Nitrogen enrichment and clustered star formation at the dawn of the Galaxy

Vasily Belokurov[1]* and Andrey Kravtsov[2,3,4]†

[1]*Institute of Astronomy, Madingley Rd, Cambridge, CB3 0HA, UK*
[2]*Department of Astronomy and Astrophysics, The University of Chicago, Chicago, IL 60637 USA*
[3]*Kavli Institute for Cosmological Physics, The University of Chicago, Chicago, IL 60637 USA*
[4]*Enrico Fermi Institute, The University of Chicago, Chicago, IL 60637*

21 July 2023

ABSTRACT

Anomalously high nitrogen-to-oxygen abundance ratios [N/O] are observed in globular clusters (GCs), among the field stars of the Milky Way (MW), and even in the gas in a $z \approx 11$ galaxy. Using data from the APOGEE Data Release 17 and the *Gaia* Data Release 3, we present several independent lines of evidence that most of the MW's high-[N/O] stars were born in situ in massive bound clusters during the early, pre-disk evolution of the Galaxy. Specifically, we show that distributions of metallicity [Fe/H], energy, the angular momentum L_z, and distance of the low-metallicity high-[N/O] stars match the corresponding distributions of stars of the *Aurora* population *and* of the in-situ GCs. We also show that the fraction of in-situ field high-[N/O] stars, $f_{\rm N/O}$, increases rapidly with decreasing metallicity. During epochs when metallicity evolves from $[{\rm Fe/H}] = -1.5$ to $[{\rm Fe/H}] = -0.9$, the Galaxy spins up and transitions from a turbulent Aurora state to a coherently rotating disk. This transformation is accompanied by many qualitative changes. In particular, we show that high N/O abundances similar to those observed in GN-z11 were common before the spin-up ($[{\rm Fe/H}] \lesssim -1.5$) when up to $\approx 50\% - 70\%$ of the in-situ stars formed in massive bound clusters. The dramatic drop of $f_{\rm N/O}$ at $[{\rm Fe/H}] \gtrsim -0.9$ indicates that after the disk emerges the fraction of stars forming in massive bound clusters decreases by two orders of magnitude.

Key words: Galaxy: disc – The Galaxy: Galaxy: (Galaxy:) globular clusters: general – The Galaxy: Galaxy: evolution – The Galaxy: Galaxy evolution – The Galaxy: Galaxy: abundances – The Galaxy: galaxies: star clusters: general – Galaxies: galaxies: high-redshift – Galaxies

1 INTRODUCTION

The dawn of the Universal galaxy assembly can now be explored in tantalizing detail via direct high-resolution infrared JWST observations of high redshift galaxies (see, e.g., Donnan et al. 2023; Harikane et al. 2023; Finkelstein et al. 2023; Robertson et al. 2023). Thanks to the combined power of JWST's NIRCam and NIRSpec instruments, galaxies beyond $z = 8$ are revealed to be small, dense, metal-poor, and actively star-forming (see, e.g., Ono et al. 2022; Tacchella et al. 2023; Robertson et al. 2023; Bouwens et al. 2023; Curtis-Lake et al. 2023).

In the Milky Way (MW), independent constraints on the physics of high-z star formation can be obtained using Galactic archaeology which interprets properties of individual ancient, low-metallicity stars from large surveys. The two approaches are complementary: direct observations of high-z galaxies probe the star

formation and state of the interstellar medium enriched by the nucleosynthetic products from the first generations of massive stars, while Galactic archeology explores properties of the surviving low-mass stars that formed in the low-metallicity environment of the MW's progenitor.

With the NIRSpec instrument, the brightest of the high-z galaxies are now amenable to unprecedented levels of scrutiny, going far beyond a simple spectroscopic redshift measurement. One such example is GN-z11, identified previously with the HST and *Spitzer* (Oesch et al. 2016). The follow-up observations with NIRSpec reported measurements of oxygen, carbon, and neon lines, as well as unusually prominent levels of nitrogen emission, indicating high relative [N/O] abundance at moderately low oxygen abundance [O/H] (Bunker et al. 2023). Using these measurements Cameron et al. (2023) estimated the relative nitrogen abundance in GN-z11 to be $\log({\rm N/O}) \gtrsim -0.25$, which is much higher than this ratio in the Sun, i.e. $\log({\rm N/O})_\odot = -0.86$ (see Lodders 2019).

In the Galaxy's stellar populations such nitrogen overabundance is rare but exists: stars with high [N/O] ratio are numer-

* E-mail:vasily@ast.cam.ac.uk
† E-mail:kravtsov@uchicago.edu

ous in Globular Clusters (see, e.g., Bastian & Lardo 2018; Gratton et al. 2019; Milone & Marino 2022). This similarity in the abundance patterns and possible connection between enrichment pathways in local globular clusters and GN-z11 has been recently pointed out in several studies (Cameron et al. 2023; Senchyna et al. 2023; Charbonnel et al. 2023).

Nitrogen over-abundance in the MW GCs is so blatant that it has been used routinely as a chemical fingerprint to identify field stars that were born in clusters (see, e.g., Martell & Grebel 2010; Carollo et al. 2013; Martell et al. 2016; Schiavon et al. 2017; Fernández-Trincado et al. 2017; Tang et al. 2019; Horta et al. 2021b; Phillips et al. 2022). The cluster origin of such field stars is evidenced by their correlations and anti-correlations of abundances of different chemical elements similar to those observed in GCs: e.g., depleted [O/Fe] and [Mg/Fe] and enhanced [Al/Fe] (see e.g. Lind et al. 2015; Schiavon et al. 2017; Fernández-Trincado et al. 2020b; Horta et al. 2021b). Curiously, not all high-[N/O] stars show the rest of the GC-specific chemical pattern. For example, a population of N-enhanced giants discovered by Fernández-Trincado et al. (2020d) in the Magellanic Clouds is consistent with the typical MW field in other projections of the chemical abundance space.

Given that the GC-born stars are straightforward to pick out in the field, a number of studies estimated the overall fraction of Galactic stellar mass contributed by clusters under various assumptions. For example, most studies agree that the overall observed fraction of stars with high [N/O] (hereafter high-[N/O] stars) at $[Fe/H] < -1$ is rather low, $\approx 2\% - 3\%$, but somewhat higher estimates can be obtained depending on the threshold in nitrogen enrichment, the metallicity range and the location in the Galaxy used (see, e.g., Martell et al. 2016; Schiavon et al. 2017; Koch et al. 2019; Horta et al. 2021b). To convert the observed high-[N/O] fraction into the total stellar mass born in clusters assumptions are made as to the initial mass of the star clusters disrupted by $z = 0$.

Another factor is the role of the so-called "first population" (1P) GC stars in clusters. The 1P stars themselves have chemical abundances indistinguishable from the rest of the Galaxy's field but they are assumed to directly contribute to the anomalous chemistry of the "second generation" (2P), which is manifested in N, Na and Al enrichment and C, Mg and O depletion (Bastian & Lardo 2018; Gratton et al. 2019; Milone & Marino 2022). Using the hypothesis that the 1P stars could have made up to ∼90% of the initial cluster's mass (e.g., D'Ercole et al. 2008; Conroy 2012; Bastian et al. 2013), the measured ∼2% implies that ∼20% of the field stars is contributed by clusters (e.g., Martell et al. 2011). Assuming a more conservative value of the fraction of 1P stars, $f_{1P} \approx 0.5$ comparable to a typical fraction observed in surviving MW clusters (Milone et al. 2017), the inferred GC contribution to the Galactic halo's stellar mass is reduced considerably to ∼ 5% (see Koch et al. 2019).

Interpretation of such estimates, however, is not straightforward. In the metallicity range of the high-[N/O] stars, stellar population is a mix of stars brought in by other galaxies (accreted) and stars born in the MW's main progenitor (in-situ). Thus, the fraction of high-[N/O] halo stars is an average of stars born in clusters in all of the progenitor galaxies that contributed to the MW's stellar halo. To address this, several studies attempted to assign high-[N/O] stars to distinct halo components based, for example, on the [Al/Fe] ratio (Kisku et al. 2021; Fernández-Trincado et al. 2022). However, given the generally anomalous chemical abundances of many of the GC-born stars, including [Al/Fe], such an assignment is bound to be biased. Orbital information has also been used but with inconclusive results (Fernández-Trincado et al. 2020a; Tang et al. 2020; Fernández-Trincado et al. 2022).

Several studies reported a prominent population of N-rich stars residing in the Milky Way's bulge (see Schiavon et al. 2017; Fernández-Trincado et al. 2020c). The increased incidence of N-rich stars towards the Galactic centre is also reported in Horta et al. (2021b). Unfortunately, the star's presence in the bulge does not elucidate where and when it formed. Multiple origins remain viable: bulge stars can be part of either accreted or in-situ halo population or even belong to the bar.

Recently, data from the *Gaia* satellite has helped to clarify and systematize the make-up of the Galactic stellar halo. In confirmation of the hypothesis put forward by Deason et al. (2013), the bulk of the accreted debris within 40 kpc from the Galactic centre appears to be donated via a single, massive and ancient merger event known as the *Gaia* Sausage/Enceladus (GS/E, Belokurov et al. 2018; Helmi et al. 2018). A sizeable population of the Milky Way GCs was shown to belong to the GS/E progenitor dwarf galaxy by Myeong et al. (2018). Details of the Galactic GC classification have been considered and re-evaluated many a time since (e.g., Massari et al. 2019; Kruijssen et al. 2019; Myeong et al. 2019; Forbes 2020; Callingham et al. 2022). A consensus amongst these works is that many of the GCs formed in situ can be identified based on either their location in the age-metallicity space or by their high angular momentum. Several recent efforts have started to verify this nomenclature via high-resolution spectroscopic studies (e.g. Koch-Hansen et al. 2021; McKenzie et al. 2022; Monty et al. 2023).

Not surprisingly, across the above classification efforts, the in-situ GCs have lower average energy compared to the accreted ones. Sometimes a fraction of the low-energy GCs is assigned to a separate accretion event (Massari et al. 2019; Forbes 2020; Kruijssen et al. 2020; Callingham et al. 2022). It is unclear however what makes these low-energy GCs distinct from the rest of the in-situ clusters. In the original analysis of Massari et al. (2019), which inspired many of the follow-up works mentioned above, the in-situ GCs were split into two groups, the tightly bound "bulge" and the clearly rotating "disk" clusters.

Recently, however, the metal-poor ($[Fe/H] < -1$) portion of the in-situ stellar halo has been shown to have a much wider range of azimuthal velocities compared to the accreted component, at least in the vicinity of the Sun (Belokurov & Kravtsov 2022, hereafter BK22). Discovered through chemical tagging, this component, dubbed *Aurora*, is the oldest Milky Way stellar population formed before the Galaxy had a coherently rotating disk.

Aurora stellar population spans a range of energies, from the lowest levels typical for the stars near the Galactic centre to that of the Sun, but its density beyond the Solar radius falls sharply. The distribution of azimuthal velocities of the Aurora stars is significantly broader than that of the GS/E. However, unlike GS/E's debris which has little net rotation, Aurora has a modest net spin of ∼ 50 km s^{-1}. BK22 show that at higher metallicities ($[Fe/H] > -1$) the kinematic behaviour of the ancient MW stars exhibits a clear trend: the azimuthal velocity increases sharply as the Galaxy spins up to become a disk (see also Conroy et al. 2022; Rix et al. 2022).

Aurora's stars also exhibit a large scatter in most elements but in particular in Al, N, O, and Si, i.e. the same elements that are considered as the best GC markers due to their anomalous behaviour. Consequently, BK22 conclude that Aurora's chemistry likely bears signs of a large contribution from massive star clusters, a hypothesis strengthened in the analysis of Myeong et al. (2022).

Thus there are multiple indications that instead of an agglomeration of numerous fragments of distinct origin, the metal-poor

stellar halo inside the Solar radius is dominated by one prominent population formed in-situ at early epochs. Throughout this Paper we refer to this early pre-disk component formed during the cosmic dawn stages of Galaxy evolution as *Aurora*.[1] This is to emphasize the in-situ origin of this component thus accepting that it may contain most or all of the alleged Kraken/Koala/Heracles structure (Kruijssen et al. 2019; Horta et al. 2021a; Forbes 2020).

The main reasons to consider such a single-origin classification scheme for the bulk of the metal-poor portion in the inner MW halo are threefold. First, detailed high-resolution chemical studies show little difference between Aurora and Kraken/Koala/Heracles (BK22, Naidu et al. 2022; Myeong et al. 2022; Horta et al. 2023a). Second, a look at the chemo-kinematics of these stars shows a clear continuity in their orbital properties across a wide range of angular momenta, from retrograde to mildly rotating and a wide range of energies, from the most bound, "bulge"-like to approximately Solar (Arentsen et al. 2020a,b, BK22, Conroy et al. 2022; Myeong et al. 2022; Rix et al. 2022). Finally, recent analysis of Semenov et al. (2023, see their Figure 10) based on the TNG50 simulations shows that mergers contribute on average $\approx 20-30\%$ to the stellar component during pre-disk epochs indicating that in-situ component dominates in most galaxies for low-metallicity stellar population. Consequently, in our study, field stars and Galactic GCs classified as in-situ have a broader range of total energies and angular momenta than considered previously.

In this paper we aim to consistently identify the in-situ and accreted components of the MW's stellar population in the metallicity range probed by the APOGEE survey and of the MW's population of globular clusters. To this end, we use both the APOGEE measurements of chemical abundances of elements and *Gaia* EDR3 measurements of proper motions of stars and globular clusters. We use the resulting classification to estimate the fraction of stars with enhanced nitrogen abundance in the in-situ and accreted populations and the fraction of low-metallicity stars born in bound clusters. Given that we select such stars using the [N/O] ratio we will refer to these stars in the context of this study as high-[N/O] stars.

The paper is structured as follows. Section 2 presents the details of our selection of field stars, high-[N/O] stars and likely GC members. In Section 3, we analyze distributions of the selected stellar populations, decipher the origin of the field high-[N/O] stars and estimate the contribution of GC-like objects to star formation in the early Galaxy. We discuss the implications of our inference in Section 4 where we list the salient changes accompanying the MW's transition from the chaotic Aurora state to the stable disk. Section 5 lists our conclusions. In the Appendix 5 we provide a table with the definition or description of the key terms and notations used in this paper.

2 DATA AND SAMPLE SELECTION

We use element abundances from the APOGEE Data Release 17 (Abdurro'uf et al. 2021), as recorded in the `allStarLite` catalogue provided on the survey's website. Following BK22, we remove stars with flags: `STAR_BAD`, `TEFF_BAD`, `LOGG_BAD`, `VERY_BRIGHT_NEIGHBOR`, `LOW_SNR`, `PERSIST_HIGH`, `PERSIST_JUMP_POS`, `SUSPECT_RV_COMBINATION`, `PERSIST_JUMP_NEG`, as well as duplicate with `EXTRATARG`

---

flag. Distances are taken from the AstroNN value-added catalogue (see Leung & Bovy 2019; Mackereth & Bovy 2018). We rely on *Gaia* EDR3 proper motions (Gaia Collaboration et al. 2021; Lindegren et al. 2021) and convert observed heliocentric stellar coordinates into the Galactocentric left-handed reference frame, assuming that the Sun is at $X = R_\odot = 8$ kpc from the Galactic Centre (c.f. a slightly larger value from Gravity Collaboration et al. 2022), and has Galactic $Z_\odot = 0$. Following Gaia Collaboration et al. (2022), we assume that the Sun's velocity is $v_\odot = \{-9.3, 251.5, 8.59\}$ km s^{-1}. Total energies E are calculated in a three-component (bulge, disk and DM halo) Galaxy potential identical to that used in Belokurov et al. (2023). In what follows, energy is reported in units of 10^5 km^2 s^{-2} and the vertical component of the angular momentum L_z in units of 10^3 kpc km s^{-1}.

2.1 Sample of field red giants with low V_ϕ

Our base sample of field stars is selected as follows. First, we remove stars within 1.3 degree of all known Galactic satellites (in particular, globular clusters) as well as all objects with `PROGRAMNAME=magclouds`. Additionally, we consider only stars within 10 kpc from the Sun that are consistent with being red giants by using the following cuts: $D < 10$ kpc, $\log(g) < 3$ and $T_{\rm eff} < 5300$ K. We also cull stars with tangential velocity errors larger than 50 km s^{-1}, [Fe/H], [N/Fe] and [O/Fe] errors larger than 0.25 dex. Note that removing measurements with large uncertainties can bias [N/O] ratios high at low metallicities. We have checked for the presence of such bias by re-running the entirety of our analysis without a cut on abundance errors and report that any changes in the measurements reported are within their associated uncertainties. Finally, to get rid of the fast-rotating, young stars in the Galaxy's thin disk, we apply a cut on the tangential component of stellar velocity $V_\phi < 160$ km s^{-1}. The combination of the above cuts leaves a total of $\sim 30,000$ stars. Given the tangential velocity cut applied, stars in this sample are predominantly halo at [Fe/H]< -1 and high-α (thick) disk at [Fe/H]> -1.

2.2 Sample of globular cluster stars

We search the APOGEE DR17 catalog for likely Galactic globular cluster members using the following strategy. For each cluster, likely members are selected within 1.5 times the GC's tidal radius from its centre as reported in the 2010 version of the GC catalog of Harris (2010). Only stars with cluster membership probabilities above 0.5, as calculated by Vasiliev & Baumgardt (2021), are kept. Finally, we apply the same $\log(g)$ and $T_{\rm eff}$ cuts as above and retain only those stars whose [Fe/H] in APOGEE is within 0.3 dex of the GC catalog value. This selection procedure yields $\sim 4,200$ candidate GC members. When available, we use globular cluster distances from Baumgardt & Vasiliev (2021a) and model estimates of their initial masses from Baumgardt & Makino (2003), fractions of the 1st cluster population from Milone et al. (2017), and GC isochronal ages are from VandenBerg et al. (2013). Total energy and L_z angular momentum for each cluster are computed using the same assumptions about the Galaxy as described in the beginning of Section 2.

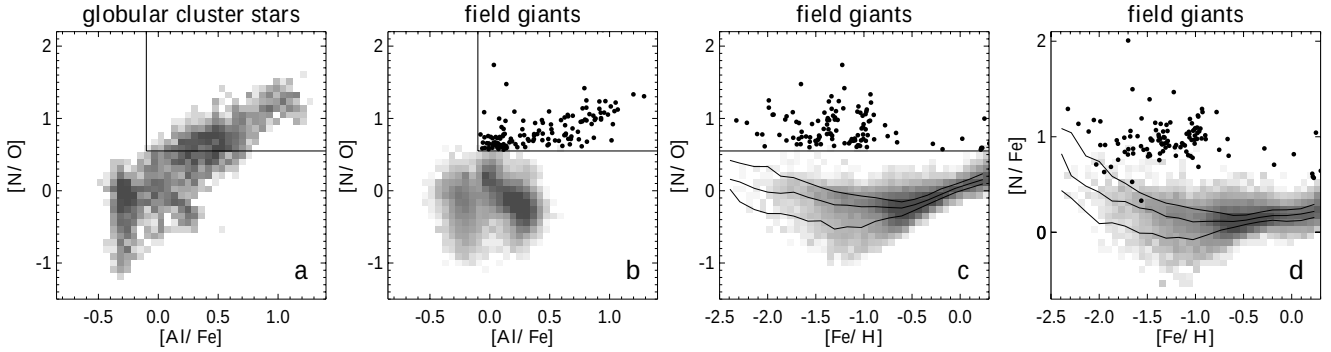
[1] After the Greek goddess of the Dawn – Aurora.

Figure 1. Stars with anomalous abundances in APOGEE DR17. **Panel a:** Greyscale shows the density of all globular cluster stars in our sample in the space of [N/O] and [Al/Fe]. Horizontal (vertical) line is the chosen [N/O] ([Al/Fe]) threshold for the selection of high-[N/O] stars with GC like chemical abundances. **Panel b:** Same as the previous panel but for field giants with $V_\phi < 160$ km/s. Small black points are the high-[N/O] stars with GC-like abundances. **Panel c:** Same as panel b, but for [N/O] vs [Fe/H]. Here and in the next panel, solid black lines show 20th, 50th and 80th percentiles of the abundance ratio distribution as a function of metallicity. **Panel d:** Same as panel c, but for [N/Fe] vs [Fe/H]. Note a stronger upward trend at low metallicity.

3   RESULTS

3.1   High-[N/O] field giants in the Milky Way

Our selection of stars with high nitrogen abundances is inspired by previous APOGEE-based studies (e.g. Schiavon et al. 2017; Horta et al. 2021b) with two minor tweaks, as illustrated in Figure 1. Instead of applying a cut on [N/Fe], we require $\mathrm{[N/O]} - \sigma_{\mathrm{[N/O]}} > 0.55$. Motivation for using [N/O] instead of [N/Fe] is twofold. First, there is a small but noticeable upward [N/Fe] trend with decreasing metallicity (see panel d of Figure 1) meaning that a selection based on a single [N/Fe] threshold is not viable. Second, we are looking to compare chemical properties of the MW stars to the extragalactic gas-phase abundances referenced to oxygen.

We choose the [N/O] threshold i) to match approximately the highest [N/O] ratios in the metal-rich disk population (see panel c of Figure 1) and ii) to reach the N/O levels observed in the high-redshift galaxy GN-z11 (see Section 3.10). Note, however, that the exact value of the adopted [N/O] threshold does not affect our results significantly because the derived fractions of GC-born stars are computed self-consistently using a calibration on surviving clusters (see Sections 3.4 and 3.6).

Panel a of Figure 1 shows the distribution of [N/O] as a function of [Al/Fe] in our sample of GC stars. Stars in GCs exhibit both anomalously high [N/O] and [Al/Fe] ratios. Therefore, for GC-like high-[N/O] stars, we use both the N/O threshold as well as a cut on aluminium, [Al/Fe]> -0.1 – these are shown in panel a with solid black lines.

Panel b of Figure 1 shows the distribution of field giants (selected using the criteria described above) in the space of [N/O] and [Al/Fe]. Although a proportionally much smaller number of high-[N/O] stars is observed in the field, most of them follow a correlation between [N/O] and [Al/Fe] very similar to GC stars. In fact, the overall distribution of GC stars in panel a and the field giants in panel b is quite similar for all values of [N/O] and [Al/Fe]. This is consistent with the conclusion we reached from the analyses presented in this paper that a large fraction of the low-metallicity field stars were born in massive bound clusters.

As demonstrated by panel c of the Figure, the number of high-[N/O] GC-like stars varies significantly with metallicity [Fe/H]. Black lines give the 20th, 50th and 80th percentiles of the [N/O] distribution as a function of [Fe/H]. At higher metallicities, i.e. at [Fe/H]> -0.5, the average [N/O] level starts to climb up due to the increased nitrogen contribution from intermediate-mass Asymp-

totic Giant Branch (AGB) stars (see, e.g., Kobayashi et al. 2020; Johnson et al. 2023). These nitrogen-rich and metal-rich stars are clearly distinct from the GC-born high-[N/O] stars because they do not exhibit any strong (anti)correlations typical of clusters (e.g., between Mg and Al). The bulk of these high-metallicity AGB-contributed disk stars with elevated [N/O] are removed by our $V_\phi < 160$ km/s cut. Also note a slight upward [N/O] trend with decreasing [Fe/H]. At least in part this is caused by culling measurements with high uncertainties. This weak [N/O] trend is however much flatter than a more noticeable increase in median [N/Fe] at low [Fe/H] as shown in panel d of the Figure (see also Figure 1 of Schiavon et al. 2017).

We proceed by classifying the field high-[N/O] GC-like stars into those born in-situ in the Milky Way (including Aurora) and those formed in dwarf galaxies (mainly GS/E) and subsequently incorporated into the accreted stellar halo.

3.2   Distinguishing accreted and in-situ stars and clusters

Due to a relatively slow pace of star formation and as a result of a strong metallicity dependence of the Al yield, dwarf galaxies never experience an over-abundance of Al compared to Fe, with the ratio [Al/Fe] staying low across a wide range of [Fe/H]. On the contrary, stars born in the Milky Way exhibit a rapid increase in [Al/Fe] around $-1.5 < \mathrm{[Fe/H]} < -0.9$. As shown in Hawkins et al. (2015), distinct behaviour of [Al/Fe] (and [Na/Fe]) can be utilized to separate accreted and in-situ halo components (see also Das et al. 2020).

Specifically, BK22 classify stars with $\mathrm{[Al/Fe]} > -0.075$ as in-situ and those with $\mathrm{[Al/Fe]} < -0.075$ as accreted. This approximate classification is supported by the observed abundance trends in the surviving massive MW dwarf satellites that typically have $\mathrm{[Al/Fe]} < -0.1$, as reported by Hasselquist et al. (2021). At low metallicities, $\mathrm{[Fe/H]} \lesssim -1.5$, the use of [Al/Fe] is not viable as the in-situ and the accreted sequences start to merge. More importantly, [Al/Fe]-based classification into accreted and in-situ is not viable for GC-like high-[N/O] stars due to their anomalously high [Al/Fe] ratios, $\mathrm{[Al/Fe]} > -0.1$.

We thus adopt a different two-stage approach to classify stars and GCs into in-situ and accreted populations. We first determine a boundary in the $E - L_z$ space that separates in-situ and accreted objects for the stars without anomalous ratios and at metallicities where [Al/Fe]-based classification is robust ($-1.4 < \mathrm{[Fe/H]} <$
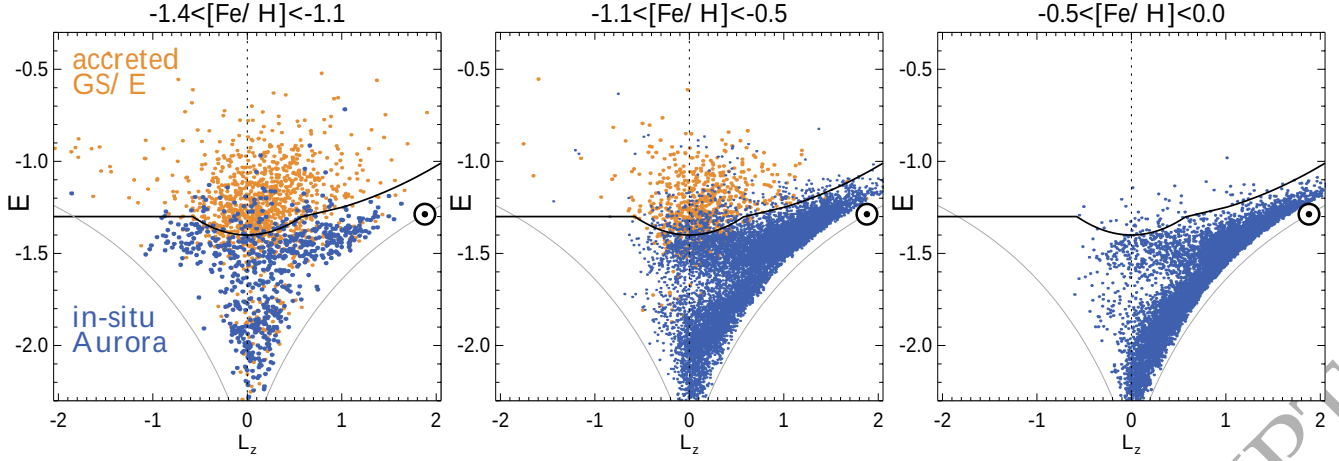
Figure 2. Orbital properties of halo giants. **Left:** Distribution of Aurora (blue, selected with $[Al/Fe] > -0.075$) and accreted (orange, selected with $[Al/Fe] < -0.075$ and an additional $[Mg/Fe]$ cut - see text for details) giants with $-1.4 < [Fe/H] < -1.1$ in the space of energy E ($\times 10^5$) and vertical component of angular momentum L_z ($\times 10^3$). Solid black line marks the decision boundary to separate stars into accreted (high energy) and Aurora (in-situ, low energy) populations. Grey solid lines correspond to the maximal angular momentum at fixed energy, i.e. orbits with circular velocity $V_\phi = V_{\rm circ}$. $\odot$ marks the location of the Sun in the chosen potential. **Middle:** Same as Left but for stars with $-1.1 < [Fe/H] < -0.5$. **Right:** Same as previous panels but for $-0.5 < [Fe/H] < 0$, note that in this metallicity range no accreted stars are visible.

-1.1), as shown in the left panel of Figure 2. The panel shows E and L_z distributions of stars classified as accreted (primarily GS/E, orange) and in-situ (Aurora, blue) using the $[Al/Fe] = -0.075$ threshold. In addition, following BK22 we apply the cuts of $[Mg/Fe] < -0.3\,[Fe/H] - 0.1$ to the accreted stars.

As the left panel of Figure 2 reveals, the distribution of accreted (GS/E) stars has a narrow range of $z-$component of angular momenta $|L_z| < 0.6$ and is limited in energy to $E \gtrsim -1.4$. This is in agreement with Belokurov et al. (2023), where a stellar sample based on the *Gaia* DR3 data was used for the analysis. The in-situ stars (blue points) on the other hand have a broader L_z distribution at energies similar to the lowest levels reached by the GS/E stars, i.e. $E \sim -1.4$ where the two groups have a small overlap. In-situ stars continue to lower energies where the occurrence of accreted stars (classified with $[Al/Fe]$ cut) is negligible and is likely to an occasional scatter of $[Al/Fe]$ values below the threshold.

The solid line approximates the boundary separating the in-situ and accreted populations in the $E - L_z$ space visible in the left panel and is described by the following equation:

$$
\begin{aligned}
L_z < -0.58 : & \ E = -1.3 \\
-0.58 < L_z < 0.58 : & \ E = -1.4 + 0.3 L_z^2 \\
L_z > 0.58 : & \ E = -1.325 + 0.075 L_z^2,
\end{aligned}
\tag{1}
$$

where E is in units of $10^5\,{\rm km}^2\,{\rm s}^{-2}$ and L_z is in units of $10^3\,{\rm kpc}\,{\rm km s}^{-1}$. We then test that the same boundary separates in-situ and accreted populations at other metallicities.

Middle panel of Figure 2 considers high metallicity stars, i.e. $-1.1 < [Fe/H] < -0.5$. While some GS/E debris is still visible, the E, L_z distribution is dominated by the in-situ stars, mainly high-α disk and the Splash (see Belokurov et al. 2020). Finally, the right panel of the Figure gives the distribution of stars with $[Fe/H] > -0.5$ and, unsurprisingly, shows no presence of any accreted debris. As we can see the same boundary of equation 1, shown by the solid line, separates the in-situ and accreted components well. Thus, in what follows, we use this boundary to classify both stars and GCs of all metallicities into in-situ and accreted.

We note that the dominance of high-α and low-$[Fe/H]$ Aurora

at low energies explains the trends reported in Donlon & Newberg (2023) without invoking additional accretion events.

Although the boundary in Equation 1 is categorical and was derived as a simple approximation to the distribution of the in-situ and accreted stars, we have also tested it using machine learning classification. Specifically, we used the "Gradient Boosted Trees" (GBT) machine learning method (Friedman 2001), implemented in the `GradientBoostingClassifier` class in the Sci-kit Learn package (Pedregosa et al. 2011), which allows for overlap between classes by assigning a class probability for stars in the overlap region. The method was trained with 80% of the sample of stars with reliable $[Al/Fe]$-based classification, while the remaining 20% of stars was used to test classification accuracy. We then computed the classification accuracy obtained with GBT and with the categorical boundary of Equation 1 for the test sample finding that both methods result in $\approx 96\%$ accuracy in classification. We also tested other machine learning methods, such as Extremely Random Trees and artificial neural networks finding comparable or lower accuracy. Thus, the accuracy of classification with Equation 1 is comparable with the accuracy of classification with machine learning methods.

3.3 Distinct chemical properties of accreted and in-situ GCs

Figure 3 tests our accreted/in-situ classification based on the position in the E, L_z space on the Galactic GCs. Left panel of the Figure shows positions of all GCs with measured orbital properties as small coloured dots. Orange marks the accreted GCs and blue the in-situ ones (including Aurora). Out of 149 Galactic GCs considered, 98, or $\approx 2/3$ are classified as in-situ. GCs with sufficient APOGEE measurements are shown as large filled circles coloured according to their classification. There are 25 in-situ and 13 accreted GCs in our APOGEE sample. For further analysis we only consider GCs with sufficient number of measurements. For example, at least 3 GC member stars are required when studying GC properties reported in this Section; when the high-$[N/O]$ stars are concerned, at least 10 candidate members are required as well as the relative uncertainty on the fraction of the high-$[N/O]$ stars $f_{\rm N/O}$
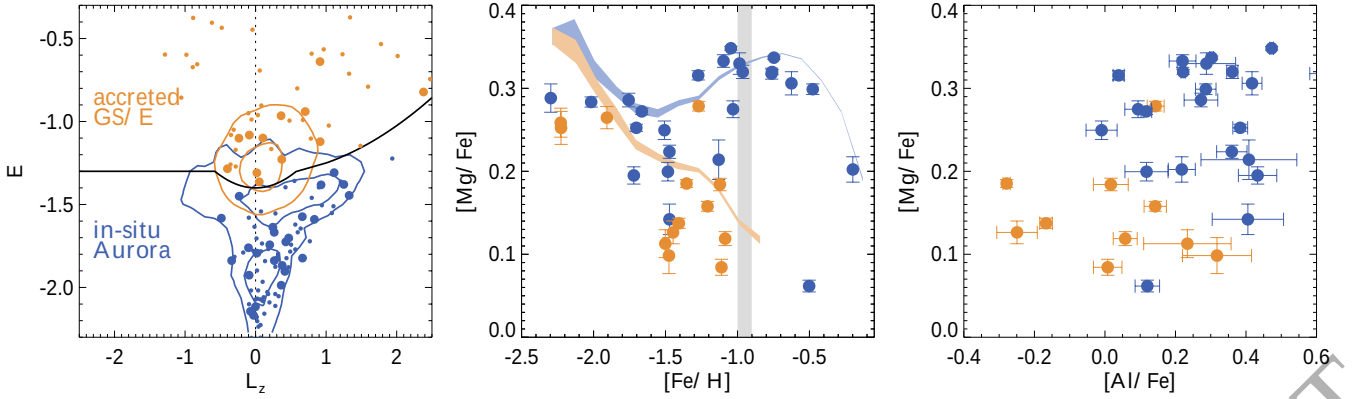
Figure 3. Properties of globular clusters classified as accreted (orange) and in-situ (including Aurora, blue). **Left:** Distribution of all Galactic globular clusters (small points) in the space of E and L_z (similar to left panel of Figure 2. Accreted (in-situ, including Aurora) GCs with high-quality $f_{\rm N/O}$ measurements are shown as orange (blue) filled circles. Blue (orange) contours show density of in-situ (accreted) stars from the left panel of Figure 2. **Middle:** Median [Mg/Fe] abundances with associated uncertainties as a function of cluster's [Fe/H] for accreted and Aurora GCs, classified as shown in the Left panel. Only GCs with median Mg uncertainties less than 0.025 dex are shown. Note that GCs classified as in-situ have systematically higher [Mg/Fe] compared to those classified as accreted. Orange (blue) bands show median stellar [Mg/Fe] abundance ratios (with associated uncertainties) for field stars as a function of metallicity. These stars have been classified using the same E, L_z boundary shown in the left panel. Note a clear peak in [Mg/Fe] of both in-situ GCs and in-situ field stars around [Fe/H]≈ -1 (the Spin-up, marked by a grey vertical band). **Right:** Orange (blue) points show median [Mg/Fe] as a function of median [Al/Fe] for stars in the Galactic GCs observed by APOGEE and classified as accreted (in-situ).

(selected according to the conditions stipulated in Section 3.1), less than 50%. The latter combination of cuts leaves only 28 clusters, of which 11 are classified as accreted and 17 as in-situ.

The middle panel of Figure 3 shows the median [Mg/Fe] ratios (with associated uncertainties) for the accreted and in-situ GCs as a function of metallicity. Note that in this Figure we show the median metallicity of the cluster's APOGEE member stars. There is little overlap between the two groups. The accreted GCs typically have median [Mg/Fe] ratios lower by some 0.15 dex than those of the in-situ clusters. The [Mg/Fe] values in Galactic GCs can be compared to the halo stars across the same metallicity range separated into accreted (orange band) and in-situ (blue band) using the same E, L_z criteria. Reassuringly, halo stars follow the same trends in the space of [Mg/Fe] and [Fe/H], in particular, at [Fe/H]> -1.5, median [Mg/Fe] for in-situ stars is ~ 0.15 dex higher than that for the accreted population (a similar conclusion is reached in Horta et al. 2020).

There is a striking trend of [Mg/Fe] with increasing metallicity exhibited by both field in-situ stars and the MW in-situ GCs in the middle panel of Figure 3. At [Fe/H]$\lesssim -1.5$ the [Mg/Fe] ratios of the accreted stars and GCs are similar, while at larger metallicities they separate. The trends exhibited by the field stars and GCs are broadly similar. This is consistent with model results of Hughes et al. (2020), which predicted that $[\alpha/\mathrm{Fe}]$–[Fe/H] distribution of GCs largely follows that of the field stars.

At [Fe/H]≈ -1.3, [Mg/Fe] starts to rise and exhibits a broad peak at [Fe/H]$\approx -1 \div -0.7$. The magnitude of this [Mg/Fe] increase is modest, just under 0.1 dex. Beyond [Fe/H]≈ -0.7, the [Mg/Fe] ratio is decreasing. As traced by field stars, this [Mg/Fe] peak can also be seen in Figure 7 of BK22 (see also the left panel of Figure 5 below) and is discussed at length in Conroy et al. (2022). Here for the first time, we show that the same pattern is displayed by the Galactic in-situ GCs.

As elucidated in Weinberg et al. (2017), bumps in α-element ratios of the order of $0.1 - 0.3$ dex are a tell-tale sign of a star-formation burst in a system converting a sizeable portion of its gas into stars, whilst retaining the core-collapse supernova enrichment

products. α-bumps similar to that reported here for the Galactic in-situ stars have been seen in several massive MW dwarf satellites by Hasselquist et al. (2021). Conroy et al. (2022) suggest that the MW in-situ α-bump can be explained by models in which the star formation efficiency increases sharply by a factor of $\sim$10 around [Fe/H] ≈ -1. As the middle panel of Figure 3 demonstrates, below [Fe/H] $= -1.8$, both the in-situ and the accreted stars show a systematic and noticeable increase in [Mg/Fe]. We believe that this [Mg/Fe] rise at low [Fe/H] may not be a genuine characteristic of the Galactic field stars but rather a sign of the APOGEE pipeline struggling with measurements of low and intermediate α abundances at low [Fe/H]. This hypothesis is based on the following tests: i) the media GC values show a less pronounced rise compared to the field stars, ii) the increases at low metallicities is reduced if no cuts are applied on abundance uncertainties or other effective signal-to-noise filters, and iii) [Si/Fe] shows a much flatter behaviour at [Fe/H]< -1.5.

Right panel of Figure 3 gives the distribution of the in-situ (blue) and accreted (orange) GCs in the space of median [Al/Fe] and [Mg/Fe]. GCs with [Fe/H]< -1.8 are excluded due to the suspected bias in [Mg/Fe] measurements mentioned above. The in-situ and accreted GC populations occupy distinct regions of [Mg/Fe]-[Al/Fe] space and show little overlap. On average, the in-situ GC have higher values of [Al/Fe]. At fixed [Al/Fe], the in-situ GCs have higher [Mg/Fe] ratios. The only two objects that buck this trend are NGC 6388 (classified as in-situ, blue point at [Mg/Fe] ≈ 0) and NGC 288 (classified as accreted, orange point with [Mg/Fe] > 0.25). The peculiar properties of these two clusters have been noted before and are discussed in the literature (see e.g. Myeong et al. 2019; Massari et al. 2019; Horta et al. 2020). Most recently, Carretta & Bragaglia (2022) argued for the in-situ origin of NGC 6388 based on the abundance pattern of iron-peak elements. For NGC 288, Monty et al. (2023) show that it does not follow the chemical trends of other GS/E GCs.

As revealed by GCs' detailed chemical properties, the misclassification rate is at a level similar to that indicated by our experiments with the GBT ML method. Additionally, we note that includ-
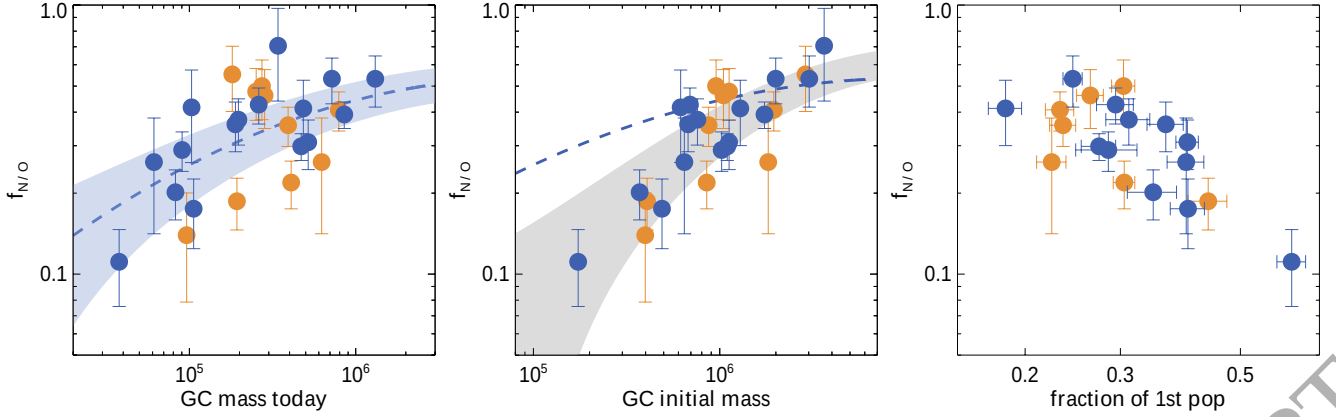
Figure 4. Fraction of high-[N/O] stars $f_{N/O}$ in the Milky Way globular clusters. **Left:** $f_{N/O}$ as a function of the present-day GC mass. Blue dashed line and band show quadratic fit to measurements of the in-situ GCs. **Middle:** $f_{N/O}$ as a function of the GC's initial mass, as estimated by Baumgardt & Makino (2003). The dashed blue line shows the quadratic fit to the correlation with present-day mass in the left panel. The grey band shows an approximate quadratic fit to all GCs shown in this panel. **Right:** $f_{N/O}$ as a function of the fraction of the GC's 1st population stars, as measured by Milone et al. (2017).

ing a rotating bar can and will affect some of the GCs' orbits and consequently might change their membership in the accreted/in-situ groups. However, when the effect of the bar is included, only 3 out of 40 clusters considered in the study of Pérez-Villegas et al. (2020) have been singled out as potential outer halo interlopers.

Notwithstanding these intricacies, the middle and right panels of Figure 3 show that the simple E, L_z-based classification of stars and GCs into accreted (primarily GS/E) and in-situ (primarily Aurora) using boundary defined by equation 1 works well. In particular, this classification results in populations of GCs with distinct chemical properties: the in-situ GCs exhibit systematically higher Mg abundance and the two GC classes occupy distinct regions in the [Al/Fe] − [Mg/Fe] space.

3.4 Trends of high-N fraction with cluster properties

Figure 4 explores the behaviour of the fraction of high-[N/O] stars in the Galactic GCs observed by APOGEE DR17. We have excluded NGC 6715 (M 54) residing in the core of the Sgr dwarf galaxy, as well as NGC 5139 (ω Cen). The left panel of the Figure shows $f_{N/O}$ as a function of the cluster's present-day mass. While there is a considerable scatter, there is a clear trend of $f_{N/O}$ increasing with increasing cluster mass, from $\approx 10\%$ in $M \approx 10^5 M_\odot$ clusters to $f_{N/O} \approx 50\%$ in clusters of $M \approx 10^6 M_\odot$. There is also an indication that the trend is more pronounced for the in-situ GCs (shown in blue).

A second-degree polynomial fit (± 0.075) to the in-situ $f_{N/O}$ values is shown as a blue band. The middle panel of the Figure shows that there is also a trend of increasing $f_{N/O}$ with the GC's initial mass. The dependence on the initial mass appears tighter and stronger compared to the correlation with the present-day mass shown in the left panel. The quadratic fit to all of the available data is shown as a grey band which includes ± 0.075 scatter around the mean model.

Finally, the right panel of Figure 4 shows that $f_{N/O}$ decreases with decreasing fraction of the 1st stellar population in clusters (where available, as measured by Milone et al. 2017). There is no obvious difference between in-situ and accreted GCs in terms of the trends of $f_{N/O}$ as a function of GC mass or fraction of 1st population in agreement with results of Milone et al. (2020).

The trends shown in Figure 4 are consistent with those pre-

viously reported in the literature (see e.g. Bastian & Lardo 2018; Gratton et al. 2019). Note however that the above dependence on the initial GC mass allows us to calibrate the $f_{N/O}$ (computed, as many previous studies, according to a somewhat arbitrary threshold) and thus link it directly to the Galaxy's original cluster population.

3.5 The majority of high-[N/O] field stars belong to Aurora

Left panel of Figure 5 shows metallicity distributions of in-situ (blue), accreted (orage), and high-[N/O] stars (black histogram). The high-[N/O] sample shows a sharp truncation at [Fe/H] ≈ -1, a broad peak at $-1.5 \lesssim$ Fe/H $\lesssim -1$ and a considerable drop off towards [Fe/H] ≈ -2. All populations, i.e. in-situ, accreted and high-[N/O] stars have approximately the same slope of the metallicity distribution at [Fe/H] < -1 (see discussion in Rix et al. 2022). At [Fe/H] ≈ -1, the in-situ distribution shows a sharp upward inflection matching the location of the cliff in the [Fe/H] distribution of the high-[N/O] stars (see also Figure 4 in BK22). At [Fe/H] ≈ -1, the slope of the in-situ metallicity distribution changes abruptly from $d \log(n)/d[Fe/H] \approx 1$ to $d \log(n)/d[Fe/H] \approx 2.3$. The metallicity distribution of the accreted stars does not show any obvious features around [Fe/H] ≈ -1, instead it exhibits a continuous downward slope from [Fe/H] ≈ -1.3 to [Fe/H] ≈ -0.7 where it is sharply truncated (in agreement with other studies, see, e.g, Belokurov et al. 2018; Mackereth et al. 2019; Belokurov et al. 2020; Feuillet et al. 2020; Sanders et al. 2021).

The middle panel of Figure 5 shows the distribution of the high-[N/O] stars in the $E -$ [Fe/H] space. Apart from the sharp truncation in the number of high-[N/O] stars at [Fe/H] ≈ -1 at all energies, there does not seem to be any strong correlation between E and [Fe/H]. Note, however, that at [Fe/H] < -1.7 the selected high-[N/O] stars do not span the entire range of energies of the high-metallicity counterparts. It is difficult to assess whether this is a sign of a genuine change in the properties of the high-[N/O] population or is simply a result of small statistics of stars in our sample.

The right panel of Figure 5 shows the $E - L_z$ distribution of the GC-like high-[N/O] stars selected as described in Section 3.1. The absolute majority of the stars fall within the blue contours
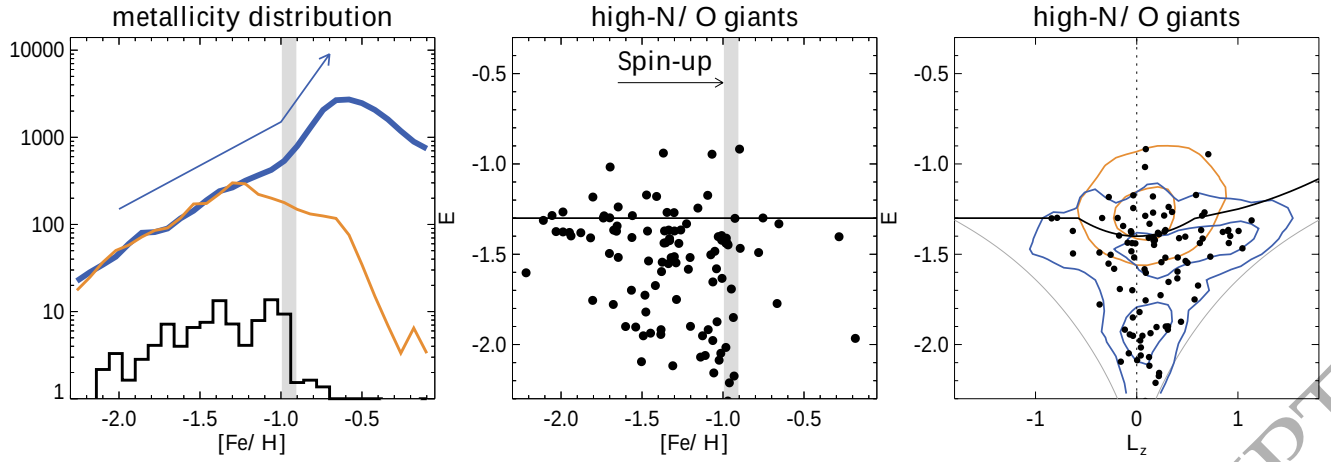
Figure 5. Left: Metallicity distribution of high-[N/O] field giants (black) is compared to the metallicity distribution of field giants classified as accreted (orange) and in-situ (blue). Only stars with low azimuthal velocities are used. Around [Fe/H] ≈ -1 (Spin-up, marked by the vertical grey band) there is a sharp upward bend in the slope of the distribution of the in-situ field stars (as indicated by the blue arrow) and a sharp truncation in the metallicity distribution of the high-[N/O] stars. **Middle:** Energy of high-[N/O] stars as a function of metallicity. The horizontal black line shows the approximate energy threshold below which the majority of stars are of in-situ nature. **Right:** E, L_z distribution of high-[N/O] stars (black dots). Blue (orange) contours show the density of in-situ (accreted) stars from the left panel of Figure 2.

Figure 6. Fraction of Aurora (accreted) stars shown in blue (orange) and fraction of high-[N/O] GC-like giants (multiplied by 50) shown in black as a function of energy. The accreted, in-situ and high-[N/O] samples are the same as shown in Figure 2. Note the similarity of $f_{N/O}$ and f_{Aurora} dependencies on energy. For comparison, dashed grey line gives the energy distribution of high-[N/O] stars with low-[Al/Fe], i.e. those likely not born in GCs.

marking the distribution of the in-situ stars. This indicates that most of high-[N/O] stars belong to the Aurora population of low-metallicity in-situ stars.

This is further clarified in Figure 6, which shows the fraction of the Aurora, accreted, and high-[N/O] stars as a function of their total energy E. Note that at low energies a few stars designated as accreted are likely misclassified in-situ stars. Due to the sharp

rise in [Al/Fe] with increasing [Fe/H], some of the low-metallicity Aurora stars drop below the nominal threshold. In the intermediate range of $-1.4 < E < -1.3$, the typical contribution of Aurora stars is $\approx 45\%$. The fraction of high-[N/O] stars is always low, typically $\sim 1\%$. Figure 6 shows that in the metallicity range considered $\approx 90\%$ of stars with $E > -1.2$ belong to the accreted population.

The striking fact shown in Figure 6 is that trends of $f_{N/O}$ and f_{Aurora} with energy are remarkably similar. As the energy decreases, the fraction of in-situ stars grows sharply, reaching $\approx 80\%$ at $E \approx -1.5$. This is a clear indication that these stellar populations are closely related.

3.6 Fraction of field high-[N/O] stars as a function of [Fe/H]

Figure 7 shows the fraction of high-[N/O] field stars in the Milky Way as a function of metallicity. At [Fe/H] < -1, the fraction is flat at approximately $1\% < f_{N/O} < 2\%$. In contrast, at [Fe/H] ~ -0.9, $f_{N/O}$ drops abruptly, decreasing with increasing [Fe/H] by an order of magnitude. This dramatic change in the $f_{N/O}$ with metallicity mirrors chemo-kinematic trends observed in the in-situ stars in BK22. They associate a rapid change in the chemical abundance spreads and the overall spin of the Galaxy's stars with the Milky Way transitioning from a chaotic pre-disk state into a coherently rotating disk.

Indeed, Figure 8 shows that the sharp increase in $f_{N/O}$ is due entirely to the stars in the Aurora population, while $f_{N/O}$ of the accreted (primarily GS/E) population does not change significantly with metallicity. Moreover, the in-situ population (Aurora) shows ≈ 5 times higher fraction of high-[N/O] stars compared to that of the GS/E. This increase somewhat depends on the selection method of the Aurora population (using E, L_z or [Al/Fe]), as shown in the right panel of Figure 7. Reassuringly, the two methods (E, L_z-based and [Al/Fe]-based) yield $f_{N/O}$ curves that are similar across a wide range of metallicity, i.e for [Fe/H] > -1.3. Around [Fe/H] ≈ -1.4 the [Al/Fe] ratio of the in-situ stars starts to drop below the chosen threshold (see Figure 2 in BK22) and, as a consequence, the [Al/Fe]-based in-situ $f_{N/O}$ is biased high compared
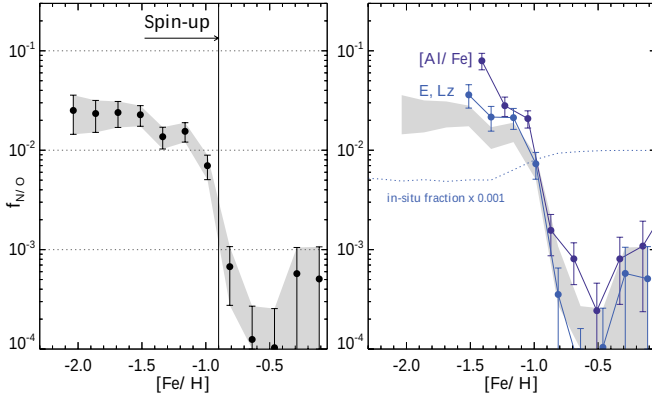
Figure 7. Left: Fraction of high-[N/O] stars $f_{N/O}$ with low tangential velocities as a function of metallicity. Stars in globular clusters and Magellanic Clouds have been excluded. Note the abrupt jump in $f_{N/O}$ around [Fe/H] ≈ -0.9. **Right:** Comparison of the in-situ $f_{N/O}$ ratio computed with two different methods of in-situ classification. Blue (purple) shows classification based on (E, L_z) (ratio of [Al/Fe]). The two estimates agree for [Fe/H] > -1.4. Around [Fe/H] ~ -1.4 the [Al/Fe] ratio of the Aurora population drops quickly under the nominal threshold used to separate the accreted and the in-situ stars. As a result, the $f_{N/O}$ ratio artificially increases. Blue dotted line shows the re-scaled in-situ fraction, note the drop from 100% to 50% at $-1.5 <$[Fe/H]< -1.

to the E, L_z one at this metallicity. However, these differences are much smaller than the overall trend itself and can be viewed as uncertainty with which $f_{N/O}$ is estimated for the Aurora population.

The sharp increase of the fraction of high-[N/O] Aurora stars $f_{N/O}$ with decreasing [Fe/H] shows that processes that produced anomalous amounts of nitrogen in the regions where high-[N/O] stars were born were much more prevalent at lower metallicities. If massive star clusters are plausible sites of high-[N/O] star formation, we can expect that the metallicity distribution of the high-[N/O] stars is similar to that of the globular clusters.

In principle, if high-[N/O] were produced in clusters, their metallicity distribution should trace distribution of metallicity of the disrupted clusters. Moreover, given that $f_{N/O}$ is higher in larger mass clusters (Figure 4) and if the fraction of massive had a dependence on metallicity, the metallicity distribution of field high-[N/O] stars could be somewhat different from the metallicity distribution of clusters. However, models indicate that the fraction of disrupted clusters is nearly independent of metallicity (O. Gnedin, priv. communication) and there is no significant trend of cluster mass distribution with metallicity for observed surviving GCs (see, however, results of the E-MOSAIC model Pfeffer et al. 2023, which do predict that disruption is metallicity dependent). We note that the fraction of mass lost by surviving MW GCs, as estimated by Baumgardt & Makino (2003), shows no trend with metallicity which is consistent with this assumption. If this assumption is correct, we can expect the metallicity distribution of high-[N/O] stars and *surviving* clusters to be quite similar.

Indeed, the upper right panel in Figure 9 shows that cumulative metallicity distributions of the in-situ high-[N/O] stars and in-situ MW's globular clusters are quite similar for stars and GCs with [Fe/H] < -1 and differ at larger metallicities. For this comparison we applied a cut of V_ϕ to the sample of GCs to account for a similar cut used in selecting high-[N/O] stats. The similarity of metallicity distributions is confirmed by the Kolmogorov-Smirnov (KS) probability that the two distribution agree at [Fe/H] < -1.
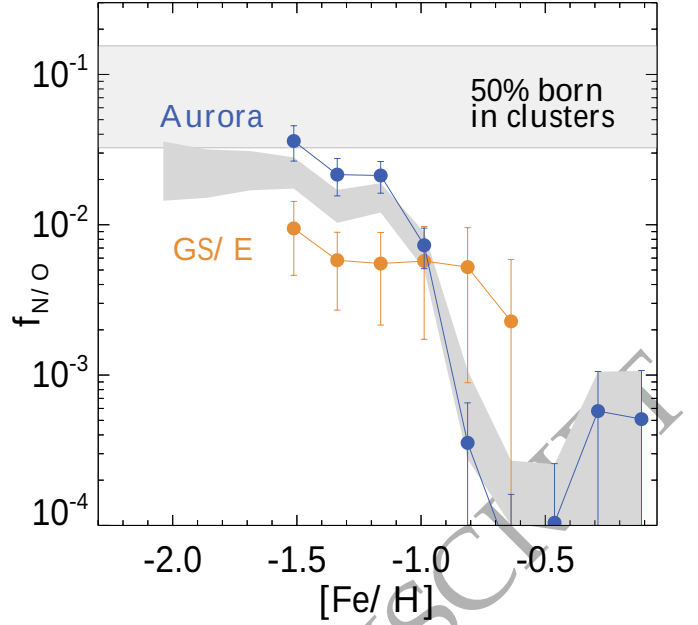

Figure 8. Grey band shows $f_{N/O}$ measurements reported in Figure 7. Blue (orange) line and symbols show measurements for in-situ (accreted) populations classified using energy boundary shown in Figure 2. Below [Fe/H]=-0.9, Aurora's $f_{N/O}$ is > 3 times higher than that for the accreted stars. Above [Fe/H]=-0.9, Aurora's $f_{N/O}$ plummets by more than an order of magnitude, while the behaviour of the accreted population population is much flatter. At [Fe/H]< -1, Aurora's $f_{N/O} = 0.04$, which is just under a half of the typical fraction in our fiducial model (see text for details), implying that close to a half of the Aurora stars (at this metallicity) were born in clusters.

The 95% range of this probability evaluated using bootstrap resampling is extending to probability of 0.21. In contrast, the metallicity distribution of the in-situ high-[N/O] stars is different from that of the ex-situ GCs and the 95% range of the KS probability extends only to 0.01 (with the mean value of $\approx 10^{-3}$).

The middle and right panels of Figure 9 show a comparison of the cumulative distributions of total energy and z-component of the angular momentum of the high-[N/O] stars and GCs with metallicities [Fe/H] < -1. It clearly shows that distributions of the energy and angular momentum of the in-situ stars and GCs agree, while in-situ stars are inconsistent with the distributions of ex-situ GCs both in metallicity and energy.

We also find that distributions of metallicity, energy, and angular momentum of the ex-situ high-[N/O] stars and ex-situ GCs are consistent. The uncertainties in this comparison are larger, however, as there are only 25 ex-situ high-[N/O] stars in our sample. Still, distribution of ex-situ stars and in-situ stars is strongly inconsistent in energy (probability to be drawn from the same distribution is 0.0), which is not surprising because boundary between in-situ and ex-situ objects indicated by [Al/Fe] corresponds to a rather sharp boundary in energy.

Our results appear to be inconsistent with results of a recent study by Kim et al. (2023) who also find that N-rich stars have metallicity distribution similar to that of GCs at [Fe/H] < -1, but who argued that distributions of orbital parameters of these stars and clusters are not the same. We note, however, that in their Figure 4 only distributions of the pericenters (r_{min}) are somewhat different. Furthermore, if similarity exists only for the low-metallicity
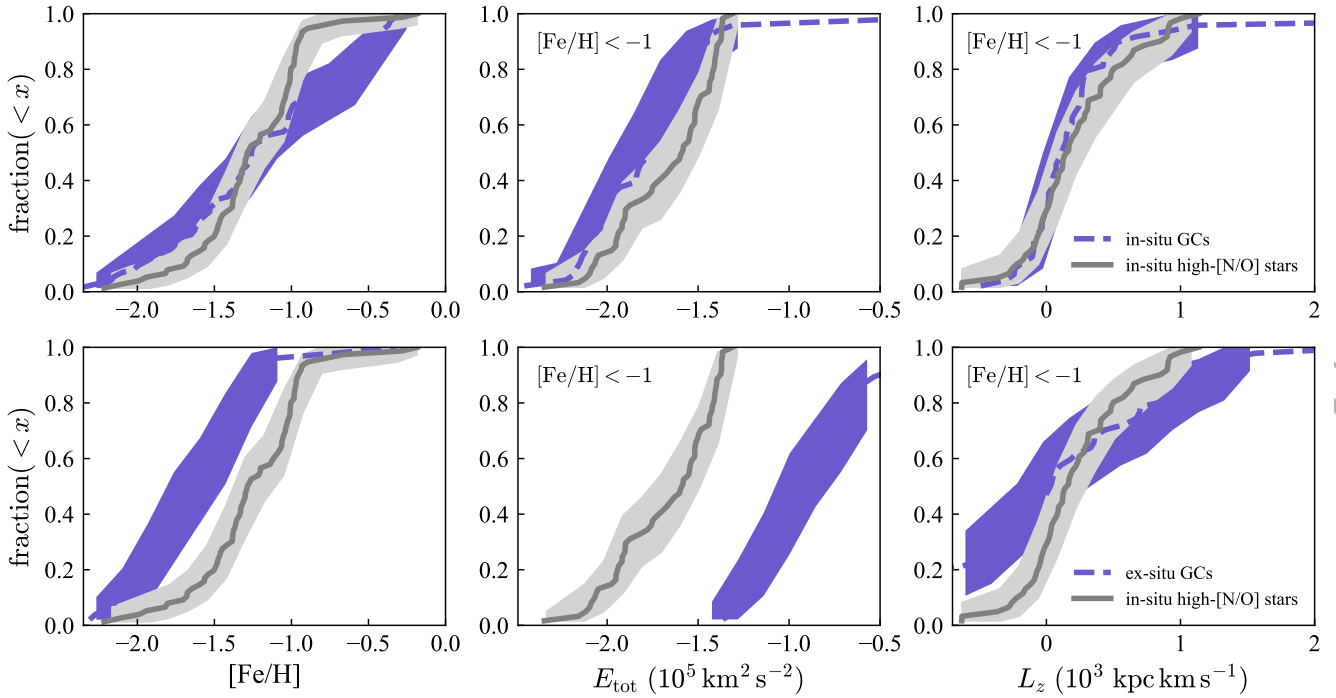
Figure 9. Cumulative distributions of metallicity ([Fe/H]), energy and L_z component of the angular momentum of the high-[N/O] stars (gray solid lines and shaded bands) and MW globular clusters (blue dashed lines and shaded bands). To account for the velocity cut $V_\phi < 160$ km/s for the stars, we applied the same cut to GCs for this comparison. The gray and blue shaded bands show corresponding 95% range estimated using bootstrap resampling. The top (bottom) row compares distributions of the in-situ high-[N/O] stars with in-situ (ex-situ) GCs. Note that [Fe/H] distributions are compared for the full range of metallicity, while $E_{\rm tot}$ and L_z distributions are compared for stars and GCs in the metallicity range where metallicity distributions of in-situ stars and GCs agree (upper left panel): [Fe/H] < −1. The figure shows that distributions of metallicity, energy, and momentum of the in-situ GCs and in-situ high-[N/O] stars are similar in this metallicity range, which indicates that these stars originated from low-metallicity GCs formed in the Milky Way.

([Fe/H] < −1) regime and the goal is to test whether similarity exists also in the orbital properties of stars and clusters, it would be more conclusive to do a comparison of orbital properties in that regime, as we do here, rather than for the entire sample.

Overall, our results presented in Figure 9 shows that distributions of metallicity, energy, and momentum of the in-situ GCs and in-situ high-[N/O] stars with [Fe/H] < −1 are similar, *which strongly indicates that low-metallicity in-situ globular clusters are the primary formation sites of these stars.*

If we assume that this is the case, the observed fraction of high-[N/O] stars in the Aurora population allows us to estimate the fraction of stars born in clusters during early, pre-disk stages of MW evolution. We present such an estimate in the next subsection.

3.7 The fraction of the Aurora stars born in clusters

The correlation between the fraction of high-[N/O] stars and the initial GC mass shown in the middle panel of Figure 4 and estimates of the $f_{\rm N/O}$ at different metallicities (Figure 8) can be used to compute $f_{\rm N/O,cl}$ – the fraction of stars born in clusters that have high N/O for a population of newly born GCs in high-z MW progenitor. This, in turn, can be used to estimate the fraction of stellar mass born in clusters in the Aurora population.

To this end, we adopt a model for the mass function of clusters at birth that follows the Schechter functional form (Schechter 1976): $dN/dM \propto M^{-\beta} e^{-M/M_c}$. We choose the cut-off mass of $M_c = 5 \times 10^6 M_\odot$ motivated by the analysis of Choksi & Gnedin

(2019) and observational estimates (see Fig. 17 in Adamo et al. 2020b), as well as $\beta = 2$, the average slope of the mass function of young compact star clusters measured in observed galaxies at $z = 0$ (see Krumholz et al. 2019, for a review), and assume that clusters are born with masses in the range $10^3 − 5 \times 10^7 M_\odot$ (see, e.g., Norris et al. 2019, for an observational justification of the upper value). The resulting estimate using such a model gives the fraction of high-[N/O] stars born in clusters $f_{\rm N/O,cl} \approx 14\%$. Increasing characteristic mass to $M_c = 10^7 M_\odot$ and the minimal cluster mass to $10^4 M_\odot$ while varying the $f_{\rm N/O} − M_{\rm ini}$ within the grey band shown in the middle panel of Figure 4 gives a range of typical high-[N/O] fractions $6\% < f_{\rm N/O,cl} < 32\%$.

Given this estimated range and the observed fraction of high-[N/O] stars in the Aurora population, we can estimate the fraction of the Aurora stars born in clusters as $f_{\rm cl} \approx f_{\rm N/O}/f_{\rm N/O,cl}$. The MW in-situ stars with metallicities of [Fe/H] $\approx −1.5$ exhibit $f_{\rm N/O} \approx 4\%$. Thus, these fractions indicate that the fraction of the Aurora stars born in clusters ranges from $\approx 15\%$ to $\approx 70\%$. The range corresponding to 50% of all stars born in clusters is shown as a horizontal gray band in the middle panel of Figure 8.

Aurora stars of higher metallicity [Fe/H] $\approx −1$ have $f_{\rm N/O} \approx 1\%$ and fractions of stars born in clusters is correspondingly four times smaller: $\approx 4 − 17\%$. These estimates can be compared to the GS/E progenitor, where a considerably smaller fraction of $2 − 10\%$ is estimated to be born in clusters independent of metallicity. At larger metallicities the implied fraction of stars born in clusters drops to smaller values, which is broadly consistent with predic-

tions of the E-MOSAIC GC formation model presented by Hughes et al. (2020).

Note that our estimate of $f_{N/O,cl}$ implicitly assumes that all of the stellar mass initially in GCs is now in the field. This is supported by modelling results of Rodriguez et al. (2023, see their Fig. 3) and Gieles & Gnedin (2023, see their Fig. 13) which indicate that the majority of the clusters contributing large numbers of high-[N/O] stars disrupt by $z = 0$. Furthermore, according to the initial mass estimates for the surviving in-situ clusters by Baumgardt & Makino (2003), that population lost $\approx 84\%$ of its initial combined mass. Thus, within the uncertainties of our estimate, the fraction of stellar mass born in GCs that is still in surviving clusters is negligible.

3.8 Estimate of mass of the Aurora stellar population

The estimate of mass fraction of the in-situ stars born in clusters presented above can be used to estimate the stellar mass of the Aurora population – the in-situ stars with metallicities $[Fe/H] \lesssim -1$. This is useful because MW globular cluster propulation is more or less complete, while Aurora stars are observed over a limited volume in a sample with a fairly uncertain selection function.

Indeed, this mass can be estimated as

$$M_{Aur} = \frac{1}{1 - f_{disrupt}} \sum_{i=1}^{N_{GC,ins}} \frac{M_{i,ini}}{f_{i,cl}}, \qquad (2)$$

where $f_{disrupt}$ is the fraction of the initial clusters that were disrupted, $N_{GC,ins}$ is the number of observed in-situ clusters with metallicity $[Fe/H] < -1$, $f_{i,cl}([Fe/H])$ is the fraction of star formation that occurred in bound star clusters at metallicity $[Fe/H]_i$, and $M_{i,ini}$ is the initial mass of the observed cluster i, respectively. We use the initial GC mass estimates from Baumgardt & Makino (2003) and approximation to $f_{cl}([Fe/H]) = f_{N/O}([Fe/H])/0.14$, where $f_{N/O}([Fe/H])$ is approximation to the trend shown in Figure 7:

$$f_{N/O} = \frac{a}{(1 + x^b)^{g/b}}, \qquad (3)$$

where $x = 10^{[Fe/H]+1.2}$, $a = 3 \times 10^{-2}$, $b = 6$, $g = 3.25$.

Using eqs 2 and 3 and our classification of the in-situ clusters in the Baumgardt & Vasiliev (2021b) catalog, we estimate stellar mass of the Aurora population $M_{Aur} \approx 4.6 \times 10^8 M_\odot$. The uncertainty of this estimate is due to uncertainty in $f_{cl}([Fe/H])$, the estimate of the fraction of high-[N/O] stars in clusters, the uncertainty in the estimate of initial cluster masses and uncertainty in the fraction of initial globular cluster population that was disrupted before $z = 0$. A rough estimate of the first uncertainty can be obtained by changing the value of constant a in equation 3 from 2×10^{-2} to 4.5×10^{-2} to roughly encompass the uncertainties of $f_{N/O}$ measurements in Figure 7. This gives estimates of the Aurora population mass ranging from $3 \times 10^8 M_\odot$ to $6.9 \times 10^8 M_\odot$. Thus, the mass of this population is approximately

$$M_{Aur} = 5 \pm 2 \times 10^8 (1 - f_{disrupt})^{-1} M_\odot, \qquad (4)$$

although the uncertainty in this estimate only includes uncertainty in $f_{N/O}$ and not the other sources listed above. The uncertainty in the initial cluster mass estimate is a systematic uncertainty of this estimate. The unknown fraction of disrupted clusters is represented by the factor $1 - f_{disrupt}$ accounts for the fact that we use surviving clusters to make the estimate, while $f_{disrupt}$ fraction of the initial mass in clusters was returned to the field due to cluster disruption.
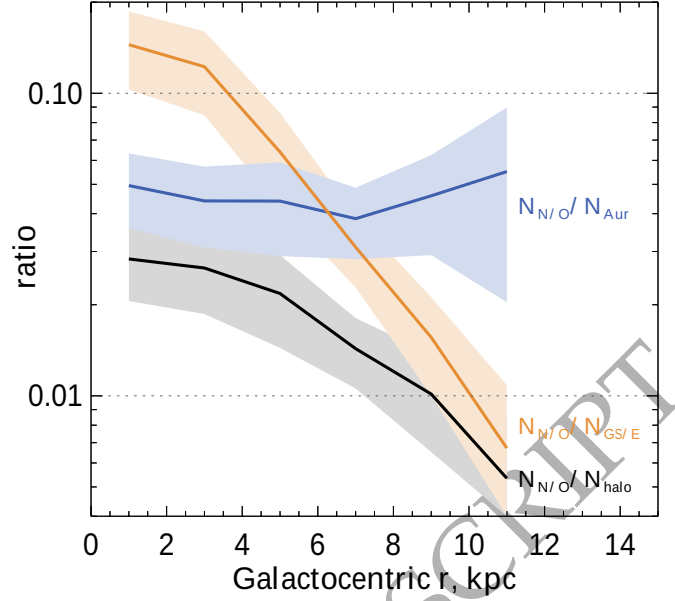

Figure 10. Ratios of the number of high-[N/O] stars to the number of stars in different Galactic components as a function of Galactocentric radius. Only stars with $-1.45 < [Fe/H] < -1$ are considered. Grey shows the ratio (and the associated uncertainty) relative to all halo stars as a function of r. The halo fraction of high-[N/O] stars increases by a factor of ≈ 6 from the Solar radius to the Galactic centre, in agreement with results of Horta et al. (2021b). The orange band shows the ratio of high-[N/O] stars to the accreted halo population (mainly GS/E) and exhibits a much steeper, a factor of ≈ 40, increase. This is because the GS/E debris does not populate the inner regions of the MW and thus has a shallower density profile compared to the halo overall and Aurora in particular. The blue band is the ratio relative to the Aurora stars. This ratio is flat at $\approx 4\%$, indicating that the high-[N/O] stars in this metallicity range belong to the Aurora population.

Discussion of the GC evolution in the previous section indicates that the fraction of disrupted clusters is $f_{disrupt} \approx 2/3$. The actual mass of the Aurora population is thus larger by the corresponding factor than the above estimate.

Notwithstanding the uncertainties, this estimate shows that the mass of the in-situ low-metallicity population of stars that were born before MW disk started to form and that have now a spheroidal distribution is comparable to the mass of the MW stellar halo of $\approx 1.4 \pm 0.4 \times 10^9 M_\odot$ (e.g., Deason et al. 2019). Given that this in-situ population is more concentrated towards the centre of the Galaxy than the overall halo (Rix et al. 2022; see comparison of the distance distributions the Aurora stars and halo stars in Figure 10), it is expected to dominate the spheroidal stellar component at $r \lesssim 10$ kpc.

By definition the Aurora population is the population of in-situ stars formed before the spin-up of the Milky Way disk. Its mass estimate above is thus also the estimate of the Milky Way progenitor's stellar mass at the time when disk started to form and indicates that this occurred when Milky Way's stellar mass was between the mass of the Small and Large Magellanic Clouds.

3.9 Radial distribution of high-[N/O] stars

Horta et al. (2021b) show that the fraction of the fraction of N-rich stars among halo stars increases by a factor of about six from

$r \approx 10$ kpc to $r = 2$ kpc. They put forward a hypothesis that such an enhancement of GC-born stars may be linked with an increase in accretion and disruption of dwarf galaxies in the high-z Milky Way. Note, however, that a much flatter trend out to ~ 40 kpc is reported (see Koch et al. 2019; Horta et al. 2021b).

Figure 10 explores the radial dependence of the fraction of high-[N/O] stars in the Galactic halo (black line, grey band). We also compute the ratio of the number of high-[N/O] stars to the number of stars in the GS/E ($N_{\mathrm{N/O}}/N_{\mathrm{GS/E}}$, orange) and Aurora ($N_{\mathrm{N/O}}/N_{\mathrm{Aur}}$, blue) components. Here we resort to the [Al/Fe]-based in-situ/accreted classification as the energy-based selection will induce a bias in the radial distribution. Therefore, the radial profiles shown in Figure 10 are limited to stars with $-1.4 <$[Fe/H]< -1.1, where [Al/Fe] selection is most reliable (see discussion above). As was shown in the right panel of Figure 7, the difference for samples selected by energy and [Al/Fe] is much smaller than the trends we are considering.

While the details of our selection of high-[N/O] stars are different to that in Horta et al. (2021b), the halo fraction in these stars shows a very similar trend with Galacto-centric radius to the one estimated in that study. As the black curve indicates, there are ~ 6 times more high-[N/O] GC-like halo stars in the centre of the Galaxy compared to their relative counts around the Sun. A much more dramatic radial change in the fraction of high-[N/O] stars can be seen when their numbers are compared to the counts of the accreted debris (composed mostly of the GS/E stars). Relative to the GS/E stars, the high-[N/O] population increases by a factor of ~ 25 from 10 kpc to 1 kpc.

This is in agreement with recent studies of the inner stellar halo which indicate that the fraction of the GS/E stars decreases within ~ 3 kpc from the Galactic centre (see Iorio & Belokurov 2021; Belokurov et al. 2023). The Galacto-centric radial distribution of the tidal debris from an accreted satellite depends on the mass of the dwarf and the merger time (see Deason et al. 2013; Horta et al. 2023b). Note that the GS/E progenitor galaxy was likely of sufficient mass to have experienced strong dynamical friction and profound loss of the orbital energy and angular momentum during its interaction with the Milky Way (see discussion in Vasiliev et al. 2022). Such rapid orbital radialization is often accompanied by "explosive" mass loss resulting in a complete disruption of the satellite before it manages to reach the centre of the Galaxy. The drop in relative density of the GS/E debris inward of the Solar radius can also be gleaned from Figure 2 where the number of orange points quickly decreases below the Sun's energy level.

Large changes in the relative fraction of high-[N/O] stars when compared to the stellar halo overall or to its accreted component signal one thing: high-[N/O] GC-like population is not a typical member of either of these. Instead, as shown by the blue curve in Figure 10, the radial distribution of the high-[N/O] stars is very similar to that of the Aurora stars and they are thus likely to be a component of the Aurora population. This is in complete agreement with the above comparisons between the trends of the numbers of high-[N/O] and Aurora stars as a function of energy (Figure 2) and metallicity (Figure 7).

Turning the argument around, the above discussion cements the view of the Aurora population as a centrally concentrated stellar halo component, as hypothesised by Belokurov & Kravtsov (2022) based on a dataset limited to the Solar vicinity. Such a radial density (strongly peaked at $r = 0$) and an extent (limited to approximately Solar radius) is in agreement with the results of numerical simulations of the Milky Way formation as illustrated in Figure 13 of BK22 and observed in the all-sky view of the metal-poor com-
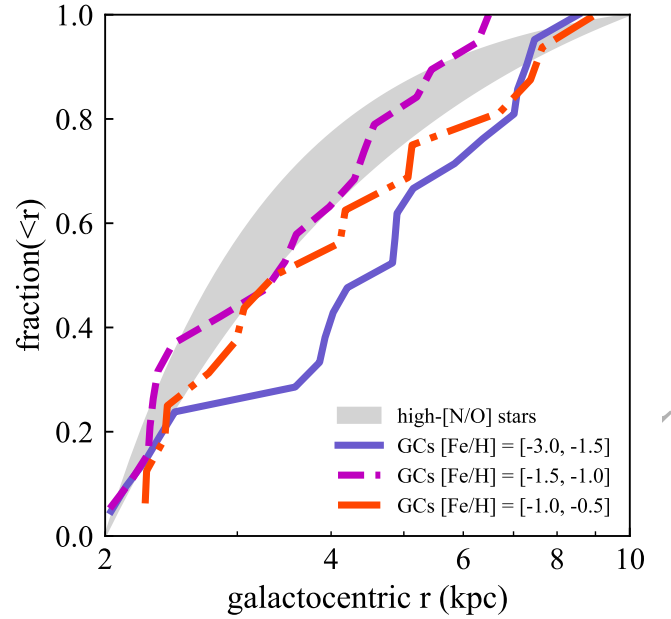

Figure 11. Cumulative distribution of the galactocentric distance for the MW in-situ GCs in three metallicity intervals (thick lines) and radial distribution of the N-rich stars from Horta et al. (2021b) with [Fe/H] < -1 shown by the grey band. The band represents the uncertainty in the slope of the derived profile of N-rich stars. Given that the distribution of the high-[N/O] stars was derived for the range $d_{\mathrm{GC}} = [2, 10]$ kpc only GCs in the same distance range are included in the cumulative distribution. The plot shows that distance distributions of the high-[N/O] stars and in-situ GCs with [Fe/H] > -1.5 are consistent indicating that these stars may have originated from the in-situ clusters in this metallicity range, which is also where metallicities of most high-[N/O] stars in the Horta et al. (2021b) sample lie.

ponent of the Galaxy (see Rix et al. 2022). These recent studies (see also Conroy et al. 2022; Myeong et al. 2022) together with the trends discussed here indicate that at low metallicities, the central stellar halo is dominated by the in-situ formed Aurora stars rather than by stars brought in by other accreted galaxies.

Given the indications that high-[N/O] stars originated in clusters that we discussed above, it is also instructive to compare the radial distribution of these stars with the radial distribution of in-situ globular clusters. Such comparison is shown in Figure 11 where the cumulative distribution of the galactocentric distance of high-[N/O] stars estimated by Horta et al. (2021b) is compared to cumulative distance distributions of the in-situ globular clusters in our classification in [Fe/H] ranges $[-3, -1.5]$, $[-1.5, -1]$ and $[-1, -0.5]$. Cumulative distributions for both clusters and high-[N/O] stars were constructed using objects with distances in the range $r \in [2 - 10]$ kpc because this was the distance range of the Horta et al. (2021b) sample.

The figure shows that the distance distribution of GCs with [Fe/H] > -1.5 is quite similar to that of the high-[N/O] stars. In particular, the distributions are closest for clusters with metallicities [Fe/H] $\in [-1.5, -1]$, which corresponds to the range of metallicities of most of the high-[N/O] stars. Recently, Gieles & Gnedin (2023) have also shown that the distance distribution of stars from disrupted massive GCs in their model matches the distribution of N-rich stars (see their Section 6.2 and Figure 12) and argued that
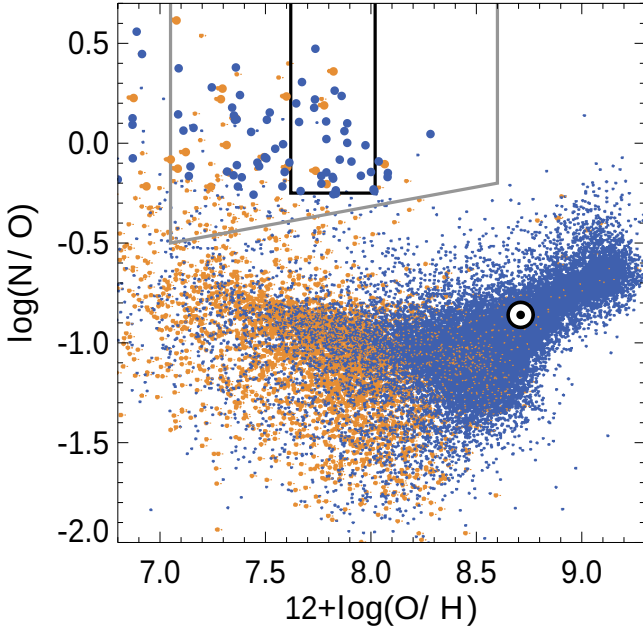
Figure 12. Nitrogen-to-oxygen $\log(O/H) = [N/O] + \log(N/O)_{\odot}$ abundance ratio as a function of oxygen abundance $12 + \log(O/H) = [O/H] + \log(O/H)_{\odot}$ for giants with low V_{ϕ} classified as accreted (orange) and those born in-situ (blue, including Aurora). Stars are classified into accreted and in-situ based on their energy E, L_z values. Larger filled circles are the high-[N/O] stars. Conservative (thin black lines) and fiducial (thick black lines) abundance measurements for GN-z11 from Cameron et al. (2023) are also shown. Solar $\log(N/O)_{\odot} = -0.86$ and $12 + \log(O/H) = 8.71$ is marked with $\odot$. Giants with high N/O ratio fall squarely within the range of measurements for GN-z11. Note that in accordance with Fig 2, at least 75% of the high-[N/O] stars belong to Aurora.

this is an indication that these stars originated in clusters. In principle, the distance distribution of the surviving clusters and disrupted clusters can be different, but our results indicate the difference is not large. The largest difference between distance distributions in Figure 11 is for in-situ GCs with [Fe/H] < -1.5, although the difference is not very significant as the number of clusters in this metallicity range is fairly small.

The similarity of metallicity, energy, and angular momentum distributions of high-[N/O] stars in clusters shown in Figure 9 and the similarity of distance distributions of these stars with the Aurora stars and in-situ GCs shown in Figures 10 and 11 strongly indicate that *the majority of high-[N/O] stars with* [Fe/H] < -1 *originated in globular clusters formed in-situ in the low-metallicity Aurora population.*

3.10 High-redhsift Milky Way vs GN-z11

Figure 12 presents the [N/O] abundance ratio as a function of the oxygen-based metallicity [O/H]. This view is similar to that shown in panel c of Figure 1, albeit here we show the absolute abundance ratios, i.e. not referenced to the solar values. Instead, the solar values are marked with the $\odot$ symbol. Below $12 + \log(O/H) \approx 8.2$ both accreted (orange) and in-situ (blue) populations are present. Above $12 + \log(O/H) > 8.2$ the stars are overwhelmingly of in-situ origin. Up to the Solar value of $12 + \log(O/H) = 8.69$, the N/O is almost flat with a small bend. Around the Solar metallicity,

the N/O abundances start to rise noticeably. Conservative (fiducial) measurements reported by Cameron et al. (2023) for the GN-z11 galaxy are shown with grey (black) lines.

The Figure shows a large number of APOGEE field giants with N/O and O/H abundance ratios similar to those reported for GN-z11. There is a total of 100 high-[N/O] low-metallicity ([Fe/H] < -0.7) stars marked with larger filled circles. Of these 73 are classified unambiguously as in-situ and 27 as accreted. Given our conclusion that the majority of the in-situ high-[N/O] stars originated in massive bound clusters and that the clusters, in turn, constituted a significant fraction of star formation at early epochs, the implication of these results is that a significant fraction of star formation in GN-z11 also occurs in bound compact star clusters.

The question, of course, is whether a significant fraction of the gas dominating the emission lines can be enriched by the 2P stars in clusters. To address such questions careful modelling is needed, which is beyond the scope of this paper. Broadly speaking, however, we think that GN-z11 emission can be arising in the core of the galaxy during a particular post-burst stage. If many clusters were born in that region in a strong star formation burst, the 1P and lower N/O stars would be born during the burst, but the observed gas can be largely 2P enriched. Given that high-z galaxies are small in mass, have much smaller sizes than galaxies at lower z ($\lesssim 100 - 500$ pc, e.g., Bouwens et al. 2021; Baggen et al. 2023), and are expected to have very bursty star formation histories, such a scenario seems plausible.

4 DISCUSSION

Results presented above strongly indicate that 1) the majority of low-metallicity high-[N/O] stars originated in compact bound star clusters and 2) up to $f_{cl} \approx 50\%$ of star formation at early epochs before MW formed its disk ([Fe/H] $\lesssim -1.5$) occurred in bound massive star clusters. This fraction decreases rapidly with increasing metallicity to $10 - 20\%$ at [Fe/H] ≈ -1 and to $\lesssim 1\%$ at [Fe/H] $\gtrsim -0.7$.

A number of recent studies have estimated the fraction of field N-enhanced stars in the Milky Way and showed it to be small, of the order of $\sim 2\%$ (e.g., Martell et al. 2016; Schiavon et al. 2017; Koch et al. 2019; Horta et al. 2021b), consistent with the measurements presented here. Some of these studies have also used the estimated fraction to derive the fraction of low-metallicity stars that formed in bound clusters, obtaining values $f_{cl} \approx 5 - 20\%$ (e.g., Martell et al. 2011; Koch et al. 2019; Horta et al. 2021b).

As we discussed in Section 1, these estimates involve the total stellar population without accounting for the contribution of in-situ and ex-situ stars. Thus, the estimated fraction is averaged over stars formed in the accreted dwarf galaxies and the main MW progenitor, which likely had very different environments and evolution. Interpretation of such an average is not straightforward. Furthermore, the estimated f_{cl} uses an uncertain factor to convert the estimated fraction of N-enhanced stars into the fraction of the first generation stars in clusters. Finally, the threshold used to select N-rich stars is somewhat arbitrary and is not calibrated on the real GCs, without correction to their initial mass, and without accounting for the GC mass function. In this study, we estimate f_{cl} rectifying all of these issues.

Direct estimates using observed globular cluster in the ultra-faint galaxy Eridanus II (Weisz et al. 2023) show that the cluster contributed up to $\approx 10\%$ of the galaxy stellar mass at birth, while Zick et al. (2018) show that $\approx 20\%$ of star formation in the Fornax
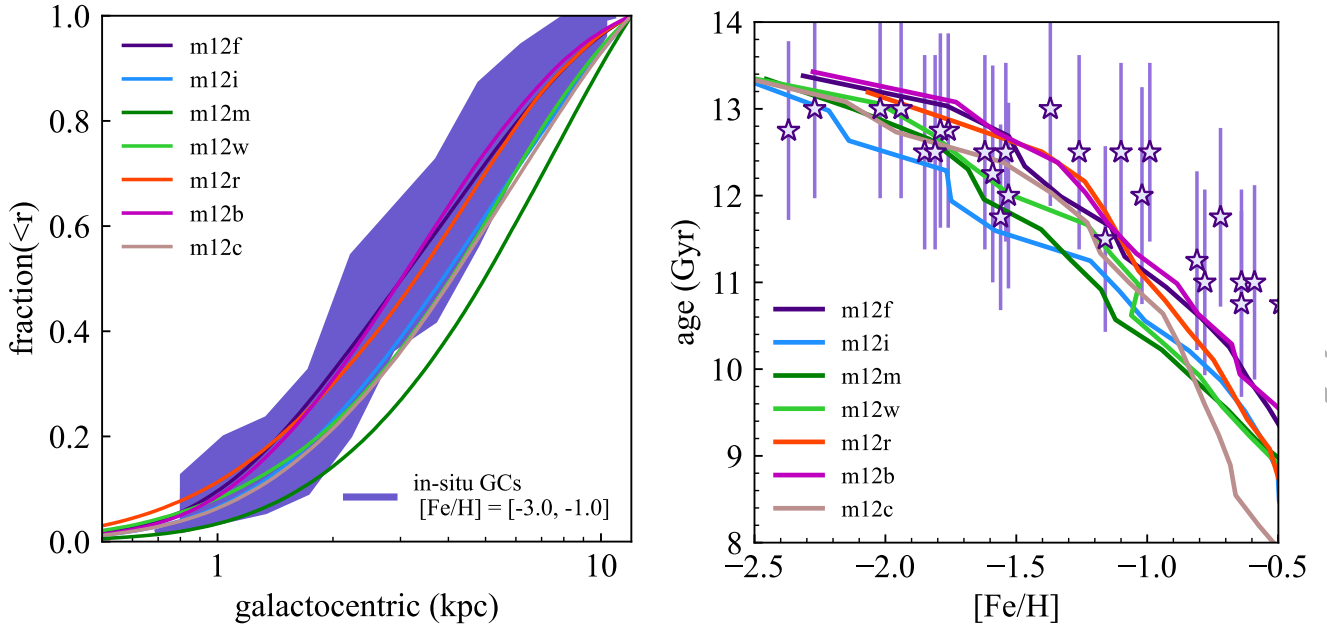
Figure 13. Left: Cumulative distribution of the galactocentric distance for the MW in-situ GCs (thick lines with shaded bands) in the metallicity interval $[\mathrm{Fe/H}] \in [-3, -1]$ compared to the distribution of galactocentric distance of the in-situ born stars in the FIRE simulations of the MW-sized galaxies in the same metallicity interval. The figure shows that the distribution of observed in-situ clusters and in-situ stars in a significant fraction of simulated galaxies agree indicating that MW GCs could form as part of the regular in-situ star formation in the early Milky Way. **Right:** age-metallicity relation for the in-situ stars in simulations and in-situ MW GCs.

dwarf occurred in bound *surviving* clusters. Similarly, we estimate that for the GS/E the fraction of stellar mass formed in GCs is $f_{\mathrm{cl}} \approx 0.06 \, f_{\mathrm{GC,GS/E}} \, M_{\star,8.7}^{-1}$, where $M_{\star,8.7}^{-1}$ is GS/E's stellar mass at the time of the merger in $5 \times 10^8 \, M_\odot$ and $f_{\mathrm{GC,GS/E}}$ is the fraction of MW's ex-situ clusters contributed by GS/E. However, $\approx 55\%$ of ex-situ GCs mass is in clusters with $[\mathrm{Fe/H}] \leqslant -1.5$, while only $\approx 15\%$ of GS/E stars are at these metallicities (see lower panel of Fig. 4 in BK22). If we assume that all ex-situ GCs were formed in GS/E, this corresponds to $f_{\mathrm{cl}} \approx 25\%$ in the GS/E at $[\mathrm{Fe/H}] \leqslant -1.5$.

We also see a direct record of GC formation in the linear relation between mass in globular cluster population in galaxies and mass of the parent halo, $M_{\mathrm{GCS}} = \eta_{\mathrm{GC}} M_{\mathrm{h}}$. Such relation is expected in the models of GC formation (Kravtsov & Gnedin 2005; Bekki et al. 2008; Choksi et al. 2018; Choksi & Gnedin 2019; El-Badry et al. 2019; Bastian et al. 2020) and is observed to hold over more than five orders of magnitude in galaxy stellar mass with $\eta_{\mathrm{GC}} \approx 3 \times 10^{-5}$ (see Spitler & Forbes 2009; Hudson et al. 2014; Harris et al. 2017; Forbes et al. 2018; Dornan & Harris 2023), although it is somewhat uncertain at the smallest masses (Bastian et al. 2020). Given that stellar mass–halo mass relation is nonlinear and in the stellar mass range of the MW progenitors is close to $M_\star \propto M_{\mathrm{h}}^\alpha$ with $\alpha \approx 1.5 - 2$ for $M \approx 10^{11} - 10^{12} \, M_\odot$ and $\alpha \approx 2 - 2.5$ for smaller masses (e.g., Kravtsov et al. 2018; Read & Erkal 2019; Nadler et al. 2020), the fraction of stellar mass in bound star clusters increases with decreasing stellar mass as $f_{\mathrm{cl}} \propto M_\star^{1/\alpha - 1}$, at least for galaxies of $M_\star \gtrsim 10^7 \, M_\odot$ that contain at least a single GC (e.g., Chen & Gnedin 2023). We thus can expect the mass fraction of stars formed in bound clusters to increase with decreasing stellar mass and metallicity for $M_\star \lesssim 10^{10} \, M_\odot$ and variations of the fraction of star formation in bound clusters are therefore quite ubiquitous.

High-z galaxies with a significant fraction of rest-frame UV light concentrated in compact (sizes of tens of parsecs) clumps are also directly observed (e.g., Livermore et al. 2015; Johnson et al. 2017; Vanzella et al. 2023), although such observations are few because they require a lucky lensing configuration. At $z \approx 0$, however, the fraction of UV light in compact bound clusters is found to be a function of star formation surface density, Σ_{SFR}, (e.g., Adamo et al. 2020b, see also Adamo et al. 2020a for a review) reaching $\sim 50 - 100\%$ in galaxies with highest Σ_{SFR}. The overall observed trend is qualitatively consistent with expectations of cluster formation models (Kruijssen 2012). Furthermore, during early epochs galaxy progenitors are 1) in the fast mass accretion regime, 2) have small sizes, and 3) are expected to have high gas fractions. We can thus expect that surface density of gas and star formation, and correspondingly the fraction of star formation in bound clusters, to be comparably high during these epochs. Observations of both local and high-z galaxies are therefore consistent with the conclusion that a significant fraction of star formation in the early MW occurred in compact, bound star clusters.

GCs are often used as probes of early evolution of the Milky Way (e.g., Kruijssen et al. 2019, 2020). The conclusion that a significant fraction of stars formed in compact bound clusters also has a straightforward implication for galaxy formation simulations. If modern simulations do a good job modelling stellar populations of observed galaxies (see Naab & Ostriker 2017; Vogelsberger et al. 2020, for reviews of recent dramatic progress in such simulations) and if GCs were a significant fraction of early star formation, we can expect that spatial distribution of the low-metallicity in-situ stars born in simulations should be similar to that of observed GCs.

Indeed, the left panel of Figure 13 shows a remarkable similarity between the galactocentric distance distribution of the low-metallicity in-situ stars in the FIRE-2 simulations (Hopkins et al.

2018a; Wetzel et al. 2023) of the MW-sized progenitors and the corresponding distribution of the MW GCs. Although the model distribution varies from object to object, the predictions are quite close to the observed GC distribution and the variations are much smaller than, say, the difference in the distance distribution of the MW in-situ and accreted GCs. In fact, six out of seven of the hosts have predicted distance distributions within the bootstrap 2σ uncertainty of the MW GC distribution (the outer shaded band).

Moreover, as can be seen in the right panel of Figure 13, the FIRE-2 galaxies that match the distance distribution of the MW GCs (m12b, m12f, m12r) also have the age-[Fe/H] distribution that is closest to that of the MW GCs. These hosts have the earliest formation (oldest ages of stars at a given metallicity), which indicates that MW star formation history corresponds to the earliest-forming tail of the MW-sized galaxies. This is consistent with independent indications of the early formation of the MW based on the occurrence and timing of the GS/E merger discussed by Fattahi et al. (2019) and Dillamore et al. (2022) and model-based arguments using comparisons of the observed and predicted $[\alpha/\text{Fe}]$-metallicity distributions in the Milky Way (Mackereth et al. 2018; Hughes et al. 2020). This is also consistent with the results and conclusion of Semenov et al. (2023) who showed that the MW-like galaxies that assemble their halo and stellar mass earliest in the Illustris TNG50 simulation match observational signatures of the MW disk spin-up best.

Interestingly, if we adopt $f_{\text{N/O}}$ as a function of [Fe/H] estimated in this study (see Fig. 8 and eq. 3) as a proxy for the functional form of the $f_{\text{cl}}([\text{Fe/H}])$ dependence, and convolve the metallicity distribution of in-situ stars in the FIRE-2 simulation with that function, the resulting distribution of metallicities has a broad peak at $[\text{Fe/H}] \sim -1 \div -1.5$ resembling broadly metallicity distribution of the low-metallicity in-situ MW GCs. This consistency check indicates that the assumption $f_{\text{cl}} \propto f_{\text{N/O}}$ is broadly consistent with observed properties of the MW in-situ GCs.

The plausible models for producing enhanced abundances of nitrogen involve massive stars (see, e.g., Bastian & Lardo 2018) – massive rotating stars ($M \gtrsim 15\,M_\odot$, e.g., Maeder & Meynet 2006; Crowther 2007) or very massive stars of $\sim 10^3 - 10^4\,M_\odot$ (Denissenkov & Hartwick 2014). If stellar collisions or some other physical processes that facilitate the formation of massive stars at low metallicities in clusters can effectively create a substantial number of massive and very massive stars, this can potentially make the initial mass function (IMF) more top-heavy (see, e.g., Dib et al. 2007; Chon et al. 2021) and significantly boost the UV luminosity per unit of stellar mass formed.

4.1 From pre-disk Aurora stage to disk formation: qualitative changes in Galaxy's evolution

This work adds new details to the emerging picture of the rapid and profound transformation that the Milky Way underwent in its youth. Below we list the main changes in the properties of the Galaxy around the high redshift epoch corresponding to metallicities $[\text{Fe/H}] \approx -1$.

Spin-up. The Milky Way was not born a disk. As we demonstrated in BK22, the earliest state of the Galaxy accessible to scrutiny through an archaeological record of its stars is that of i) high velocity dispersion and ii) little net rotation. Around $[\text{Fe/H}] \approx -1.5$, the velocity ellipsoid of the in-situ stars is quasi-isotropic with a dispersion of ~ 90 km/s. At these metallicities, there is some evidence for a modest net spin characterized by a median $V_\phi \approx 50$ km/s. The median V_ϕ then increases promptly with in-

creasing metallicity, and by $[\text{Fe/H}] \approx -0.9$, the Galaxy has an established coherently rotating disk with a median tangential velocity of $V_\phi \approx 150$ km/s. This is revealed in Figures 4 and 5 of BK22 (see also Figure 3 of Conroy et al. 2022 and Figure 6 of Rix et al. 2022). Numerical simulations show that the record of the original kinematic state of the Galaxy remains largely unaltered until the present day notwithstanding the intervening merger history (see, e.g., BK22; McCluskey et al. 2023).

Scatter in elemental abundances. As its bulk stellar kinematics is reshaped, Galaxy's chemical properties evolve as well. The *Aurora* pre-disk stellar population exhibits a large scatter in most chemical abundance ratios. The origin of this scatter has not been pinned down yet, but it is hypothesised that continuous gas accretion and bursty stat-formation may play a role. Figure 7 of BK22 shows that dispersions in abundance ratios of most elements shrink on average by a factor of ~ 2 on crossing the $[\text{Fe/H}] \approx -1$ threshold. Some elements, such as nitrogen, exhibit a more dramatic decrease in abundance scatter, i.e. by a factor of ~ 4 as the Galaxy forms its disk.

Transition from clustered to regular star formation. In BK22 we showed that the elements that exhibit the largest scatter in the Aurora population are N, Al, O, and Si, i.e. all of the established markers of anomalous chemical evolution characteristic of stars in GCs. It is therefore surmised that massive GCs had an important role in the formation and evolution of the high-redshift Milky Way, imprinted in the properties of the Aurora stellar population. In this work, we explore further the contribution of GCs to the Galactic stellar field and show that the fraction of the stars form in bound clusters is at its highest in Aurora, i.e. at $[\text{Fe/H}] < -1$, and drops by more than an order of magnitude at $[\text{Fe/H}] > -1$, where the Galactic disk forms.

The α-bump. There is evidence that not only the structural properties of the Galactic star formation changed drastically with the creation of the disk, but also its overall efficiency. The first obvious signature of the likely star-formation burst taking place around $[\text{Fe/H}] \approx -1$ is in the rise and fall of the median α-abundance ratios of both in-situ field stars and in-situ GCs as a function of metallicity (see middle panel of Figure 3), which was pointed out and explored in detail by Conroy et al. (2022).

Change of slope of the metallicity distribution. An increase in the pace of star formation betrayed by the rise of [Mg/Fe] is naturally imprinted in the shape of the metallicity distribution function. As left panel of Figure 5 shows, the MDFs of the in-situ and the accreted stars agree at $[\text{Fe/H}] < -1.5$, where the slopes of both distributions are close to $d\log(n)/d[\text{Fe/H}] \approx 1$ (also see Rix et al. 2022, for discussion). At $[\text{Fe/H}] \approx -1$, the MDF of the in-situ population has a clear inflection point: its slope abruptly changes to $d\log(n)/d[\text{Fe/H}] \approx 2.3$. In comparison, the MDF of the accreted stars (mainly GS/E) shows no features around $[\text{Fe/H}] \approx -1$.

5 CONCLUSIONS

Over the last decade, a consensus has been reached that a small number of the Milky Way field stars exhibit a chemical abundance pattern characteristic of that found in Galactic globular clusters. In particular, excess nitrogen abundance has been used widely as an effective GC fingerprint. While the globular cluster origin of these nitrogen-rich field stars is likely beyond doubt, the birthplace of the clusters themselves has remained unclear. This work provides a fresh re-assessment of the origin of the field stars enriched in nitrogen. For the first time, the origin of the bulk of such stars – the

low-metallicity in-situ Aurora population – is identified unambiguously. Below we summarize our main results and spell out their implications.

(i) We start by defining the selection criteria that relate directly to the properties of the observed GC stellar populations. The GC stars not only show elevated levels of nitrogen but also exhibit reduced oxygen as well as increased aluminium abundances. Accordingly, we identify field stars with high [N/O] and [Al/Fe] ratios, namely with [N/O]> 0.55 and [Al/Fe]> -0.1 (see Figure 1). Relying on [N/O] instead of [N/Fe] also helps us to run a comparison with measurements of gas-phase abundances in high-redshift galaxies, which are referenced to oxygen.

(ii) At the basis of our analysis is a new classification of the in-situ and accreted field stars and GCs. We first use [Al/Fe] in the metallicity range [Fe/H] $\in [-1.4, -1]$ to estimate the boundary between in-situ and accreted objects in the plane of total energy and vertical component of the angular momentum (E, L_z), as the high-[Al/Fe] and low-[Al/Fe] stars have very different (E, L_z) distributions with a small overlap (see Figure 2). We approximate the boundary with a simple parametric form (see Equation 1), but we have also tested our results against the classification done with a "Gradient Boosted Trees" machine learning method instead (Friedman 2001), using `GradientBoostingClassifier` class from the Sci-kit Learn package, and found no significant difference in classification accuracy.

(iii) We show that simply using the (E, L_z) boundary informed by distributions of stars with distinct [Al/Fe] ratios results in two population of GCs with distinct properties (see Figure 3). For example, the accreted GCs show systematically lower [Mg/Fe] ratios at fixed [Fe/H] than the in-situ objects, in-situ and accreted clusters occupy different regions of the [Mg/Fe]-[Al/Fe] space. Note that in our in-situ/accreted classification (i) we do not separate MW bulge and disk components and ii) we see no evidence for an additional low-energy component sometimes referred to as the Low-Energy group, Kraken/Heracles and Koala (see Massari et al. 2019; Kruijssen et al. 2020; Forbes 2020; Horta et al. 2021a).

(iv) We show that the distribution of the selected high-[N/O] stars in the (E, L_z) plane matches that of the Aurora population with high-[Al/Fe] and $-1.4 <$ [Fe/H] < -1 (see Figures 5 and 6). Furthermore, we show that the radial distribution of high-[N/O] stars is very similar to that of the Aurora population (see Figure 10). This indicates that the majority of the high-[N/O] stars formed in situ. This is not surprising because, if these stars form in GCs, not only the MW is expected to produce a lot more GCs than any of the accreted dwarfs due to the strong correlation between the total GC mass and the mass of the host (e.g., Hudson et al. 2014; Forbes et al. 2018), the MW-born GCs are on average older and have more time to dissolve and disrupt. While the (E, L_z) distributions of Aurora and high-[N/O] stars are very similar, a small fraction of field stars enriched in nitrogen can possibly be assigned to the accreted population, more precisely to the GS/E merger.

(v) We show that distributions of high-[N/O] stars in metallicity, energy, L_z, and galactocentric distance is similar to those of the in-situ GCs at metallicities [Fe/H] $\lesssim -1$ (see Figures 9, 11) and is different from the corresponding distributions of GCs classified as accreted. At [Fe/H] > -1 the distributions of high-[N/O]

stars and GCs are different, which indicates that at later times corresponding to such metallicities these stars form via a different route.

(vi) We estimate the high-[N/O] fraction of the total stellar mass at a given metallicity, $f_{\rm N/O}$([Fe/H]), and show that this fraction decreases rapidly with increasing metallicity (Figure 8) from $2\% < f_{\rm N/O} < 4\%$ at [Fe/H] ≈ -1.4 to $f_{\rm N/O} \approx 0.04\%$ at [Fe/H] > -0.9. In contrast, even if all high-[N/O] stars in the overlapping region of (E, L_z) are assigned to the GS/E, the high-[N/O] fraction in the progenitor dwarf galaxy is some 5 times lower at [Fe/H] < -1, i.e. $f_{\rm N/O} \approx 0.8\%$ and shows no obvious trend with metallicity.

(vii) Given that the estimated $f_{\rm N/O}$ fraction depends sensitively on the [N/O] threshold adopted for the selection of these peculiar stars, we measure the fraction of high-[N/O] stars in surviving Galactic GCs and show that $f_{\rm N/O,cl}$ scales with cluster initial mass (see Figure 4). Combining the relation between $f_{\rm N/O,cl}$ and initial mass with a model of the initial mass function of a population of freshly-born GCs, we show that the observed in-situ $f_{\rm N/O}$ implies that up to $f_{\rm cl} \approx 50\% - 70\%$ of the stellar mass being formed in bound clusters at [Fe/H] $\lesssim -1.5$ in the high-redshift MW (see Figure 8) and up to $\approx 4 - 15\%$ at [Fe/H] $= -1$ (see S 3.7).

(viii) These results show that star formation in bound star clusters was a significant mode of star formation in the early Milky Way. In this context, we show that low-metallicity ([Fe/H] < -1) in-situ stellar particles in the FIRE-2 simulations of Milky Way-sized galaxies have distributions of galactocentric distance and age-metallicity relations quite similar to those of the in-situ MW globular clusters (see Figure 13).

(ix) We use the estimated mass fraction of star formation in bound clusters as a function of metallicity to estimate the stellar mass of the Aurora population (stars with [Fe/H] < -1 formed in-situ) of $M_{\rm Aur} = 5 \pm 2 \times 10^8 (1 - f_{\rm disrupt})^{-1} M_\odot$, where $f_{\rm disrupt}$ is a fraction of clusters disrupted by $z = 0$ expected to be $\approx 0.5 - 0.7$ in the models of cluster evolution (see § 3.8 for details). This indicates that the mass of this population is comparable to that of the overall MW stellar halo. Given that the Aurora stars are more centrally concentrated than the accreted halo, this component is expected to dominate the spheroidal population in the inner ≈ 10 kpc of the Galaxy.

(x) We argue that if the MW evolution is typical of the early stages of galaxies, the high fraction of star formation in massive bound clusters may help explain the anomalous high [N/O] ratio observed in the $z \approx 10.6$ galaxy GN-z11 (see Section 3.10) and the abundance of UV-bright galaxies at $z \gtrsim 10$ (see Section 4).

ACKNOWLEDGMENTS

DATA AVAILABILITY

This study uses `allStarLite-dr17-synspec_rev1` and `apogee_astroNN-DR17` catalogues publicly available at https://www.sdss.org/dr17/irspec/spectro_data/. The catalog of the MW globular clusters with distances used in this study is publicly available at https://people.smp.uq.edu.au/HolgerBaumgardt/globular/.

REFERENCES

Abdurro'uf et al., 2021, arXiv e-prints, p. arXiv:2112.02026
Adamo A., et al., 2020a, Space Sci. Rev., 216, 69
Adamo A., et al., 2020b, MNRAS, 499, 3267
Arentsen A., et al., 2020a, MNRAS, 491, L11
Arentsen A., et al., 2020b, MNRAS, 496, 4964
Baggen J. F. W., et al., 2023, arXiv e-prints, p. arXiv:2305.17162
Bastian N., Lardo C., 2018, ARA&A, 56, 83
Bastian N., Lamers H. J. G. L. M., de Mink S. E., Longmore S. N., Goodwin S. P., Gieles M., 2013, MNRAS, 436, 2398
Bastian N., Pfeffer J., Kruijssen J. M. D., Crain R. A., Trujillo-Gomez S., Reina-Campos M., 2020, MNRAS, 498, 1050
Baumgardt H., Makino J., 2003, MNRAS, 340, 227
Baumgardt H., Vasiliev E., 2021a, MNRAS, 505, 5957
Baumgardt H., Vasiliev E., 2021b, MNRAS, 505, 5957
Bekki K., Yahagi H., Nagashima M., Forbes D. A., 2008, MNRAS, 387, 1131
Belokurov V., Kravtsov A., 2022, MNRAS, 514, 689
Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, MNRAS, 478, 611
Belokurov V., Sanders J. L., Fattahi A., Smith M. C., Deason A. J., Evans N. W., Grand R. J. J., 2020, MNRAS, 494, 3880
Belokurov V., Vasiliev E., Deason A. J., Koposov S. E., Fattahi A., Dillamore A. M., Davies E. Y., Grand R. J. J., 2023, MNRAS, 518, 6200
Bouwens R. J., Illingworth G. D., van Dokkum P. G., Ribeiro B., Oesch P. A., Stefanon M., 2021, AJ, 162, 255
Bouwens R., Illingworth G., Oesch P., Stefanon M., Naidu R., van Leeuwen I., Magee D., 2023, MNRAS,
Bunker A. J., et al., 2023, arXiv e-prints, p. arXiv:2302.07256
Callingham T. M., Cautun M., Deason A. J., Frenk C. S., Grand R. J. J., Marinacci F., 2022, MNRAS, 513, 4107
Cameron A. J., Katz H., Rey M. P., Saxena A., 2023, arXiv e-prints, p. arXiv:2302.10142
Carollo D., Martell S. L., Beers T. C., Freeman K. C., 2013, ApJ, 769, 87
Carretta E., Bragaglia A., 2022, A&A, 660, L1
Charbonnel C., Schaerer D., Prantzos N., Ramírez-Galeano L., Fragos T., Kuruvandothi A., Marques-Chaves R., Gieles M., 2023, arXiv e-prints, p. arXiv:2303.07955
Chen Y., Gnedin O. Y., 2023, MNRAS,
Choksi N., Gnedin O. Y., 2019, MNRAS, 486, 331
Choksi N., Gnedin O. Y., Li H., 2018, MNRAS, 480, 2343
Chon S., Omukai K., Schneider R., 2021, MNRAS, 508, 4175
Conroy C., 2012, ApJ, 758, 21
Conroy C., et al., 2022, arXiv e-prints, p. arXiv:2204.02989
Crowther P. A., 2007, ARA&A, 45, 177
Curtis-Lake E., et al., 2023, Nature Astronomy,
D'Ercole A., Vesperini E., D'Antona F., McMillan S. L. W., Recchi S., 2008, MNRAS, 391, 825
Das P., Hawkins K., Jofré P., 2020, MNRAS, 493, 5195
Deason A. J., Belokurov V., Evans N. W., Johnston K. V., 2013, ApJ, 763, 113
Deason A. J., Belokurov V., Sanders J. L., 2019, MNRAS, 490, 3426
Denissenkov P. A., Hartwick F. D. A., 2014, MNRAS, 437, L21
Dib S., Kim J., Shadmehri M., 2007, MNRAS, 381, L40
Dillamore A. M., Belokurov V., Font A. S., McCarthy I. G., 2022, MNRAS, 513, 1867
Donlon T., Newberg H. J., 2023, ApJ, 944, 169
Donnan C. T., et al., 2023, MNRAS, 518, 6011
Dornan V., Harris W. E., 2023, arXiv e-prints, p. arXiv:2304.11210
El-Badry K., Quataert E., Weisz D. R., Choksi N., Boylan-Kolchin M., 2019, MNRAS, 482, 4528
Fattahi A., et al., 2019, MNRAS, 484, 4471
Fernández-Trincado J. G., et al., 2017, ApJ, 846, L2
Fernández-Trincado J. G., Chaves-Velasquez L., Pérez-Villegas A., Vieira K., Moreno E., Ortigoza-Urdaneta M., Vega-Neme L., 2020a, MNRAS, 495, 4113
Fernández-Trincado J. G., Beers T. C., Minniti D., Tang B., Villanova S., Geisler D., Pérez-Villegas A., Vieira K., 2020b, A&A, 643, L4
Fernández-Trincado J. G., Beers T. C., Minniti D., 2020c, A&A, 644, A83
Fernández-Trincado J. G., et al., 2020d, ApJ, 903, L17
Fernández-Trincado J. G., et al., 2022, A&A, 663, A126
Feuillet D. K., Feltzing S., Sahlholdt C. L., Casagrande L., 2020, MNRAS, 497, 109
Finkelstein S. L., et al., 2023, ApJ, 946, L13
Forbes D. A., 2020, MNRAS, 493, 847
Forbes D. A., Read J. I., Gieles M., Collins M. L. M., 2018, MNRAS, 481, 5592
Friedman J. H., 2001, The Annals of Statistics, 29, 1189
Gaia Collaboration et al., 2021, A&A, 649, A1
Gaia Collaboration et al., 2022, arXiv e-prints, p. arXiv:2206.06207
Gieles M., Gnedin O., 2023, arXiv e-prints, p. arXiv:2303.03791
Gratton R., Bragaglia A., Carretta E., D'Orazi V., Lucatello S., Sollima A., 2019, A&ARv, 27, 8
Gravity Collaboration et al., 2022, A&A, 657, L12
Harikane Y., et al., 2023, ApJS, 265, 5
Harris W. E., 2010, arXiv e-prints, p. arXiv:1012.3224
Harris W. E., Blakeslee J. P., Harris G. L. H., 2017, ApJ, 836, 67
Hasselquist S., et al., 2021, ApJ, 923, 172
Hawkins K., Jofré P., Masseron T., Gilmore G., 2015, MNRAS, 453, 758
Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, Nature, 563, 85
Hopkins P. F., 2015, MNRAS, 450, 53
Hopkins P. F., et al., 2018a, MNRAS, 480, 800
Hopkins P. F., et al., 2018b, MNRAS, 480, 800

Horta D., et al., 2020, MNRAS, 493, 3363
Horta D., et al., 2021a, MNRAS, 500, 1385
Horta D., et al., 2021b, MNRAS, 500, 5462
Horta D., et al., 2023a, MNRAS, 520, 5671
Horta D., et al., 2023b, ApJ, 943, 158
Hudson M. J., Harris G. L., Harris W. E., 2014, ApJ, 787, L5
Hughes M. E., Pfeffer J. L., Martig M., Reina-Campos M., Bastian N., Crain R. A., Kruijssen J. M. D., 2020, MNRAS, 491, 4012
Hunter J. D., 2007, Computing In Science & Engineering, 9, 90
Iorio G., Belokurov V., 2021, MNRAS, 502, 5686
Johnson T. L., et al., 2017, ApJ, 843, L21
Johnson J. W., Weinberg D. H., Vincenzo F., Bird J. C., Griffith E. J., 2023, MNRAS, 520, 782
Jones E., Oliphant T., Peterson P., et al., 2001–, SciPy: Open source scientific tools for Python, http://www.scipy.org/
Kim C., Lee Y. S., Beers T. C., Kim Y. K., 2023, arXiv e-prints, p. arXiv:2305.04025
Kisku S., et al., 2021, MNRAS, 504, 1657
Kobayashi C., Karakas A. I., Lugaro M., 2020, ApJ, 900, 179
Koch-Hansen A. J., Hansen C. J., McWilliam A., 2021, A&A, 653, A2
Koch A., Grebel E. K., Martell S. L., 2019, A&A, 625, A75
Kravtsov A. V., Gnedin O. Y., 2005, ApJ, 623, 650
Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, Astronomy Letters, 44, 8
Kruijssen J. M. D., 2012, MNRAS, 426, 3008
Kruijssen J. M. D., Pfeffer J. L., Reina-Campos M., Crain R. A., Bastian N., 2019, MNRAS, 486, 3180
Kruijssen J. M. D., et al., 2020, MNRAS, 498, 2472
Krumholz M. R., McKee C. F., Bland-Hawthorn J., 2019, ARA&A, 57, 227
Leung H. W., Bovy J., 2019, MNRAS, 489, 2079
Lind K., et al., 2015, A&A, 575, L12
Lindegren L., et al., 2021, A&A, 649, A2
Livermore R. C., et al., 2015, MNRAS, 450, 1812
Lodders K., 2019, arXiv e-prints, p. arXiv:1912.00844
Mackereth J. T., Bovy J., 2018, PASP, 130, 114501
Mackereth J. T., Crain R. A., Schiavon R. P., Schaye J., Theuns T., Schaller M., 2018, MNRAS, 477, 5072
Mackereth J. T., et al., 2019, MNRAS, 482, 3426
Maeder A., Meynet G., 2006, A&A, 448, L37
Martell S. L., Grebel E. K., 2010, A&A, 519, A14
Martell S. L., Smolinski J. P., Beers T. C., Grebel E. K., 2011, A&A, 534, A136
Martell S. L., et al., 2016, ApJ, 825, 146
Massari D., Koppelman H. H., Helmi A., 2019, A&A, 630, L4
McCluskey F., Wetzel A., Loebman S. R., Moreno J., Faucher-Giguere C.-A., 2023, arXiv e-prints, p. arXiv:2303.14210
McKenzie M., et al., 2022, MNRAS, 516, 3515
Milone A. P., Marino A. F., 2022, Universe, 8, 359
Milone A. P., et al., 2017, MNRAS, 464, 3636
Milone A. P., et al., 2020, MNRAS, 491, 515
Monty S., et al., 2023, MNRAS, 522, 4404
Myeong G. C., Evans N. W., Belokurov V., Sanders J. L., Koposov S. E., 2018, ApJ, 863, L28
Myeong G. C., Vasiliev E., Iorio G., Evans N. W., Belokurov V., 2019, MNRAS, 488, 1235
Myeong G. C., Belokurov V., Aguado D. S., Evans N. W., Caldwell N., Bradley J., 2022, ApJ, 938, 21
Naab T., Ostriker J. P., 2017, ARA&A, 55, 59
Nadler E. O., et al., 2020, ApJ, 893, 48
Naidu R. P., et al., 2022, ApJ, 926, L36
Norris M. A., van de Ven G., Kannappan S. J., Schinnerer E., Leaman R., 2019, MNRAS, 488, 5400
Oesch P. A., et al., 2016, ApJ, 819, 129
Oliphant T. E., 2015, Guide to NumPy, 2nd edn. CreateSpace Independent Publishing Platform, USA
Ono Y., et al., 2022, arXiv e-prints, p. arXiv:2208.13582
Pedregosa F., et al., 2011, J. Mach. Learn. Res., 12, 2825–2830

Pérez-Villegas A., Barbuy B., Kerber L. O., Ortolani S., Souza S. O., Bica E., 2020, MNRAS, 491, 3251
Pfeffer J., Kruijssen J. M. D., Bastian N., Crain R. A., Trujillo-Gomez S., 2023, MNRAS, 519, 5384
Phillips S. G., et al., 2022, MNRAS, 510, 3727
Read J. I., Erkal D., 2019, MNRAS, 487, 5799
Rix H.-W., et al., 2022, ApJ, 941, 45
Robertson B. E., et al., 2023, Nature Astronomy,
Rodriguez C. L., Hafen Z., Grudić M. Y., Lamberts A., Sharma K., Faucher-Giguère C.-A., Wetzel A., 2023, MNRAS, 521, 124
Sanders J. L., Belokurov V., Man K. T. F., 2021, MNRAS, 506, 4321
Schechter P., 1976, ApJ, 203, 297
Schiavon R. P., et al., 2017, MNRAS, 465, 501
Semenov V. A., Conroy C., Chandra V., Hernquist L., Nelson D., 2023, arXiv e-prints, p. arXiv:2306.09398
Senchyna P., Plat A., Stark D. P., Rudie G. C., 2023, arXiv e-prints, p. arXiv:2303.04179
Spitler L. R., Forbes D. A., 2009, MNRAS, 392, L1
Tacchella S., et al., 2023, arXiv e-prints, p. arXiv:2302.07234
Tang B., Liu C., Fernández-Trincado J. G., Geisler D., Shi J., Zamora O., Worthey G., Moreno E., 2019, ApJ, 871, 58
Tang B., Fernández-Trincado J. G., Liu C., Yu J., Yan H., Gao Q., Shi J., Geisler D., 2020, ApJ, 891, 28
VandenBerg D. A., Brogaard K., Leaman R., Casagrande L., 2013, ApJ, 775, 134
Vanzella E., et al., 2023, ApJ, 945, 53
Vasiliev E., Baumgardt H., 2021, MNRAS, 505, 5978
Vasiliev E., Belokurov V., Evans N. W., 2022, ApJ, 926, 203
Vogelsberger M., Marinacci F., Torrey P., Puchwein E., 2020, Nature Reviews Physics, 2, 42
Weinberg D. H., Andrews B. H., Freudenburg J., 2017, ApJ, 837, 183
Weisz D. R., Savino A., Dolphin A. E., 2023, ApJ, 948, 50
Wetzel A., et al., 2023, ApJS, 265, 44
Zick T. O., Weisz D. R., Boylan-Kolchin M., 2018, MNRAS, 477, 480

Table 1. Definitions for terms and notations used in this work

Term	Definition/description
The main progenitor	progenitor of the Milky Way with the largest halo mass at any given time
In-situ stars	Stars formed inside the main progenitor of the Milky Way. This study identifies these stars using their [Al/Fe] ratio and the boundary in the $E - L_z$ space based on this ratio (see Section 3.2 and eq. 1).
Accreted (ex-situ) stars	Stars formed in other progenitor galaxies, which were tidally stripped from their galaxies or incorporated into the Milky Way's halo when galaxy was disrupted
The Aurora stellar component	In-situ stars of [Fe/H] < -1 formed during the early pre-disk stage of evolution.
GS/E	*Gaia* Sausage/Enceladus galaxy: tidal debris stars of a massive dwarf galaxy that merged with the Milky Way $\approx 8 - 11$ Gyr ago. Other commonly used abbreviations include GES, GE/S, GS, GE.
Disk spin-up	Rapid increase in the net angular momentum of in-situ stars over a narrow range of metallicity. In the Solar neighborhood, the median azimuthal velocity of the in-situ stars increases rapidly from $V_\phi \approx 50$ km/s at [Fe/H]$= -1.3$ to $V_\phi \approx 130$ km/s at [Fe/H]$= -1$ (BK22).
1P/2P stellar populations	The two populations of stars in globular clusters that have very different elemental abundance patterns. The 2P population is thought to form after 1P population and be enriched by it (see, e.g., Bastian & Lardo 2018, for a review).
$f_{\rm N/O}$	The fraction of stars with high [N/O] ratio selected according to the conditions stipulated in Section 3.1.
$f_{\rm N/O,cl}$	The fraction of stars born in clusters that have high N/O for a population of newlu born GCs in high-z MW progenitor.
$f_{\rm cl} = f_{\rm N/O}/f_{\rm N/O,cl}$	The fraction of the Aurora stars born in clusters (see Section 3.7).
$M_{\rm Aur}$	The stellar mass of the Aurora stellar population (see Section 3.8).
$f_{\rm disrupt}$	The fraction of the initial globular clusters that were disrupted (see Section 3.8).