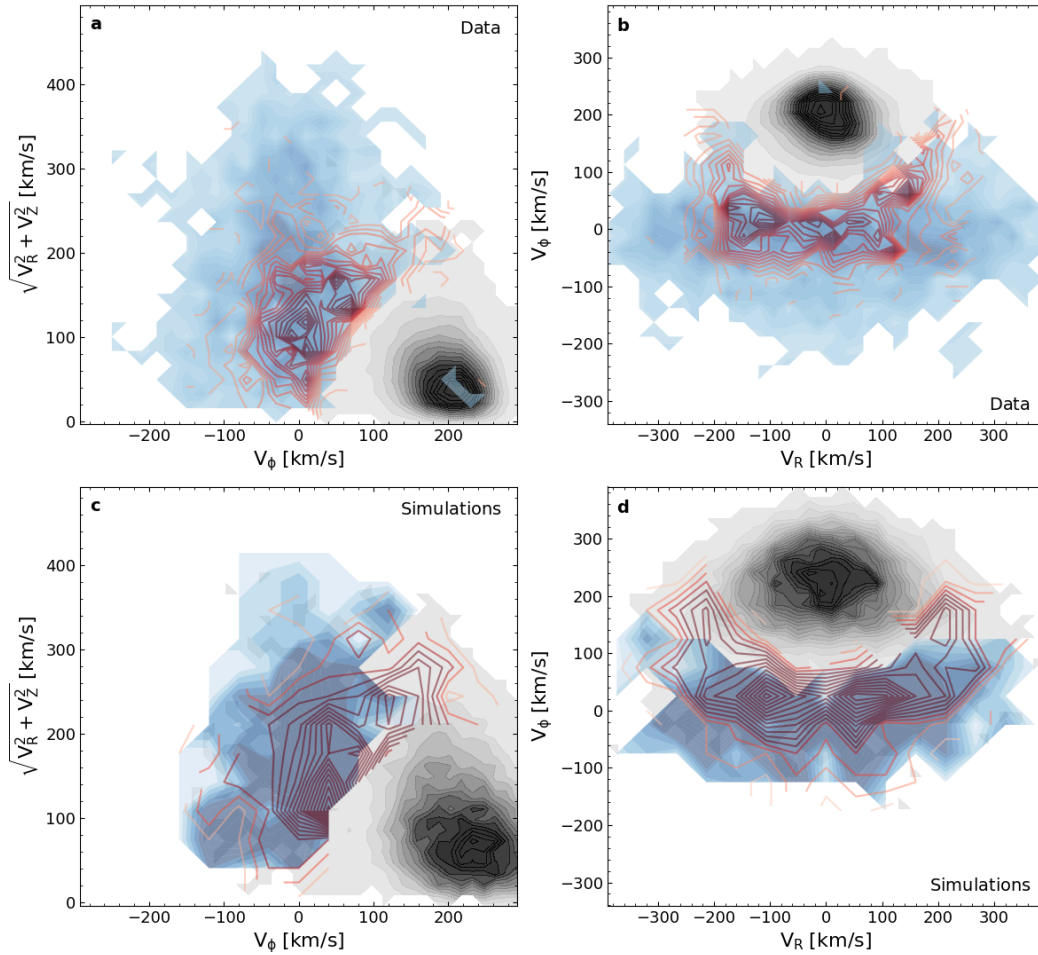


Figure 4. Comparison of the stellar kinematics between observations of Milky Way stars and the simulated Milky Way analogue. **a** and **c** show the Toomre Diagram for Milky Way stars (obtained by combining Gaia DR2 astrometry and LAMOST radial velocities(33)), and for the simulated Milky Way analogue galaxy, respectively. **b** and **d** display similar information but in an azimuthal velocity (V_ϕ) versus radial velocity (V_R) diagram. In the four panels, stars belonging to the observed halo blue sequence or the simulated accreted satellite are represented by blue shades, the observed halo red sequence or the simulated progenitor as red contours, and observed or simulated thick disc in grey scale. Note that some stars in our halo sample turn out to have disc properties when the full kinematic information is taken into account. This can help explain the presence of a small fraction of young stars in the halo sequences (panel a of Figure 2) as contamination from thin disc stars in the sample. The same is true for the inferred thick disc young population.