Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS **545,** 1–24 (2026) https://doi.org/10.1093/mnras/staf1551

# The Milky Way in context: the formation of galactic discs and chemical sequences from a cosmological perspective

Matthew D. A. Orkney [1,2]★ Chervin F. P. Laporte [1,2,3,4] Robert J. J. Grand [5] and Volker Springel[6]

[1]*Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona, Martí i Franquès 1, E-08028 Barcelona, Spain*
[2]*Institut d'Estudis Espacials de Catalunya (IEEC), E-08034 Barcelona, Spain*
[3]*LIRA, Observatoire de Paris, Université PSL, Sorbonne Université, Université Paris Cité, CY Cergy Paris Université, CNRS, F-92190 Meudon, France*
[4]*Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan*
[5]*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool, L3 5RF, UK*
[6]*Max Planck Institut für Astrophysik, Karl-Schwarzschild-Straße 1, D-85748 Garching bei München, Germany*

Accepted 2025 September 10. Received 2025 September 6; in original form 2025 June 6

## ABSTRACT

We study the formation of chemical sequences in stellar discs of Milky Way (MW)-mass galaxies in a full cosmological context with the AURIGA simulations, with implications for both the MW and external galaxies like M31. The analysis focuses on the conditions giving rise to bimodal $\alpha$-chemistry and the potential influence of mergers (e.g. Gaia–Enceladus, GSE). We find a wide diversity of chemical sequences, without correlation between the emergence of dichotomous $\alpha$-chemistry and GSE-like mergers. The transition between multiple $\alpha$-sequences is sequential, and is mediated by modulations in the star formation rate (SFR). In some cases, this can be caused by the starburst and subsequent quiescence induced by mergers. In others, by exhaustion or violent disruption of the gas disc. Realizations with singular sequences either lack significant modulations in their SFR, or form too late to have a significant high-$\alpha$ sequence. The metallicity overlap between the high- and low-$\alpha$ sequences (as seen in the Solar neighbourhood of the MW) arises from accretion of metal-poor gas from the circumgalactic medium. This depends on gas disc thickness, with thinner discs losing their metal-poor extremities. Gas donation from singular gas-rich merger events are incapable of driving long-lived metal dilution ($\Delta$[Fe/H] $\gtrsim 0.3$), and we rule-out this scenario for the low-$\alpha$ sequence in the MW. Finally, the shape of $\alpha$-sequences in the [Fe/H] versus [Mg/Fe] plane is related to long-term SFR trends, with a downward slanted locus (as observed in the low-$\alpha$ sequence of the MW) owing to a sustained or declining SFR.

**Key words:** methods: numerical – Galaxy: abundances – Galaxy: disc – Galaxy: evolution – galaxies: abundances.

## 1 INTRODUCTION

Our vantage point within the Milky Way (MW) has allowed for comprehensive spectroscopic surveys into the chemodynamic properties of individual stars across large statistical samples and wide volumes – see the Radial Velocity Experiment (RAVE, Steinmetz et al. 2006), the Sloan Extension for Galactic Understanding and Exploration (SEGUE and SEGUE-2, Yanny et al. 2009; Rockosi et al. 2022), Gaia-ESO (Gilmore et al. 2012), the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST, Deng et al. 2012), the GALactic Archaeology with HERMES (GALAH, De Silva et al. 2015), and the APO Galactic Evolution Experiment (APOGEE, Majewski et al. 2017). This, combined with precise proper motions from *Gaia* (Gaia Collaboration 2016, 2021; Lindegren et al. 2018), has led to a broader understanding of how the MW formed and established our Galaxy as a benchmark for studying the underlying physics of galaxy formation as a whole.

High-resolution spectroscopic studies of stars in the Solar neighbourhood have revealed that disc stars form a clear 'bimodal' distribution in the [Fe/H] versus [$\alpha$/Fe] chemical abundance planes, with two distinct sequences that are separated in their $\alpha$-chemistry but overlap in their metallicity distributions (e.g. Fuhrmann 1998; Haywood et al. 2013; Anders et al. 2014; Nidever et al. 2014; Hayden et al. 2015). This bimodality varies along with the location within the disc, both in terms of radius and height above or below the plane (e.g. see Queiroz et al. 2020). For inner radii (within $\approx 6$ kpc), the sequences form a single track mediated by an underpopulated gap (e.g. Hayden et al. 2015; Queiroz et al. 2020). For outer radii (beyond $\approx 12$ kpc), the proportion of stars in the high-$\alpha$ sequence is strongly diminished (e.g. Bensby et al. 2011; Anders et al. 2014).

The high- and low-$\alpha$ sequences are well correlated with stars in the thick and thin discs respectively, and share similar qualitative properties. Whereas the high-$\alpha$ disc is uniformally old (having formed the majority of its stars earlier than 10 Gyr ago), centrally concentrated and vertically extended – the low-$\alpha$ disc is thin, radially extended and formed over a wide time period up to the present day (Haywood et al. 2013; Mackereth et al. 2017; Bonaca et al. 2020; Miglio et al. 2021; Queiroz et al. 2023). This is consistent with extra-Galactic observational evidence that galaxy formation proceeds in an inside–out and top–down manner (Franx et al. 2008; van Dokkum et al. 2013; Patel et al. 2013a; van der Wel et al. 2014; Lian & Luo 2024), a result that is backed up by idealized and fully cosmological simulations (e.g. Bournaud, Elmegreen & Martig 2009; Brook et al.

★ E-mail: morkney@icc.ub.edu

2012; Bird et al. 2013). Yet, it is increasingly believed that the high/low-$\alpha$ and thick/thin discs are distinct populations of their own (e.g. see arguments in Robin et al. 1996; Jurić et al. 2008; Bovy, Rix & Hogg 2012; Bovy et al. 2016). None the less, the populations are often treated as interchangeable and chemical cuts are commonly employed to select the thick and thin disc populations.

Whilst observational data show that thick and thin discs are a frequent attribute of galaxies as a whole (e.g. Mould 2005; Yoachim & Dalcanton 2006; Tsukui et al. 2025), the same determination has not yet been made for chemical bimodalities. Extragalactic observations rely on integrated light, which blends stellar populations across disc regions (e.g. Scott et al. 2021). Furthermore, observations at higher redshift increasingly rely on gas-phase chemistry, which may not confer the chemical distribution of the stars. The one exception is our nearest neighbour M31. Early Keck/DEIMOS (Deep Imaging Multi-Object Spectrograph) spectroscopy of 70 M31 stars found no sign of dichotomous chemistry (Escala et al. 2020), and near-infrared spectroscopy of around 100 stars from the *James Webb Space Telescope* (*JWST*) shows only a single sequence at high-$\alpha$ (Nidever et al. 2024). This is an early sign that dichotomous chemistry may not be an ubiquitous quality of massive disc galaxies, which suggests a diversity of formation channels.

If the formation mechanism behind the chemical bimodality was well known, one could debate whether it was a unique development to the MW. Despite many competing hypotheses, there has been no universal agreement on a dominant mechanism. These hypotheses fall into two categories: secular or episodic formation. An example of a secular mechanism would be the chemical evolution models of Schönrich & Binney (2009a, b). These find that low-metal and high-$\alpha$ stars form in the centrally compact proto-Galaxy,[1] and are then driven to the Solar neighbourhood by radial migration (as in Sellwood & Binney 2002; Roškar et al. 2008). There, they overlap with younger, low-$\alpha$ stars in the radially extended thin disc. This scenario is backed up by chemical enrichment models in Sharma, Hayden & Bland-Hawthorn (2021) and Chen et al. (2023), $N$-body simulations in (Loebman et al. 2011), and could also explain other properties of the disc such as the radial dependence of the metallicity distribution function (Loebman et al. 2016). However, other models with radial migration fail to reproduce two chemical sequences (Minchev, Chiappini & Martig 2013, 2014a; Johnson et al. 2021; Dubay, Johnson & Johnson 2024).

Otherwise, the two $\alpha$-sequences may form at the same time but in different Galactic environments. The authors of Clarke et al. (2019) and Beraldo e Silva et al. (2021) argue that high-$\alpha$ stars could form in the dense intermediate-mass gas clumps that arise from disc fragmentation in the early Universe (see Elmegreen & Elmegreen 2005; Elmegreen et al. 2007; Dekel, Sari & Ceverino 2009; Livermore et al. 2012), within which the star formation rate (SFR) is raised and the $\alpha$-enrichment from core-collapse supernovae (SNe) is enhanced. Then, low-$\alpha$ stars can form concurrently in a more evenly distributed gas disc. This scenario could simultaneously explain the elevated velocity dispersion of high-$\alpha$ stars via strong scattering processes (Bournaud et al. 2009; Amarante et al. 2020). However, observations from the CANDELS survey (the Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey) find that only up to 20 per cent of the integrated SFR comes from within clumps (see Wuyts et al. 2012, for galaxies over a similar redshift range and mass scale as the proto-MW).

An episodic formation mechanism might suggest that a bimodality requires a hiatus in the star formation between the formation of the high- and low-$\alpha$ sequences (as in Haywood et al. 2016). This could be due either to a shrinking and then growth of the star-forming gas disc (i.e. Grand et al. 2018), or a suppression of the SFR following merger interactions or feedback from active galactic nucleus (AGN, Beane 2025; Beane et al. 2025). The pause in the star formation allows the delayed iron enrichment from Type-Ia SNe to 'catch up', thereby reducing the ratio of [$\alpha$/Fe] prior to the onset of the low-$\alpha$ sequence. Indeed, there is tentative evidence for a quenching period around 8 Gyr ago from closed-box chemical evolution models in Haywood et al. (2018) and APOGEE DR17 data in Spitoni et al. (2024).

It could also depend on distinct epochs of gas inflow, as in the two-infall model of (Chiappini, Matteucci & Gratton 1997; Chiappini, Matteucci & Romano 2001). This proposes that the high-$\alpha$ sequence formed with a high SFR from a primordial rapid gas inflow, followed by a lower SFR from subsequent metal-poor gas accretion. The second infall 'resets' the ambient metallicity of the gas disc to lower values, thereby also explaining the signature [Fe/H] overlap between the two sequences in the Solar neighbourhood (see Grisoni et al. 2017; Spitoni et al. 2019, 2021, 2023). Prior works have discussed gas-rich mergers as the sources of a second gas infall (e.g. Richard et al. 2010; Snaith et al. 2016; Mackereth et al. 2018; Buck 2020), or as a trigger that stimulates disc growth leading to low-$\alpha$ formation (see examples in Grand et al. 2018).

The low-$\alpha$ disc may have formed from a prolonged fuelling from the circumgalactic medium (CGM). Studies have identified cold inflowing gas filaments (e.g. Anglés-Alcázar et al. 2017; Renaud et al. 2021), re-accretion of early gas outflows (e.g. Khoperskov et al. 2021), and smooth accretion from the CGM (e.g. Parul et al. 2025) as the dominant sources of gas during the formation of low-$\alpha$ sequences in simulations.

Evidence from the stellar halo has now revealed the signature of an ancient massive merger, hereafter the Gaia–Sausage Enceladus or GSE (Belokurov et al. 2018; Helmi et al. 2018). This merger is predicted to have arrived in the MW around 8–11 Gyr ago (Vincenzo et al. 2019; Belokurov et al. 2020; Naidu et al. 2021; Xiang & Rix 2022), roughly during the transition between the two $\alpha$-sequences (e.g. Linden et al. 2017; Sahlholdt, Feltzing & Feuillet 2022), with a mass scale similar to that of the present-day Small Magellanic Cloud. Numerous works argue that the GSE played an important role in the history of the MW, with implications for its disc chemistry (e.g. Buck 2020; Buck et al. 2023; Ciucă et al. 2024), the hot kinematics of the thick disc population (Helmi et al. 2018; Bonaca et al. 2020; Grand et al. 2020; Chandra et al. 2024) and the warp observed in the outer disc (Dillamore et al. 2022; Dodge et al. 2023; Deng et al. 2024). While evidence for the GSE merger is strong, its exact mass, infall angle, and impact on the proto-Galaxy remain uncertain. In particular, its exact relation to the chemical bimodality is still circumstantial. This is a topic we intend to probe further.

In this work, we will investigate the properties and formation of chemical sequences in the AURIGA simulation suite of MW-mass galaxies. Our ultimate goal is to apply insights from a statistical sample of cosmological realizations to the formation of chemical sequences, and to improve our understanding of the MW formation history in a cosmological context. We seek to address the following scientific questions:

(i) What are the physical channels through which chemical sequences form in disc galaxies?

[1]In which a proto-galaxy is defined as the most ancient progenitor to modern-day galaxies, that has undergone minimal star formation or chemical enrichment (see for e.g. Belokurov & Kravtsov 2022).

(ii) How do simulated chemical sequences qualitatively compare to that of the MW, and what does this tell us about the physical processes at play?

(iii) Are mergers, GSE-like or otherwise, a necessary component in forming a chemical bimodality at all?

(iv) How do GSE-like mergers affect the formation and evolution of the disc, and how is this translated into the present-day structure and kinematic properties of the disc?

We describe our methods and post-processing in Section 2. Our results are given in Section 3, where we consider the properties of chemical distributions at $z = 0$ in Section 3.1 and the chemical evolution of sequences in Section 3.2. In Section 3.3, we investigate how distinct $\alpha$-sequences can arise. We discuss how our findings can be interpreted with respect to observations and other simulations in Section 4. Finally, we present our conclusions in Section 5.

## 2 METHODS

### 2.1 AURIGA simulation suite

The simulations analysed in this paper are drawn from the AURIGA project (Grand et al. 2017, 2024), a suite of thirty magnetohydrodynamic simulations of isolated MW-mass galaxies evolved within a full cosmological context. The initial conditions are generated at $z = 127$ using the Gaussian white-noise realization PANPHASIA (Jenkins 2013), and have a periodic cosmological box with a side length of 100 cMpc. The suite uses cosmological parameters from Planck Collaboration XVI (2014), which are $\Omega_{\rm m} = 0.307$, $\Omega_{\rm b} = 0.04825$, $\Omega_\Lambda = 0.693$, and a *Hubble* constant of $H_0 = 100h\,{\rm km^{-1}\,Mpc^{-1}}$, where $h = 0.6777$.

The AURIGA simulations incorporate a comprehensive suite of subgrid physics to model galaxy formation in a cosmological setting. These include treatments for a spatially uniform photoionizing ultraviolet background, primordial and metal-line cooling, star formation, stellar evolution, feedback from SNe, supermassive black hole growth and feedback, as well as magnetohydrodynamics. A detailed overview of these physical prescriptions is provided in Grand et al. (2017, 2024). The simulations successfully reproduce a variety of key galactic properties expected in a cosmological context (Vogelsberger et al. 2013; Genel et al. 2014; Marinacci, Pakmor & Springel 2014). The AURIGA galaxies themselves exhibit characteristics consistent with observations of MW-like galaxies, including their halo mass–metallicity relations (Monachesi et al. 2019), vertical disc oscillations (Gómez et al. 2017), rotation curves and SFRs (Grand et al. 2017), and detailed chemodynamical structures in their central regions (Fragkoudi et al. 2020). Crucially, they also produce the presence of thick and thin disc components (Grand et al. 2018), both in terms of geometric, kinematic, and chemical properties.

The AURIGA galaxy formation models employ a two-phase interstellar medium, consisting of cold clouds immersed in a hot gaseous background. Star formation occurs stochastically in gas above a threshold density of $0.13\,m_{\rm p}\,{\rm cm^{-3}}$, with the target stellar particle mass matching the initial gas cell resolution. Each star particle represents a simple stellar population following a Chabrier initial mass function (Chabrier 2003). Mass and metals are returned to the gas via subgrid models for AGB winds, Type-II, and Type-Ia SNe. Here, the returns from Type II SNe are assumed to occur instantaneously, whilst those from Type Ia SNe follow a time-delay distribution resulting in a characteristic enrichment time-scale of $\approx 1$ Gyr. Further details on the physics prescriptions, including the adopted yield tables, are provided in the AURIGA public release (Grand et al. 2024) and introductory paper (Grand et al. 2017). Grand et al. (2024) also discuss how these choices compare to the physics implemented in the TNG simulations (Nelson et al. 2019).

Throughout this work, we use the 'level-4' resolution versions of each AURIGA simulation. Here, the high-resolution target regions of the cosmological box are resolved with dark matter (DM) particle masses of $\sim 3 \times 10^5\,{\rm M_\odot}$ and initial baryonic gas cell masses of $\sim 5 \times 10^4\,{\rm M_\odot}$. For convenience, we hereafter refer to the AURIGA simulations as Au-$i$, where $i$ indicates the particular halo.

### 2.2 Post-processing

The group and subhalo properties are calculated using the SUBFIND halo finder (Springel et al. 2001), and the virial masses and radii are derived using a sphere of mean density $200\times$ the critical density of the universe. The evolution of halo structure is tracked across time-steps using the LHALOTREE merger tree algorithm (Springel et al. 2005). When necessary, inter-snapshot subalo properties are estimated using fitted splines. These are 3D cubic splines for subhalo orbits, and a PCHIP (Piecewise Cubic Hermite Interpolating Polynomial) 1D monotonic cubic interpolation for subhalo properties.

The AURIGA simulations track nine chemical species. We use Magnesium (Mg) for our estimation of $\alpha$-abundance because it has been shown to be a reliable tracer of the $\alpha$-process, with a high separation between the thick and thin discs (Jönsson et al. 2018). All chemical abundances in this work are normalized to Solar values using Asplund et al. (2009), with further corrections of order 0.4 dex used to lower the metallicity ratio ([Fe/H]). This work is often focused on the distribution of stars in the [Fe/H] versus [Mg/Fe] elemental abundance ratios, and for brevity we will hereafter refer to this simply as the 'chemical abundance plane'.

We focus exclusively on stars in the disc, following selection criteria based on both the stellar orbital circularity (using the definition for orbital circularity given in Abadi et al. 2003) and metallicity as described in the methods section of Orkney et al. (2023). To summarize, we calculate the circularity distribution of all star particles and assume all retrograde stars belong to the stellar halo. Then, we mirror this distribution around $\eta = 0$ to find the probability that prograde stars belong to the halo. Prograde star particles are probabilistically assigned to the halo using a random number generator, and the remaining stars are classified as disc. Additionally, we enforce that the retrograde and prograde halo components share an equivalent metallicity distribution function, though we note this has little implication on the final selections.

Unless stated otherwise, we include stars from both *in-situ* and *ex-situ* sources, where the *in/ex-situ* designation is calculated based on whether the star particle is bound to the host halo at the earliest available snapshot. This is necessary to account for the presence of any '*ex-situ* disc' (see Gómez et al. 2017). We note that the *ex-situ* contributions neither materially alter our reported results, nor do they create features in chemical abundance space that might be misinterpreted as independent chemical sequences (with Au-1 being the one exception). Furthermore, it is unlikely that the MW disc is dominated by *ex-situ* stars (e.g. Ruchti et al. 2011). The galaxies are aligned at every snapshot on the angular momentum of the youngest (formed in the last 3 Gyr) and innermost ($R_{\rm G} < 0.1 \times R_{200}$, where $R_{\rm G}$ is the galactocentric radius) *in-situ* star particles.

We calculate merger mass ratios based on the ratio between the total SUBFIND masses (DM, gas, and stars) as opposed to the commonly used ratio of $M_{200}$ virial masses. This is because the virial mass will often include massive substructures, and as a consequence the time evolution of the mass ratio is more stochastic. The quoted merger mass ratio is then found as the maximum ratio after the

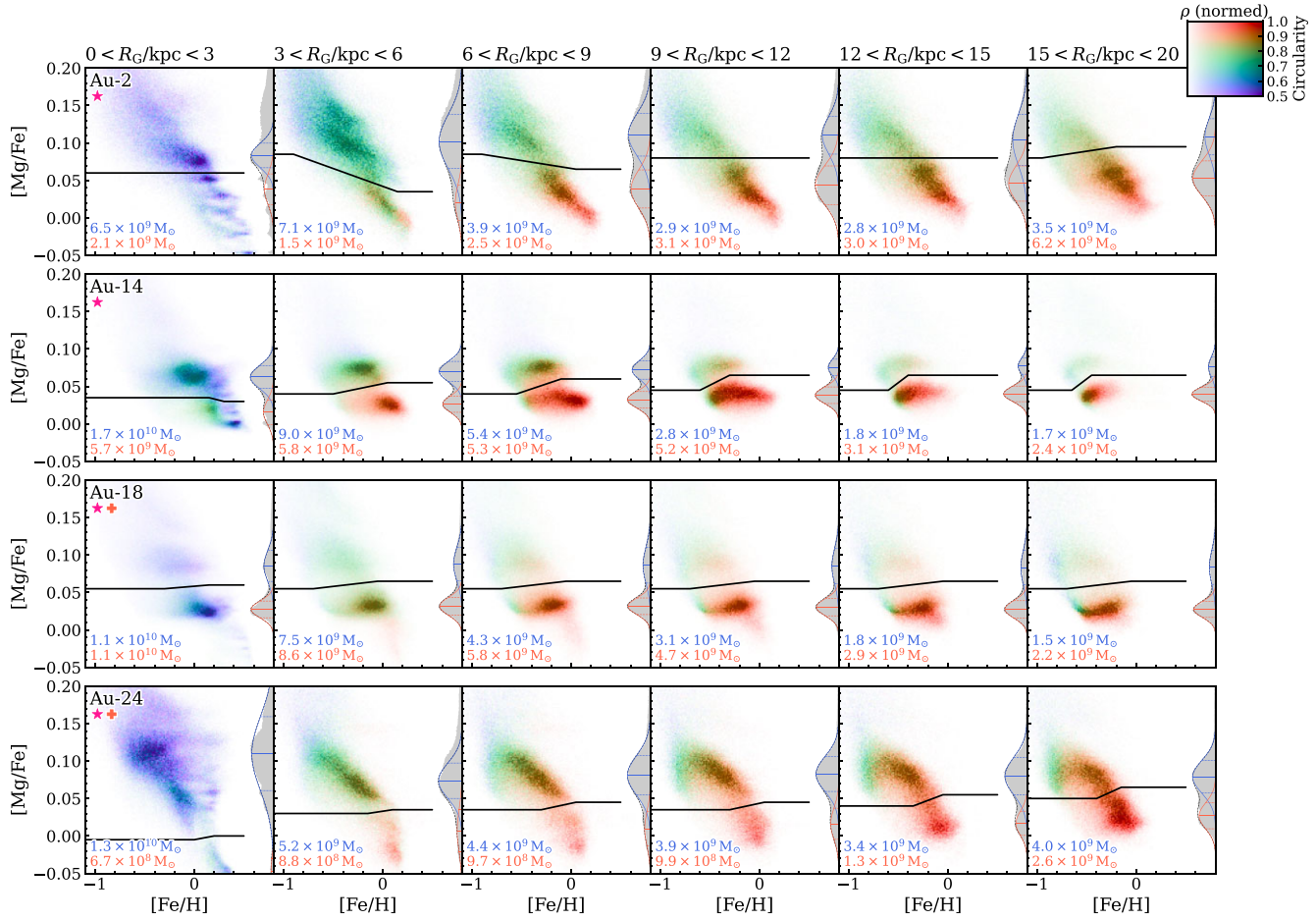


**Figure 1.** Examples of chemical sequences in four AURIGA galaxies (where the simulation name is indicated in the top-left corner of the leftmost panels), over a range of radial bins. Stars with halo-like kinematics have been excluded, following the methods described in Section 2. The 2D histograms indicates the normalized mass-weighted density of star particles in units of $M_\odot$ dex$^{-1}$ dex$^{-1}$. Whereas stars in the MW are typically depicted with a logarithmic weighting, we opt for a linear weighting to better emphasize the main overdensities. The histogram colour represents the mean orbital circularity within each pixel, as indicated by the colour bar in the top-right. Each panel includes vertical 1D histogram that show the [Mg/Fe] abundance ratio distributions, where we overplot Gaussian profile fits to the high- and low-$\alpha$ selections. These selections are shown with a thick solid line.

satellite has joined the FoF (Friends of Friends) group of the host structure. This usually occurs just after the satellite joins the FoF group, because gravitational tides and ram pressure stripping will begin to degrade it thereafter.

## 3 RESULTS

### 3.1 A diversity of chemical sequences in AURIGA

We show stellar mass-weighted histograms in the chemical abundance plane for a sample of AURIGA galaxies in Fig. 1, where the full suite is included in an online repository.[2] The four chosen examples were selected to exhibit different behaviours; *Au-2:* with a slanted and narrow track through the chemical abundance plane, reminiscent of the distribution observed in the MW at sub-Solar radii (e.g. fig. 6 of Queiroz et al. 2020); *Au-14 :* with a clear bimodality exhibiting a centrally dominant high-$\alpha$ sequence, formed in a galaxy that experiences multiple gas-rich mergers on shallow infall trajectories; *Au-18:* with a clear chemical dichotomy that forms in a realization with a GSE-like merger occurring at a lookback time of $\tau = 9.3$ Gyr, and a significant metallicity overlap between each chemical sequence; and *Au-24:* with another chemical dichotomy forming in a realization with a GSE-like merger, this time at $\tau = 8.5$ Gyr. This simulation exhibits a slanted high-$\alpha$ sequence similar to the chemical distribution in the MW, but a less massive low-$\alpha$ sequence that has a narrow metal distribution ($\Delta$[Fe/H] $\approx 0.3$) and hardly any overlap with the metallicity of the high-$\alpha$ sequence, much unlike the MW.

Each panel features a black line, determined by eye, which separates the stars into high- and low-$\alpha$ selections. Manual selections like these are commonly employed in both observational data and simulated analogues, and our reported results are not sensitive to the exact selection choices. We adjust the selections with each radial bin in order to account for the radial dependence of the chemical sequences in some realizations, which ensures a minimal contamination between the high- and low-$\alpha$ sequences. The total stellar mass in each selection is indicated in the lower left corner of each panel. For the purposes of this investigation, we always separate the stars into a high- and low-$\alpha$ selection even when just

[2]We echo the acknowledgment in Grand et al. (2018) that the $\alpha$-abundances in AURIGA do not necessarily align with the chemistry expected for the MW, and refer the reader to their discussion section where these concerns are explored.

one dominant sequence is present. All chemical sequences display some level of granularity in their distribution, and so it is always possible to define an appropriate separation even in the case of single sequences.

We further divide the chemical abundance plane into six radial bins spanning $0 < R_G/\mathrm{kpc} < 20$, which brackets the majority of disc stars at $z = 0$. We note that there are also trends with height above and below the disc plane, as seen in the MW (e.g. Nidever et al. 2014; Minchev, Chiappini & Martig 2014a; Hayden et al. 2015; Queiroz et al. 2020) and also in AURIGA at a higher resolution level (Grand et al. 2018). Whilst we do not show these trends explicitly, we find that the proportion of stars in each $\alpha$-sequence is dependent on the $|z|$-height, with stars at higher [Mg/Fe] typically gaining mass dominance at greater $|z|$-heights (especially around and within Solar radii). This is in agreement with observational data of the MW from APOGEE (e.g. Hayden et al. 2015), and is therefore likely to be a general outcome of galaxy formation at MW-mass scales rather than a property that is unique to the MW.

The histograms are coloured by the mean orbital circularity of stars within each pixel (see the definition in Orkney et al. 2023), revealing some interesting features. First, stars in the high-$\alpha$ selection tend to have lower orbital circularities than those in the low-$\alpha$ selection, in-keeping with expectations from our own Galaxy (see for example fig. 5 in Chandra et al. 2024; fig. 11 in Khoperskov et al. 2025). This pattern is consistent with a top–down, inside–out formation scenario: the older high-$\alpha$ stars formed when the galaxy was more radially compact but vertically extended, resulting in, on average, less circular orbits. In line with this picture, high-$\alpha$ stars tend to contribute more mass at lower orbital radii than their low-$\alpha$ counterparts. We will show that this is indeed the case for one example simulation in Section 3.3.3.

Secondly, stars at lower [Fe/H] tend to have lower orbital circularities even at a fixed [Mg/Fe]. This behaviour can also be seen in the MW: see the low-$\alpha$ sequence below a metallicity of approximately $[\mathrm{Fe/H}] = -0.3$ in fig. 6 of Chandra et al. (2024). One interpretation would be that these low-metal stars formed at older times, meaning they have undergone more kinematic perturbations that might heat their orbits, or that they formed at a time when the disc was of a lower orbital circularity. However, both here and in the MW, this wide metallicity spread in the low-$\alpha$ sequence is more accurately a proxy of the underlying radial metallicity gradient than any temporal evolution (i.e. Frankel et al. 2018, 2020; Patil & Bovy 2024; Cerqui et al. 2025). This is because the stellar disc has kinematically cooled and grown to greater radii during the formation of the low-$\alpha$ sequence, such that the radial metallicity gradient dominates over the temporal metallicity evolution. We will argue in Section 3.3.3 that the lower circularity of these low-metal and low-$\alpha$ stars owes to their formation from recently accreted gases that have not yet fully settled into the rotating gaseous disc.

### 3.1.1 *Quantifying chemical bimodalities*

From a cursory examination of the full suite, there is clearly a wide variation in the shape and number of $\alpha$-sequences. To aid comparison, we manually classify the chemical distributions into four groups:

(i) ★ A bimodality.
(ii) ★ A unimodality.
(iii) ★ A multimodality, for cases where there are more than two substantial loci in the chemical abundance plane.
(iv) ★ Smeared chemistry, for cases with a diffuse smear of stars across the chemical abundance plane, that do not adhere to obviously independent sequences.

In some instances, the classification would change depending on the radial bin considered, and so we prioritize the classification that seems to fit the best overall. Coloured stars corresponding to each of our four classifications are appended to the first column of panels in Fig. 1, and will be used elsewhere in this work where appropriate. We also include a ✚ symbol to indicate simulations that possess a radially anisotropic stellar halo feature loosely comparable to the GSE in the MW. These realizations are taken from the grey-dashed box selection in fig. 3 of Fattahi et al. (2019), which includes stellar haloes that have features corresponding to a velocity anisotropy of $\beta > 0.8$ and a stellar halo contribution fraction of $> 0.5$. The majority of the stellar mass in these features owes to ancient massive mergers with radial infalls (and see also Orkney et al. 2023), though not all of these mergers are GSE-like in the totality of their properties (i.e. their infall time or merger mass ratio). The mergers in Au-18 and Au-24 are perhaps the most comparable to expectations of the GSE.

For a more quantitative analysis, we fit a Gaussian profile to the manual selections used in Fig. 1, and then estimate the overlap between the two profiles by using the Bhattacharyya (1946) distance ($D_B$):

$$D_B(p,q) = \frac{1}{4}\ln\left[\frac{1}{4}\left(\frac{\sigma_p^2}{\sigma_q^2}+\frac{\sigma_q^2}{\sigma_p^2}+2\right)\right]+\frac{1}{4}\frac{(\mu_p-\mu_q)^2}{\sigma_p^2+\sigma_q^2}, \quad (1)$$

where $p, q$ correspond to two normal distributions, $\mu$ is the mean, and $\sigma^2$ is the variance.

We use the assumption that the chemical sequences are wel described by Gaussian profiles, despite the lack of direct physical justification for this choice. Nevertheless, the Gaussian models provide good fits in nearly all cases (e.g. see the histograms along the right $y$-axis in Fig. 2). We also remind the reader that our manual selections divide the chemical plane into the most representative high- and low-$\alpha$ populations, even in cases where the distribution appears predominantly unimodal or smeared. In the case of smeared distributions, the Gaussian fits are far poorer.

This simple estimate will yield higher values in cases where the two Gaussians have less overlap. However, there may be cases where one of the two sequences has only a negligible mass fraction, and this should never be considered as a 'strong' bimodality. Therefore, we modify $D_B$ as such:

$$D_{\mathrm{Bmod}}(p,q) = D_B(p,q)\times\sqrt{\frac{\min(A_p,A_q)}{\max(A_p,A_q)}}, \quad (2)$$

where $A$ is the Gaussian amplitude, meaning the final value is reduced in the case of uneven Gaussian amplitudes.

We calculate $D_{\mathrm{Bmod}}$ over each radial bin and display the results in Fig. 2. This shows that there is a wide variation in the 'goodness' of bimodalites within AURIGA, both between realizations and also at different radii. Generally, the values of $D_{\mathrm{Bmod}}$ are well correlated with our visual classifications. Bimodal realizations have higher values and unimodal realizations have lower values, but realizations with 'smeared' chemistry are not well modelled by two Gaussians and should be discounted.

The figure also demonstrates that radial variation, even between neighbouring radial bins, can have a significant impact on the goodness of the chemical bimodality. This can be due to both a variation in the mass budget in high- and low-$\alpha$ sequences, but also due to the changing chemical abundances of stars. An example of

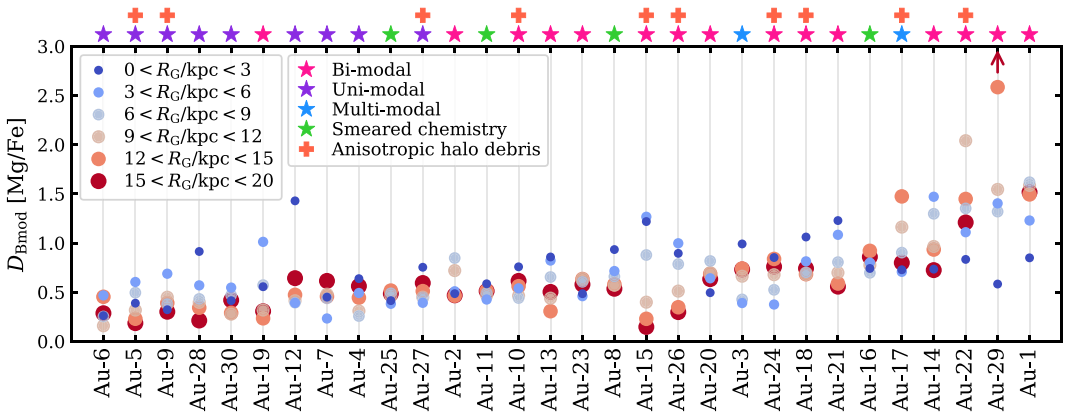


**Figure 2.** A measure of the bimodality strength across the AURIGA suite, as defined by a modified Bhattacharyya distance in equation (2). Stars are split into a range of radial bins, as given in the left legend. The ordering of simulations along the $x$-axis is determined by their average value of $D_{\rm Bmod}$. The markers along the top $x$-axis denote the type of chemical distribution as judged by eye (see the right legend), with an additional '+' symbol to highlight galaxies that have a stellar halo debris feature (as found in the dashed box of fig. 3 in Fattahi et al. 2019, galaxies that have a stellar halo with a component of velocity anisotropy $\beta > 0.8$ and contribution fraction $> 0.5$). These features were linked to specific massive mergers with radial infalls (and see also Orkney et al. 2023). There is no strong correlation between $D_{\rm Bmod}$ and the presence of anisotropic halo debris.

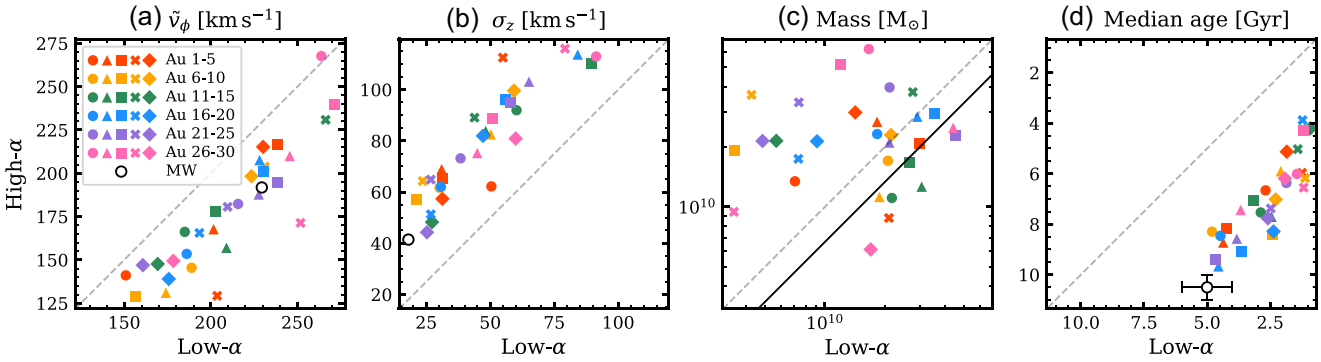


**Figure 3.** A comparison of four salient properties for the high- and low-$\alpha$ selections described in Section 3.1.1 (for stars in the radial range $0 < R_{\rm G}/{\rm kpc} < 20$). In order; (a) the median azimuthal velocity, (b) the standard deviation of the velocity in the $z$-direction, (c) the summed stellar mass, and (d) the median of the stellar age. The values for each AURIGA simulation are provided by unique markers as indicated in the legend. We show expected values for the MW with an unfilled circle, where kinematic properties are taken from Anguiano et al. (2020), and the high/low-$\alpha$ mass ratio value is taken from Khoperskov et al. (2025).

the latter effect can be seen in Au-15, where two distinct sequences present at $3 < R_{\rm G}/{\rm kpc} < 6$ begin to converge towards a single mean [Mg/Fe] value at progressively higher radii.

### *3.1.2 Correlation with GSE-like halo debris*

For the reasons discussed in our introduction, it may be expected that realizations with GSE-like mergers would be more likely to have two distinct chemical sequences. However, from Fig. 2 we find that simulations hosting GSE-like debris features (marked with a + symbol) have chemical abundance distributions at both low and high values of $D_{\rm Bmod}$. Furthermore, there are examples of galaxies with strong chemical dichotomies that do not experience GSE-like mergers (i.e. Au-14 and Au-29).

Whilst GSE-like mergers do not seem to be a prerequisite in forming multiple chemical sequences, there may still be a requirement for some kind of merger activity. As an example, Au-25 has a particularly quiescent accretion history with the last $> 1:30$ mass ratio merger occurring before a lookback time of $\tau = 11$ Gyr. The evolution through the chemical abundance plane is smooth and almost featureless. This is in opposition to results from Prantzos et al. (2023), which finds that a chemical bimodality can develop from secular evolution even in the absence of mergers. We will return to this idea in Section 3.2.

### *3.1.3 Properties of each chemical sequence*

We present a comparison of four different properties for the high- and low-$\alpha$ selections in Fig. 3: (a) the median azimuthal velocity, (b) the standard deviation of the velocity in the $z$-direction, (c) the total stellar mass, and (d) the median stellar age. These properties are sampled for stars at $z = 0$, over the radial range $0 < R_{\rm G}/{\rm kpc} < 20$, and once again exclude stars with halo-like kinematics following the methods described in Section 2. The figure shows that, whilst many of the galaxies contain no dichotomous chemical distribution, there are still very predictable overall trends which are shared with the MW. The low-$\alpha$ selections rotate at higher velocities (with Au-26 being the one exception), are kinematically cooler in the vertical direction, and are younger by 2 Gyr or more. These results are an inevitable outcome of the progressive enrichment in MW-mass galaxies, whereby

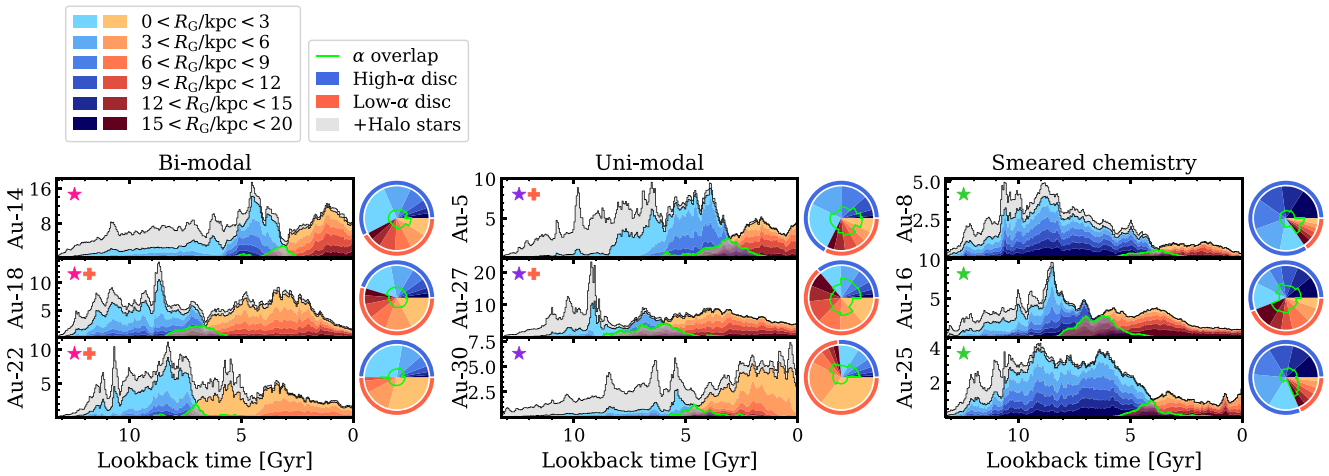


**Figure 4.** Included are a selection of AURIGA realizations with bimodal (column 1), unimodal (column 2), and smeared chemistry (column 3). The entire suite can be seen in Appendix A. Each panel shows the star formation history in units of $M_\odot$ yr$^{-1}$. The blue/left and red/right histograms are based on the high- and low-$\alpha$ selections defined in Section 3.1.1. These selections are subdivided into stacked histograms corresponding to distinct radial bands (based on the stellar positions at $z = 0$), as described in the left legend. The total overlapping region is highlighted with a green outline, which shows the contribution of stars that formed coevally. A lighter histogram shows the additional contribution of stars on halo-like orbits. The pie charts on the right-hand sides of each panel show the total mass ratios within each radial band. The coeval fraction is once again represented with a green line, but here it is calculated per radial band.

high-$\alpha$ stars are born at higher lookback times. The stellar disc at those times is often rotating at a lower velocity and was kinematically hotter, and/or has been subjected to a greater amount of kinematic heating (e.g. mergers and secular evolution).

We include reference values for the MW in black. The kinematic properties are taken from APOGEE and *Gaia* DR2 data in Anguiano et al. (2020), who decomposed the Galaxy based on chemical selections of the thick and thin discs. In comparison, the AURIGA simulations possess stellar discs that are kinematically hotter. This may owe to the numerical limitations inherent to simulations that resolve stellar distributions with discrete mass elements (e.g. particle noise), and cosmological simulations often struggle in reproducing the cooler kinematics observed in the MW (e.g. House et al. 2011; Sanderson et al. 2020). The mass ratio between our high- and low-$\alpha$ selections has a near random scatter both above and below the 1 : 1 dashed-grey line. We compare against the high- to low-$\alpha$ mass ratio estimate of 4 : 6 from the work of Khoperskov et al. (2025), which is based upon an orbit superposition approach. We compare the median age distributions of each sequence with loose predictions that are cited throughout the academic literature. The comparative late formation of the AURIGA sequences may either be an indication that the MW itself is an outlier with an especially front-loaded assembly history, or that the AURIGA chemical evolution favours later disc assembly.

A key result in panel (d) is the wide variation of formation times for both high- and low-$\alpha$ sequences. The formation of a bimodality can then be considered to be the time when a low-$\alpha$ sequence has grown a significant mass fraction, which is qualitatively similar to the median low-$\alpha$ formation times shown here. This highlights that chemical bimodalities depend on the unique assembly and star formation history of each galaxy rather than intrinsic universal formation epochs, a topic we explore in the next section.

### 3.2 The time formation of chemical sequences

The star formation histories for a selection of AURIGA galaxies are presented in Fig. 4, where we consider examples that have bimodal, unimodal, and smeared chemistry. The full suite can be found in Appendix A. Each star formation history is represented with two overlaid histograms for the high-$\alpha$ (blue) and low-$\alpha$ (red) selections, including all stars in the range $0 < R_G/\text{kpc} < 20$ at $z = 0$.

The lookback times where the two distributions overlap is outlined in green. Whilst many realizations appear to have a significant fraction of coeval stars, these are not necessarily indicative of a coeval formation of the high- and low-$\alpha$ sequences. In fact, the formation of two chemical sequences in AURIGA progresses mostly sequentially in all cases. Instead, the fraction of coeval stars are primarily made up of the intermediate-$\alpha$ abundances between the higher and lower sequences. Regardless, it is immediately evident that there are a negligible number of stars in the low-$\alpha$ selections forming at $\tau > 9$ Gyr, similar to what is observed in the MW (though see Borbolato et al. 2025).

These two histograms are subdivided into six stacked histograms, representing the radial regions considered throughout this work. The total mass budgets in each region are shown in the pie charts appended to the right-hand side of each panel, which once again includes a green line to show the overlap between high- and low-$\alpha$ stars (though here, the overlapping fraction is restricted to stars within each radial bin). In most cases, the stars found in the innermost radial range of $0 < R_G/\text{kpc} < 3$ form either the majority or the plurality of the total stars for both high- and low-$\alpha$ selections (though with the low-$\alpha$ stars in 'smeared chemistry' examples proving an exception). Yet, it is also common that the overlapping fraction is minimized in this innermost region. This is due to a combination of a few factors. First, the SFR is the most bursty in the inner few kpc, mainly due to its sensitive response to subhalo interactions and inflows. Secondly, gas mixing in inner regions occurs on shorter time-scales, such that there is a narrow chemical abundance range at any one time. Thirdly, there is a lower fraction of stars that migrated from other birth radii where the disc chemistry might be different (which we will show in Section 3.3.5). These effects lead to a higher definition or 'clumpy' distribution of stars in the chemical abundance plane.

We stack additional histograms in light-grey, where the thickness represents all remaining stars that do not have disc-like orbits at $z = 0$. These can be stars that formed before the development of the stellar disc, stars that naturally formed in regions outside of the disc, or disc stars that have transitioned to halo-like kinematics due to kinematic heating. In some cases, large numbers of disc stars are scattered into the halo and onto higher radial orbits during a major merger event. Signs of these major mergers can be seen with sudden and short-lived spikes in the SFR, such as Au-27 at $\tau = 9.5$ Gyr. In the most extreme cases, the scattering depletes the mass budget of stars in the high-$\alpha$ sequence, which in turn reduces the strength of any potential bimodality.

The transitionary epoch between high- and low-$\alpha$ selections are often preceded by either a period of rising and falling SFR, or a short-lived merger-induced starburst. This is especially evident in realizations with strong dichotomous chemistry, i.e. the 'bimodal' examples. We now explain how the interplay between SFR and chemical enrichment can contribute to the formation of a bimodality: a period of rapid star formation results in a pileup of stars at the current $\alpha$-enrichment of the proto-galaxy. The ensuing $\alpha$-enrichment via rapidly exploding Type-II SNe further encourages the formation of stars with raised [$\alpha$/Fe] abundances. Both of these mechanisms combined result in the formation of a high-$\alpha$ sequence. Then, as the SFR falls, the iron enrichment from Type-Ia SNe begins to take effect (with a lag of approximately 1 Gyr). The $\alpha$-to-iron ratio drops to a new level, and the low-$\alpha$ sequence begins to form. This is, in essence, a description of the $\alpha$-knee as in Tinsley (1979, 1980) and McWilliam (1997).

Whilst some 'unimodal' realizations share this SFR feature, there are circumstances that have inhibited the formation of a bimodality. For example, in Au-27, the merger-induced starburst at $\tau = 9.5$ Gyr occurred when the host galaxy had not grown much beyond 3 kpc. As a result, some chemical dichotomy is present in the innermost regions but does not reach as far as the Solar radius. The same scenario can be seen occurring for Au-12 in Appendix A. There is no significant rise and fall in the SFR outside of the galactic centre, and this is shown to greater depth in Section 3.2.1.

Across the entire suite, there is a great variation in the star formation history for realizations with bimodal and unimodal chemistry. Certainly, there are multiple routes to forming different chemical morphologies. In contrast, the 'smeared chemistry' examples display very similar characteristic histories: an early radial growth of the disc, and then a gradual reduction in the SFR until the present day. This early radial growth is a consequence of high angular momenta accretions from the cosmic web during the early galactic assembly. The relatively steady SFR over all radial bins, and the scarcity of disruptive radial merger events when compared with the full suite, results in a smooth chemical enrichment through time. As such, these realizations lack a chemical bimodality.

We focus on the star formation history of Au-18 in Fig. 5, showing only the *in-situ* stars. As in Fig. 4, there is a peak in the star formation history at $\tau = 9$ Gyr that is most prevalent in the innermost radial region, which corresponds to a starburst that is induced by the accretion of a GSE-like merger. This is one example of starburst-induced formation of the chemical thick disc that is described in Grand et al. (2018). However, these stars contribute a much lower mass fraction to the high-$\alpha$ sequence in the Solar neighbourhood ($6 < R_G/\mathrm{kpc} < 9$).

The high-$\alpha$ stars in the Solar neighbourhood arise mostly from a second SFR peak at around $\tau = 7$ Gyr. Curiously, this peak is due to the passage of an extremely minor satellite galaxy with a merger

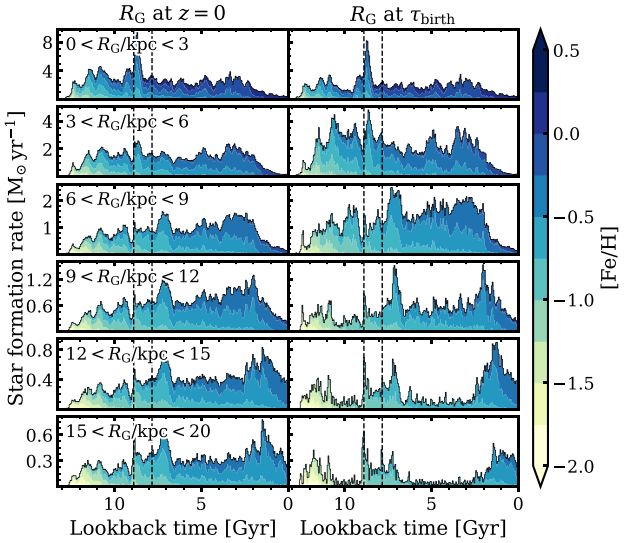


**Figure 5.** Each row shows the *in-situ* SFR of Au-18 for stars in different radial regions (as indicated in the top-left corners). The left column is for stars at their $z = 0$ positions, whereas the right column is for positions closest to the snapshot that the star was born in (approximately the birth radius). In all panels, the histograms are subdivided into stars of different metallicities, as described in the colour bar. Vertical dashed lines mark the pericentres of important mergers. Additional plots for the full suite are included online.

mass ratio of less than 1 : 100 at the time of infall, and a peak total mass of just $8 \times 10^9\,M_\odot$. It passes beyond the edge of the gas disc at $\tau = 7.8$ Gyr (at which time the disc is mostly confined to within 10 kpc) and briefly disrupts its symmetry, causing star-forming gas to flow beyond the Solar region on one side of the galaxy. This yields a delayed star formation peak that is most prevalent around the Solar neighbourhood, but specifically not at inner radii as with other merger-induced starbursts.

From the decomposition with [Fe/H], there is no obvious sign of a long-lived dilution in the metallicity, as might be expected following the two-infall model of Chiappini et al. (1997, 2001). Instead, there is a sudden and short-lived dilution corresponding to the starburst at $\tau = 9$ Gyr, though this occurs during the formation of the high-$\alpha$ sequence and is not relevant for the later formation of the low-$\alpha$ sequence. Otherwise, there is a gradual progression from metal-poor to metal-rich, with almost a dex of variation at all times and across most radial regions.

The impact of radial migration can be seen when comparing the first and second columns of Fig. 5. There is an extinction in the SFR in the two highest radial bins at birth time ($12 < R_G/\mathrm{kpc} < 15$, $15 < R_G/\mathrm{kpc} < 20$) between $7 > \tau/\mathrm{Gyr} > 2$. Before this time period, the compact gas disc is regularly disturbed by merger passages, which act to temporarily propagate star-forming gases to higher radii. For the duration of the extinction time period, the accretion history is extremely quiescent and the gas disc remains undisturbed. The few stars that do form have a low and narrow range of metallicities. By $z = 0$, this star formation valley has been repopulated by migrating stars from the inner radial regions. In turn, the range of metallicities broadens and extends to higher values. We will discuss the affects of radial migration further in Section 3.3.5. We note that this star formation valley at the highest radii does not contribute to the formation of the two $\alpha$-sequences in Au-18, as it does in some other AURIGA simulations (Grand et al. 2018).

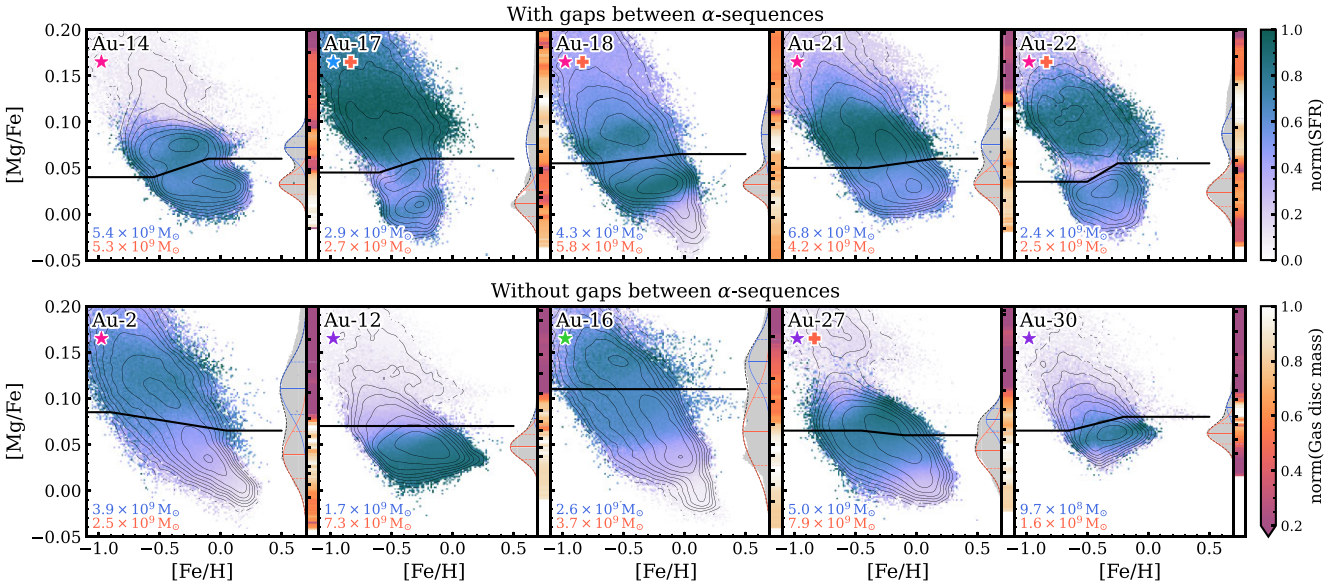


**Figure 6.** The chemical abundance planes for a selection of AURIGA galaxies, showing only disc stars in the Solar region ($6 < R_G/\mathrm{kpc} < 9$). The histograms are coloured by the mean instantaneous SFR during the birth times of each star particle, where the SFR is evaluated over 50 Myr intervals. The stellar mass distribution is overlaid with contours. Included on the right-hand side of each panel is a coloured bar that represents the mass in the gas disc at a time corresponding to the median stellar age in slices of [Mg/Fe]. This is intended to show how the gas disc varies along with the formation of $\alpha$-sequences. Median time intervals of 1 Gyr are marked with horizontal ticks. The formatting is otherwise as in Fig. 1. Upper panels: examples of galaxies which experience an under-populated 'gap' between their high- and low-$\alpha$ sequences. Lower panels: examples where there is no obvious gap. Additional plots for the full suite are included online.

### *3.2.1 The bimodality as a consequence of star formation rate*

A key feature of the chemical bimodality in the MW is that the high- and low-$\alpha$ sequences are mediated by an underpopulated 'gap', most prominent at metallicities of [Fe/H] $< -0.25$. In the AURIGA simulations, several realizations display comparable gaps between the two $\alpha$-sequences, including Au-1, Au-14, Au-17, Au-18, Au-21, Au-22, and Au-29. Both Au-1 and Au-29 are strongly affected by recent major mergers, so for clarity we opt to focus on the other examples here.

The chemical abundance planes for these simulations are shown in the upper panels of Fig. 6, where the pixels are colour-coded by the mean instantaneous SFR at the birth time of each star particle (this SFR is calculated with a histogram of the stars found within $6 < R_G/\mathrm{kpc} < 9$ at $z = 0$). This reveals that the high-$\alpha$ sequences are associated with a high SFR, which sharply declines in the chemical region corresponding to the gap. The SFR then increases again for stars within the low-$\alpha$ sequence. The exception is Au-14, where the two sequences are close in their [Mg/Fe] and so the reduced intermediary SFR cannot be seen.

There are multiple mechanisms by which the SFR can be reduced. These include a disruption of the gaseous disc by satellite fly-bys or mergers, feedback from stellar formation processes and/or AGN, or simply an exhaustion of the local gas reservoir. We select the mass in the gas disc with a crude cut of $\eta > 0.7$ on the orbital circularity and $\rho > 10^5\,\mathrm{M_\odot\,kpc^{-3}}$ on the density. We show this mass in the purple–orange bar on the right side of each panel in Fig. 6. This bar represents the normalized gas disc mass at the median birth time of the corresponding stars in each [Mg/Fe] bin. Since the evolution of the stellar [Mg/Fe] is decreasing almost monotonically with time, this effectively provides an analogue of the evolution of mass in the gas disc. Intervals of 1 Gyr are indicated using ticks along the axis.

In all cases, the reduction in the SFR during the gap is coincident with a drop in the gas disc mass. This is a result of gas exhaustion following the epoch of elevated SFR. For Au-14, the gas disc also faces violent disruption both before and during the gap due to the accretion of multiple gas-rich mergers. For Au-21, the raised SFR and gas disc mass are both caused by two simultaneous major mergers with mass ratios of approximately 1 : 2. For Au-18, there is an instantaneous drop in the gas disc mass just before the high-$\alpha$ sequence begins to form (i.e. at higher [Mg/Fe]). This is due to the previously discussed GSE-like merger briefly disrupting the gas disc. There is only a very minor rise and fall of the gas disc mass surrounding the gap itself, indicating that the two chemical sequences in Au-18 are formed by variations in the star formation efficiency rather than changes in the available gas reservoir. As we discussed in Section 3.2, star formation bursts in this realization are stimulated by the action of minor mergers at approximately $\tau = 9$ and 7 Gyr, the earliest of which is a GSE-like merger.

We draw special attention to Au-17 and Au-22, for which the gas disc both loses mass and rapidly shrinks into a compact structure less than 3 kpc wide. This results in a total suppression of star formation around the Solar neighbourhood, as can also be seen in Fig. 4. After a quiescent period of several Gyrs, the disc is partially re-established with gases from the CGM and from low-mass mergers. This mechanism was first described in Grand et al. (2018), for which the higher resolution variants of Au-6 and Au-23 are found to form multiple chemical sequences through this same process. Interestingly, we do not find clearly distinct chemical sequences in Au-6 at the lower resolution that is examined in this work, indicating that these mechanisms are sensitive to subtle variations in resolution.

In the lower panels, we show examples of AURIGA simulations that do not exhibit strong gaps between the high- and low-$\alpha$ sequences. In the case of Au-2 and Au-16, transitions in the SFR correlate with

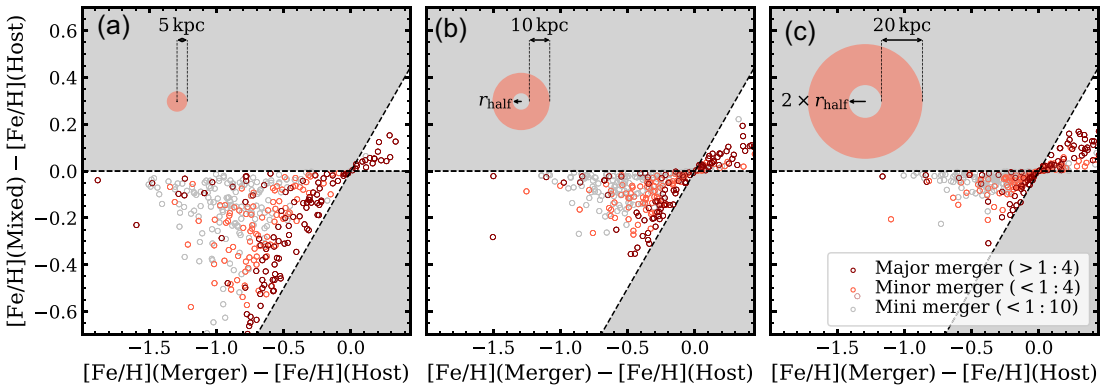


**Figure 7.** On the $x$-axis, the difference between the gas metallicity of merging subhaloes (determined at the snapshot before crossing over the virial threshold of the host halo) and the gas metallicity of the host galaxy (the main progenitor galaxy, at the same time as the merging subhalo). On the $y$-axis, the difference between the 'mixed' gas metallicity and the gas metallicity of the host, where the mixed metallicity is calculated differently for the first, second, and third panels: (a) the merger gas is assumed to mix with the host gas in the inner 5 kpc, (b) with the gas contained between the stellar $r_{\rm half}$ and $r_{\rm half} + 10$ kpc, and (c) with the gas contained between $2 \times r_{\rm half}$ and $2 \times r_{\rm half} + 20$ kpc. These mixing volumes are illustrated with schematics in the top-left corners. Mergers of different mass ratios are indicated with separate colours, down to a minimum merger gas mass of $10^8$ M$_\odot$. Forbidden regions are shaded.

bumps in the [Mg/Fe] histograms (shown on the right side of each panel). However, these transitions are more gradual, and the SFR continues to decline during the formation of the low-$\alpha$ sequence. In the case of Au-27, Au-12, and Au-30, the SFR was very low in ancient times, and so the conditions leading to an SFR gap never arose.

In general, different overdensities in the chemical abundance plane are well correlated with step-like transitions in the SFR. As such, the formation of chemical bimodalities are dependent on the star formation history of the host galaxy, which in turn is sensitive to the stochastic assembly history. This, then, explains why the formation times of chemical sequences in Fig. 3 display such great diversity.

### 3.3 How do the high- and low-$\alpha$ sequences arise?

#### *3.3.1 Merger gas dilution*

One key feature of the MW is that the high- and low-$\alpha$ sequences share a substantial overlap in their [Fe/H] chemistry around the Solar neighbourhood. This would seem contrary to the expectation that the Galactic gas metallicity is continuously enriched by SNe. As discussed in Introduction, one explanation is that the Galaxy underwent two periods of gas infall, or the 'two-infall model' (Chiappini et al. 1997). In this model, the Galactic metallicity is reset to a less-enriched state by a second infall of primordial gas. In recent years, this infall has been frequently attributed to the GSE merger event, which is predicted to have been a massive gas-rich merger that occurred around the same time as the formation of the high-$\alpha$ sequence ($\approx$ 10 Gyr ago). Here, we investigate whether comparable merger events in AURIGA can yield a similar effect.

GSE-like mergers could be important for the creation of a low-$\alpha$ sequence by providing gaseous fuel and by diluting the gas metallicity. At this point, we stress that we are focused on the kinds of gas dilution that can explain the overlapping metallicities between two $\alpha$-sequences as is observed in the MW. Therefore, the gas dilution must be sufficient long-lived that a significant number of stars are formed, must be of at least a magnitude $\approx 0.5$ dex, and must apply to the Solar neighbourhood.

To further explore these possibilities, we create a toy model for the metallicity dilution of all subhaloes across the AURIGA suite with a gas content of $10^8$ M$_\odot$ or more. The most straightforward estimate would be to calculate the combined [Fe/H] ratio for both the host and the merging subhalo. However, the gas from a subhalo may not mix evenly with all of the gas in the host galaxy. The dilution effect will vary depending on whether the merger is able to effectively penetrate the hot gaseous halo and deliver its gas directly into the galactic centre, or if its gas is stripped at higher radii and accretes onto the edge of the star-forming gas disc (which is what happens in the higher resolution version of Au-23 in Grand et al. 2018). These different outcomes will depend on the infall trajectory of the subhalo, on the ram-pressure stripping, on the respective matter density distributions, and more. Then, the magnitude of the chemical dilution will depend on four different properties; (i) the gas mass of the merger, (ii) the mass of host gas that is diluted, (iii) the metallicity of the merger gas, and (iv) the metallicity, and likewise the metallicity gradient, in the host gas.

In an attempt to account for such possibilities, we consider three different mixing scenarios for the donated gas:

(i) The donated gas mixes with the central host gas (within 5 kpc).
(ii) The donated gas mixes with the host gas around the Solar neighbourhood (between the stellar half mass radius at the time of the merger ($r_{\rm half}$), and $r_{\rm half} + 10$ kpc).
(iii) The donated gas mixes with the host gas towards the edge of the gas disc (between $2 \times r_{\rm half}$, and $2 \times r_{\rm half} + 10$ kpc).

These toy scenarios, especially (i), assume an improbably small mixing volume. This is by design, and is intended to maximize the magnitude of the gas dilution.

We show the results in Fig. 7, where on the $x$-axis we plot the difference between the mass-weighted average metallicity for the merger and host gas, and on the $y$-axis we plot the metallicity difference between the new 'mixed' gas and the original host gas. The mergers are divided across three groups corresponding to their pre-infall merger mass ratio; major ($> 1:4$), minor ($< 1:4$), and mini ($< 1:10$).

Whilst major mergers have the highest proportional gas masses, they also tend to be more chemically enriched. The effect of this can be seen in Fig. 7; the major merger scatter points tend to group-up at lower values of [Fe/H](Merger) − [Fe/H](Host). There are some mergers with gas that is more enriched than that of the host. In the cases of panels (b) and (c), this is primarily because the mixing volume excludes the central enriched gas of the host. However, there

are still some examples of super-enriched mergers in panel (a), and these are extremely ancient mergers ($\tau > 12$ Gyr) that come from a time when merger mass ratios are typically much higher, and when galaxy formation was more stochastic.

The dilution is sensitive to the gas mixing volume, and this is because of the gas metallicity gradient in the host galaxy. The galactic centre is the most enriched, meaning it has a greater potential for dilution, and this is the only scenario that sees dilutions of 0.5 dex or more. In comparison, the gas at higher radii is far less enriched, and there is often no opportunity for dilution at all ([Fe/H](Merger) − [Fe/H](Host) $\simeq 0$).

Some realizations may experience multiple gas-rich and metal-poor mergers, and this would lead to greater combined dilution. One such example is Au-14, which undergoes three minor mergers at $8 > \tau/\mathrm{Gyr} > 6$ followed by three mini mergers at $6 > \tau/\mathrm{Gyr} > 5$. The combined gas supplied by these merging groups is $1.67 \times 10^{10}$ and $1.07 \times 10^{10}$ $\mathrm{M}_\odot$, respectively, and has important repercussions on the formation of the stellar disc. However, Au-14 is a late-forming galaxy and all of these mergers occur before or during the emergence of the high-$\alpha$ sequence, and are therefore not responsible for a metallicity dilution between the formations of the high- and low-$\alpha$ sequences.

Metallicity dilution can also result from the inward transport of metal-poor gas located at the edge of the gas disc, driven inward by gravitational torques from interacting subhaloes. This mechanism is investigated by Bustamante et al. (2018), who find that major mergers are the most effective at inducing such gas inflows. However, the resulting metallicity dilution is around 0.1 dex for major mergers, and even lower for minor mergers. None the less, the dilution from this mechanism should be considered alongside our toy mixing model.

### 3.3.2 *GSE-like merger gas donation*

To determine how many stars form from merger-donated gas, we analyse a novel reproduction of Au-24 in which we have activated Monte Carlo tracer particles as in Genel et al. (2013). These track the gas flow and enables us to determine where star-forming gas originated. This realization was chosen because there is a merger with an infall time of $\tau = 9.1$ Gyr and a mass ratio of 1 : 8, broadly comparable to expectations of the GSE. This merger brings in a gas mass of $9.3 \times 10^9$ $\mathrm{M}_\odot$ with an average [Fe/H](gas) $= -1.1$.

In the left panel of Fig. 8, we show the fraction of stars across the chemical abundance plane that formed from the gas associated with this GSE-like merger. The chemistry of the pre-merger gas is overlaid with black contours. The fraction of stars that formed from the merger gas is $\sim 10$ per cent for stars in the locus of the high-$\alpha$ sequence, and then drops to $< 10$ per cent for stars in the locus of the low-$\alpha$ sequence. We also note that there are no other mergers in Au-24 which contribute a greater overall fraction than this. We have excluded the inner 3 kpc from this comparison, and this is because there is a higher metallicity group of stars with a fraction of $\approx 20$ per cent in this region. This tendency is typical of GSE-like mergers, as the low impact parameters naturally funnels a higher proportion of their gas toward the galactic centre, where metal enrichment is naturally high.

The fractional variation across the chemical abundance plane is relatively low, indicating that the merger gas rapidly becomes well mixed with the host gas prior to forming stars. There is a region at [Mg/Fe] $\approx 0.17$ in which the fraction is around 15 per cent, which represents the proportionally low number of stars that form before the merger gas is well mixed. These stars share a similar chemistry to that of the merger gas.

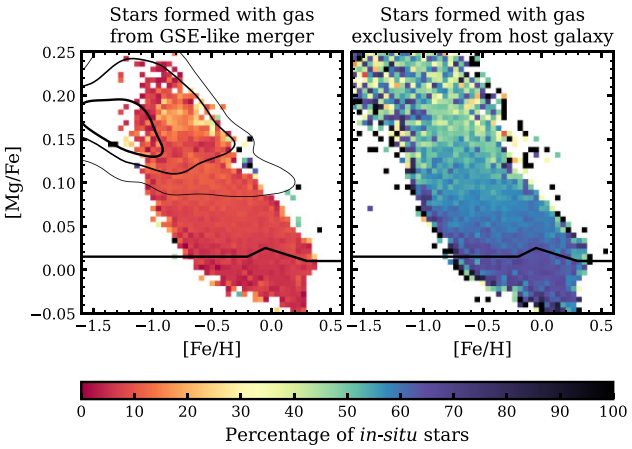


**Figure 8.** Left panel: a 2D histogram showing the percentage of *in-situ* stars which formed from the gas that was donated by the GSE-like merger in Au-24. Overlayed are contours showing the $1\sigma$–$3\sigma$ limits of the merger gas chemical abundances, taken at a time before infall. Right panel: the same 2D histogram, but this time showing the percentage of stars that formed from the host gas (i.e. gas that was not donated from any merger above a merger mass ratio of 1 : 30, but rather smooth accretion from the CGM). All stars are from the radial range $3 < R_\mathrm{G}/\mathrm{kpc} < 20$ at $z = 0$, as this is the range where the chemical sequences are most prevalent. The thick solid line marks the division between the high- and low-$\alpha$ selections.

In the right panel of Fig. 8, we show the mass fraction of stars that formed from the host gas. This is defined as gas that was not bound to mergers of mass ratio $> 1 : 30$, and should therefore represent stars that formed from local gas, gas that accreted from the smooth CGM, or gas from a spectrum of mini mergers. The result is that the disc chemistry is comprised from host gas at around 60 per cent or higher, where the remaining fraction is comprised of gas from the GSE-like merger (left panel) and several other massive mergers (not shown).

In Fig. 9, we show the evolution of the mass-weighted gas metallicity in the CGM in radial bins out to 200 kpc (upper panels) and for lines of constant radius (lower panels). We select Au-18 and Au-27 for this comparison, where both realizations experience a GSE-like merger but only Au-18 exhibits a chemical dichotomy in its disc stars. The trends shown here are representative of the broader behaviour observed across most of the AURIGA suite.

Au-18 experiences a GSE-like merger with an infall time of $\tau = 9.3$ Gyr and a mass ratio of 1 : 7. This merger brings $7.7 \times 10^9$ $\mathrm{M}_\odot$ of gas with [Fe/H](gas) $= -1.17$, a full 0.66 dex lower than the gas metallicity in the inner 10 kpc of the host galaxy that it merges with. This same realization was previously investigated in Ciucă et al. (2024), where it is shown that the merger causes a sudden but short-lived drop in the stellar metallicity in Solar neighbourhood stars of $\approx 0.25$ dex (see their fig. 4).

We see a much lower gas dilution in our analysis, where a dilution of up to $\approx 0.25$ dex is only achieved in the galactic centre (3 kpc). There, the dilution is short-lived and is entirely recovered within a few 100 Myr. As we will consider later in Section 3.3.5, some of the stars that form from this gas will ultimately migrate into the Solar neighbourhood and will therefore contribute to the feature described in Ciucă et al. (2024). None the less, this metal dilution is not relevant for the overlapping [Fe/H] values of the two chemical sequences in Au-18. This is because the dilution occurs during the formation of the high-$\alpha$ sequence, and the gas metallicity has entirely returned to

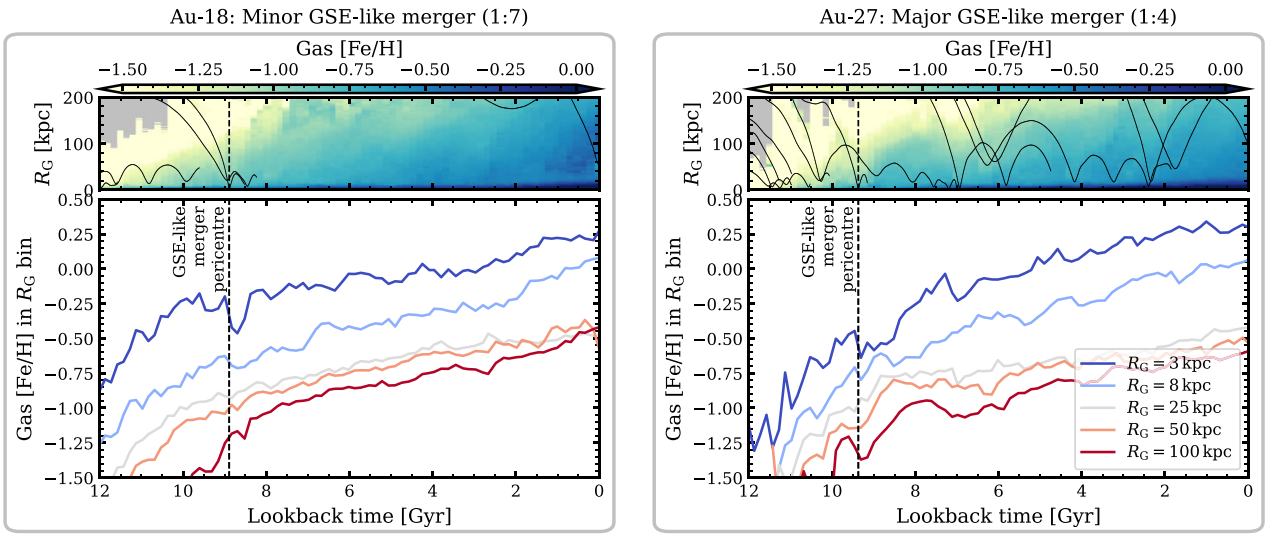


**Figure 9.** Upper panels: the mass-weighted metallicity of the gas surrounding the main progenitor galaxy, shown with galactocentric radius versus lookback time. The infall paths of merging galaxies are plotted with solid lines (merger mass ratios $> 1:20$). The bin size is 0.4 kpc. Lower panel: the mass-weighted gas metallicity at a selection of constant radii, as indicated in the legend. The verticle dashed line marks the pericentre time of a GSE-like merger. Additional plots for the full suite are included online.

pre-dilution levels before the time period that the low-$\alpha$ sequence forms.

The GSE-like merger in Au-27 has an infall time of $\tau = 9.7$ Gyr and a higher mass ratio in excess of $1:4$. This merger brings $1.6 \times 10^{10}\,\mathrm{M}_\odot$ of gas with [Fe/H](gas) $= -0.98$, which is 0.41 dex lower than the gas metallicity in the inner 10 kpc of the corresponding host galaxy. Despite the greater quantity of donated gas, the only perceptible metallicity dilution occurs at 3 kpc and is of a magnitude 0.2 dex. In both realizations considered here, the metallicity of galactic gas around the Solar neighbourhood (the 8 kpc line) is minimally impacted by the GSE-like merger accretion.

These results demonstrate the same point made in Fig. 8: the gas donated by a GSE-like merger is not the reason why low-$\alpha$ sequences overlap with the metallicity of high-$\alpha$ sequences. Whilst there is a measurable metallicity dilution, it is not of a great enough magnitude to explain the overlapping chemical sequences. This is not to say that the merger is unimportant in other aspects, such as its effect on the SFR within the host galaxy, its kinematic influence on the gas (e.g. Merrow et al. 2024), or on pre-existing disc stars (Orkney & Laporte 2025). The figure also reveals another significant detail: there is a reservoir of consistently metal-poor gas in the wider CGM (beyond 25 kpc), which remains at [Fe/H] $< -0.5$ even up to $z = 0$. We will discuss how this can impact the development of the low-$\alpha$ sequence, next.

### *3.3.3 Chemical variation across the disc.*

If gas donation from mergers is relatively unimportant for the creation of $\alpha$-sequences in AURIGA, then why do some AURIGA realizations manifest dichotomous chemical distributions with substantial overlaps in [Fe/H] (e.g. Au-14 and Au-18 in Fig. 1)?

To investigate this further, we extract the precise birth locations of stellar particles in one example AURIGA realization, whilst noting that the resulting trends are general across the suite. Then, we derive the galactocentric radius and height above or below the disc plane for all disc stars at the time of their birth. In Fig. 10, we show normalized Gaussian kernel density estimates for the stars that formed in the high-$\alpha$ (top panels) and low-$\alpha$ (bottom panels) sequences.

The left set of panels shows the stars as partitioned by their birth time, demonstrating the typical inside–out (first column) and top–down (second column) formation that has been widely predicted for galaxy formation. The right set of panels shows the stars as partitioned by their metallicities. This reveals consistent trends for both high- and low-$\alpha$ sequences; (i) metal-poor stars form at greater radii, and (ii) metal-poor stars form at greater heights above and below the disc. This is an intuitive result, galaxies naturally form a negative metallicity gradient in their gas with both radius and height. This is imprinted on the kinematic properties of the stars at $z = 0$: stars from the metal-poor ends of each sequence have slightly lower mean orbital circularities in Fig. 1.

We expand upon this result in Fig. 11, which shows histograms of the chemistry in different regions of the gas disc for Au-18 and Au-24 at times corresponding to the formation of their high- and low-$\alpha$ sequences. These histograms are shown in terms of distance in the $x - y$ plane and height above and below the disc. Black contour lines show the $1\sigma - 3\sigma$ contours of the SFR in the gas, and we emphasize that the SFR in AURIGA is a function of the gas density (see Grand et al. 2017). Here, Au-18 is chosen as an example galaxy for which there is a large [Fe/H] overlap between the two $\alpha$-sequences, and Au-24 is chosen as an example galaxy for which there is almost no overlap (see Fig. 1). In the upper panels, we include plots of the chemical abundance plane for the contemporaneous stars as a helpful reference.

During the formation of the high-$\alpha$ sequences, the gas discs in both examples are thick and flared, and so there is star formation occurring up to and beyond 10 kpc above and below the disc plane. The central gas disc is enriched with metals, but the surrounding gases are approaching metal-poor values of $\langle$[Fe/H]$\rangle \simeq -0.8$.

During the formation of the low-$\alpha$ sequence, a difference emerges between the two examples. The gas in Au-18 is now less flared, but there are still dense metal-poor gases at heights up to $\sim$ 10 kpc above and below the disc plane, leading to continued star-forming

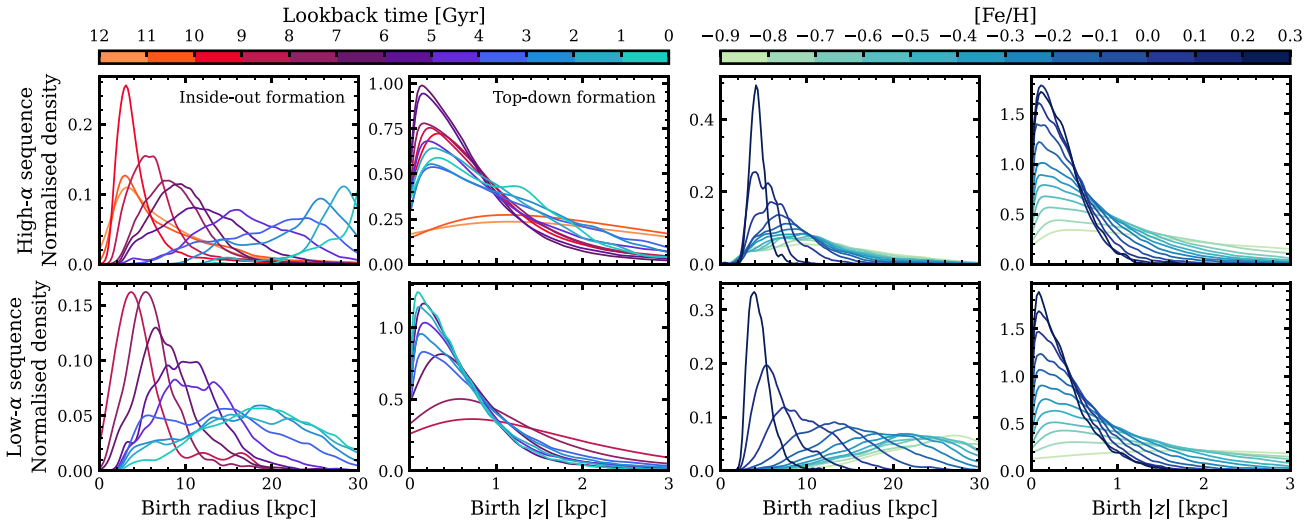


**Figure 10.** This plot shows the normalized density functions of stars as a function of birth galactocentric radius and birth height above/below the disc plane in one example realization. The left four panels segregate the stars by their birth time, and the right four panels do so by their [Fe/H]. The upper panels show stars within the high-$\alpha$ selection, whereas the lower panels show stars within the low-$\alpha$ selection. The trends here are seen across all realizations, no matter what their chemical sequences look like.

activity in those regions. On the other hand, the star formation in Au-24 is almost exclusively concentrated in the disc plane. Here, the gas is metal-rich and forms a tight radial metallicity gradient. There is still metal-poor gas above and below the disc plane, but it is no longer sufficiently dense to form stars. The explanation is that the gas disc has become stable at late times, ultimately collapsing into a thin plane in which the metal-poor and metal-rich gases are mixed together.

From closer examination of Au-18 in Fig. 1, there is a faint tail of stars below the main overdensity of the low-$\alpha$ sequence, and these have an extremely narrow metallicity range of $\Delta$[Fe/H] $\approx 0.2$. This tail is formed entirely from stars in the final 2 Gyr, by which time the gas disc in Au-18 has settled into a configuration that is almost as thin and stable as in Au-24. We show this epoch in the column titled 'Low-$\alpha$ epoch II'. The behaviour is seen across the whole simulation suite: a narrow metal distribution is correlated with a thinner gas disc, whereas a wide metal spread is correlated with a thicker gas disc. Then, it is apparent that a wide spread of metal values in these simulations is dependent on the presence of dense metal-poor gases at elevated heights above and below the disc plane. Therefore, we argue that [Fe/H] overlaps in AURIGA come not from metallicity dilution following a merger event, but from a natural variation in gas chemistry and its geometric distribution throughout time.

In the lower panels of Fig. 11, we include the same plots but now weighted by the gaseous [Mg/Fe] abundance ratio. When the [Fe/H] and [Mg/Fe] gas abundance ratios are considered together, we can explain the radial dependence of the chemical sequences shown in the upper set of panels. In both examples, there are modest [Fe/H] gradients and [Mg/Fe] gradients during the formation of the high-$\alpha$ sequence. This leads to $\alpha$-sequences with chemistry that is strongly radially dependent. Both examples have a metallicity gradient during the 'low-$\alpha$ epoch I', but only Au-24 has a corresponding gradient in its [Mg/Fe] abundances. The outcome is that the low-$\alpha$ sequence of Au-18 has a radial dependence on [Fe/H] but not on [Mg/Fe]. For comparison, the $\alpha$-sequences in the MW show no radial evolution for the high-$\alpha$ sequence, and evolution in both [Fe/H] and [Mg/Fe] for the low-$\alpha$ sequence (see fig. 10 of Imig et al. 2023).

### *3.3.4 The shape of chemical sequences*

In Fig. 1, there are sequences that form a flat pancake (fixed [Mg/Fe] over a range of [Fe/H]; Au-14 and Au-18) and sequences that slant down from left-to-right (an inverse relation between [Fe/H] and [Mg/Fe]; Au-2 and Au-24). We find that the slant of chemical sequences is correlated with several properties of the gas disc, as was already seen for two examples in Fig. 11. There, a thinner and more radially extended gas disc is associated with a downward slanted chemical sequence.

We show how the slope of low-$\alpha$ sequences relates to the height of the gas disc across the whole AURIGA suite in Fig. 12. Here, we have chosen representative snapshots of the low-$\alpha$ formation epochs that avoid moments of disruptive merger interactions. The gas disc height is defined as the average vertical distance within which more than 10 per cent of the gas is denser than the AURIGA star formation threshold of $\rho_{\rm SF} = 0.13\, m_{\rm p}\, {\rm cm}^{-3}$. We calculate this value in a series of cylindrical bins corresponding to the radial cuts used throughout this work. This definition is designed to accommodate for the very clumpy gas distribution often found at greater vertical heights, which would be ignored in a volume-averaged gas density. Then, the slant of the chemical sequence is determined as the gradient of a first-order polynomial fit to the stars that formed at a birth time interval of $\tau_{\rm epoch} \pm 1$ Gyr. Our stellar selection is based on the $z = 0$ properties.

The figure shows a clear relation: thinner gas discs are associated with low-$\alpha$ sequences that have a negative relation between [Mg/Fe] and [Fe/H]. These relations become increasingly negative as we sample stars further toward the galactic centres, because the gaseous abundance gradients are steepest there.

We attribute the origin of this relation to long-lived trends in the SFR. The $\alpha$-elements are produced by Type-II SNe of short-lived massive stars, which rapidly enrich the ISM. Iron is mainly produced by the Type-Ia SNe of longer lived binary stars, and can take $\sim 1$ Gyr to enrich the ISM. Therefore, an increasing SFR will raise or maintain the [Mg/Fe] ratio because the iron enrichment is lagging behind the $\alpha$-enrichment. An SFR that is decreasing or unchanging over time-scales of a Gyr or more will drive the same ratio down. Therefore,

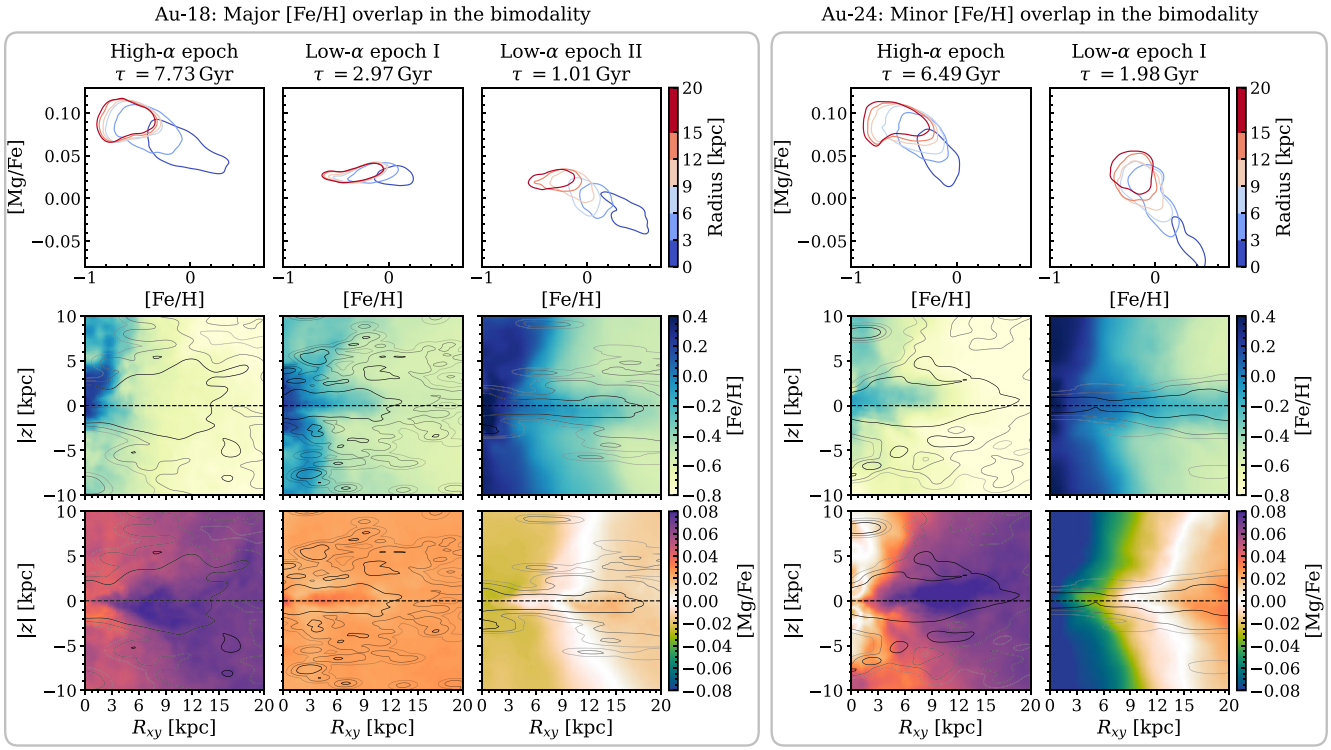


**Figure 11.** This figure shows the chemistry of the galactic gas and stars at various critical epochs for two example realizations, Au-18 and Au-24. Upper panels: the $1\sigma$ contour encircling the stars that formed in each time epoch (where the time epoch is indicated in the title, and stars are selected from a birth time interval of $\tau \pm 0.5$ Gyr). The stars are grouped by their $z = 0$ orbital radius, as indicated in the colour bar. Middle panels: mass-weighted [Fe/H] abundance ratio for galactic gas in the $Rxy$ versus $|z|$ plane. The overlaid contours indicate the $1\sigma$–$3\sigma$ of the instantaneous SFR in the gas cells. This roughly corresponds with the gas density, due to the star formation criteria used in AURIGA (Grand et al. 2024). Bottom panels: the same, but now for the [Mg/Fe] abundance ratio.

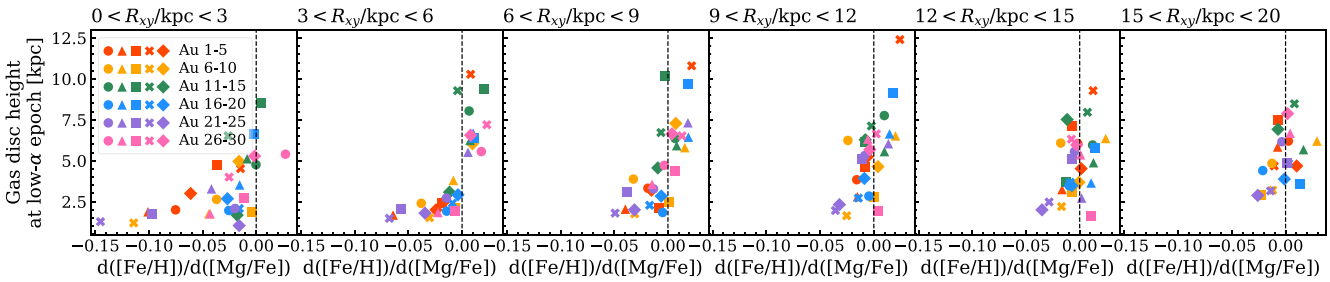


**Figure 12.** An estimate of the gas disc height versus the gradient of low-$\alpha$ sequences in the chemical abundance plane, for a series of cylindrical bins. The gas is selected from a single representative lookback time, whereas stellar properties are gathered at $z = 0$ from a birth time interval of $\tau \pm 1$ Gyr.

chemical sequences with steep downward 'slants' should correlate with an unchanging or decreasing SFR.

We provide evidence that this occurs in AURIGA in Fig. 13, which shows the mean gradient of the radially binned SFR in the chemical abundance plane for a selection of realizations. Here, we are calculating the SFR for stars that were born in the range $6 < R_G/\mathrm{kpc} < 9$, though the same results are recovered when considering stars found in the same radial range at $z = 0$. The upper five panels show examples where the $\alpha$-sequences are flat. There, we see that the SFR is rising for both high- and low-$\alpha$ sequences. For realizations with bimodal chemistry (see the pink coloured markers in the top-left), we can also see that the SFR declines for the underpopulated region between the two sequences. This is similar to the result shown in Fig. 6.

The lower five panels show examples where the $\alpha$-sequences are 'slanted' from top-left to bottom-right (an inverse relation between [Fe/H] and [Mg/Fe]). Each example has an initially growing SFR at high-$\alpha$, corresponding to the growth of the proto-galaxy at early times. Then, the SFR is either unchanging or decreasing throughout the formation of the low-$\alpha$ sequence. These trends are consistent across the rest of the AURIGA suite and for other radial ranges not shown here. This behaviour is the same as is reported in the EAGLE simulation suite by Mason et al. (2024), in which the formation of 'knees' in the [$\alpha$/Fe] versus [Fe/H] plane are attributed to a sustained decline in the SFR.

Considering both Figs 12 and 13 together helps to build a more complete picture of the processes behind the morphology of $\alpha$-sequences. A declining SFR means a reduction in turbulent stellar

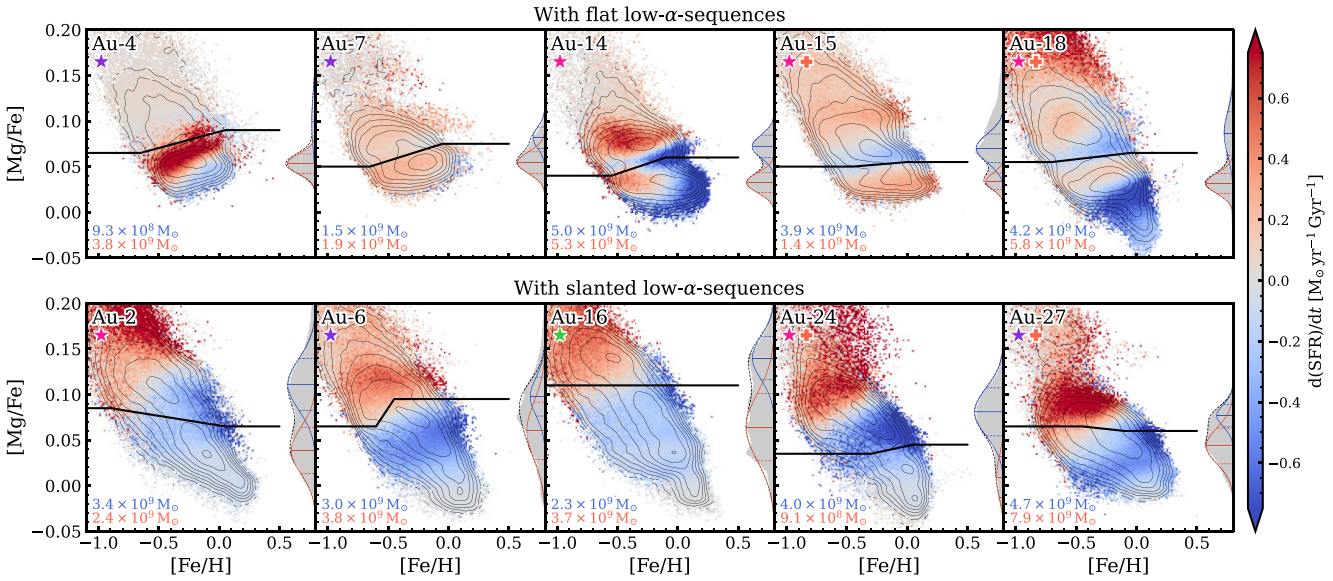


**Figure 13.** The chemical abundance planes for a selection of AURIGA galaxies, showing only disc stars that were born in the Solar region ($6 < R_G/\mathrm{kpc} < 9$). The histograms show the gradient of the local SFR associated with each star particle, evaluated over a time period of $\tau_{\mathrm{birth}} \pm 1$ Gyr. We calculate the SFR for stars that were born in the Solar region rather than stars found in the Solar region at $z = 0$. This is so that the gradient of the SFR better reflects the birth environment of each star particle, which may change due to radial mixing and migration. The stellar mass distribution is overlaid with contours. The formatting is otherwise as in Fig. 1. Upper panels: examples of galaxies with a flat distribution in their $\alpha$-sequences. Lower panels: examples where there is a downward slanted distribution in their $\alpha$-sequences. Additional plots for the full suite are included online.

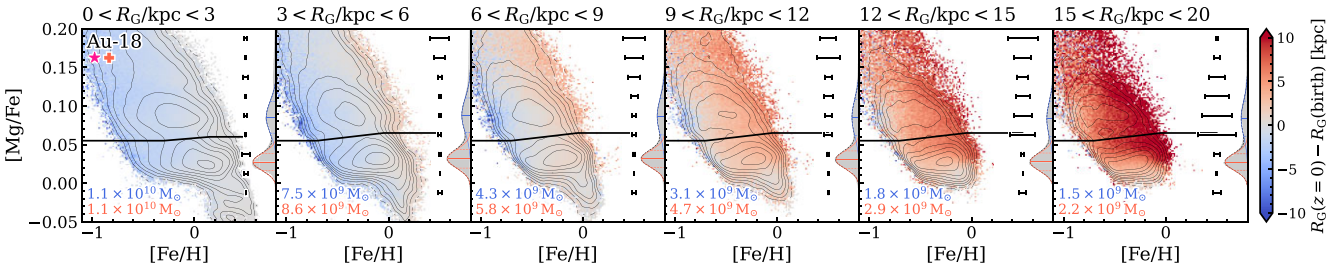


**Figure 14.** In this plot, we show the radial migration of stars in the chemical abundance plane for one example galaxy (Au-18). The histograms show the mean change in orbital radius, from the closest birth snapshot to $z = 0$. The stellar mass distribution is overlaid with contours. Error bars to the right of each panel illustrate how the spread of the [Fe/H] distribution has increased due to radial migration, calculated as the difference between the 5th and 95th mass-weighted percentiles for stars at their birth radii and at their $z = 0$ radii. Additional plots for the full suite are included online.

feedback processes, which enables the development of a thin and stable gas disc.

### *3.3.5 Radial migration*

A potential complication in the formation of galactic chemical sequences is the tendency for stars to migrate away from their initial birth sites (radial migration). This is caused by a change in their orbital angular momenta, originating from interactions with non-axisymmetric disc structure (e.g. spiral arms, Sellwood & Binney 2002; Roškar et al. 2008, central bars, Minchev & Famaey 2010; Minchev et al. 2011, or mergers and subhaloes, Minchev et al. 2014b; Carr et al. 2022).

The impact of radial migration on the AURIGA level-3 simulations was previously discussed in Grand et al. (2018), where it was found to have little importance for the development of chemical sequences. This is consistent with idealized simulations that show radial migration is not the dominant factor with respect to the bimodality (Minchev et al. 2014a). For completeness, we have performed our own analysis on all of the level-4 simulations, for which we show just one example in Fig. 14.

This figure is generally representative of the rest of the suite, and exhibits a few key features. First, there is very little bulk migration occurring for the high- and low-$\alpha$ sequences around the Solar neighbourhood ($6 < R_G/\mathrm{kpc} < 9$). This shows that the bimodality at these radii does not arise through the radial migration of stars, but specifically from the gas chemistry and SFR local to the Solar neighbourhood.

Secondly, the outward migration of stars becomes increasingly prevalent at regions beyond the Solar neighbourhood, with the high-$\alpha$ sequence having migrated the most. This would be because stars that formed in the high-$\alpha$ sequence have had a longer time over which to migrate, and also because this population is more likely to have exclusively formed in the compact proto-galaxy.

Thirdly, there is a strong trend whereby stars at lower [Fe/H] have migrated from higher birth radii, and stars at higher [Fe/H] have

migrated from lower birth radii. This is an inevitable outcome of the negative metallicity gradient present in the gaseous disc throughout time. Whilst this radial migration or mixing is not responsible for forming the bimodality itself, it does cause a broadening of the [Fe/H] distribution over radii near the Solar region. We include horizontal black error bars to show the magnitude of this broadening. In the Solar region, the broadening is only significant for the most ancient high-$\alpha$ stars.

This result is qualitatively in agreement with the broadening of [Fe/H] distributions reported in idealized simulations of an MW-mass galaxy in Grand, Kawata & Cropper (2015), which was also caused by radial migration processes. Similarly, Haywood et al. (2013) use observational data to argue that radial migration would not have significantly contaminated the Solar neighbourhood in the MW.

## 4 DISCUSSION

Since its discovery, the origins of the chemical bimodality have remained elusive. None the less, the wealth of observational data in recent times have provided us with important clues. In this work, our intention has been to use the AURIGA simulation suite to explore some of these clues in detail using a representative statistical sample of MW-like realizations. From this, we were able to gain a better understanding of the kinds of processes responsible for forming chemical sequences, and more specifically the most likely scenario for our own Galaxy. This provides insight into not only the chemical evolution of the MW, but of its overall formation history too. Furthermore, it is relevant for the many other disc galaxies (which may or may not resemble the chemical structure of the MW) that will be uncovered by future spectroscopic surveys of the Local Volume in the era of Extremely Large Telescopes.

Here, we discuss the implications of our results on the formation time of the two $\alpha$-sequences in the MW, the roles that the GSE merger may or may not have played, and provide our own interpretation for the timeline of events that built the MW disc and other notable stellar populations.

### 4.1 A consecutive formation of the chemical sequences

We have found that, in all cases except for one unusual outlier (Au-1), the formation of $\alpha$-sequences is sequential. Any apparent temporal overlap between two sequences is due to the spread in chemical abundances for mono-age populations rather than a faithful coeval formation. This, then, is accurate to first order to what is observed in the MW, with many studies finding that the high-$\alpha$ sequence formed rapidly before $\sim$ 8 Gyr ago and the low-$\alpha$ sequence formed over a much longer period of time thereafter (e.g. Fuhrmann 1998, 2011; Bensby et al. 2005; Haywood et al. 2013; Hayden et al. 2017; Mackereth et al. 2017).

However, there is also observational evidence for a fraction of low-$\alpha$ stars that formed at lookback times older than 10 Gyr ago, in consort with the high-$\alpha$ sequence (e.g. Beraldo e Silva et al. 2021; Gent et al. 2024; Borbolato et al. 2025). The authors in Gent et al. (2024) show that up to 20 per cent of stars in the low-$\alpha$ sequence formed prior to 9 Gyr ago using Gaia-ESO data for main-sequence turnoff stars and subgiants. Similarly, Borbolato et al. (2025) use the STARHORSE code (Santiago et al. 2016) to determine stellar ages for Solar neighbourhood stars in both APOGEE and LAMOST spectroscopic survey data, finding the fraction of low-$\alpha$ stars with ages older than 11 Gyr to be over a third in the case of APOGEE and nearly a half in the case of LAMOST. These suggest the early Galaxy was very capable of achieving low-$\alpha$ chemistry even before the accretion of the GSE at around 10 Gyr ago.

One possible explanation is that these stars were born within intermediate-mass gas clumps that are predisposed to form at high redshifts when gas fractions are higher (see Bournaud et al. 2009; Clarke et al. 2019). The subgrid physics model in AURIGA does not resolve the formation of such gas clumps, and further studies would be needed to explore the impact of these physics in fully cosmological simulations.

Observational studies find that the high-$\alpha$ disc transitioned to almost exclusively low-$\alpha$ formation very rapidly at approximately 8–9 Gyr ago (Linden et al. 2017; Sahlholdt et al. 2022; Borbolato et al. 2025). This, then, is just 1–2 Gyr after the likely accretion time of the GSE merger (e.g. Belokurov et al. 2018; Helmi et al. 2018; Kruijssen et al. 2019; Mackereth et al. 2019). However, these investigations are often restricted to the Solar neighbourhood and/or limited stellar populations (e.g. see the review of Soderblom 2010), and by high uncertainty in stellar age determinations. Likewise, the exact accretion time of the GSE is still uncertain to within a precision of 1–2 Gyr. In this work, we find that gas from GSE-like mergers was not instrumental for the formation low-$\alpha$ sequences via their gas donation or metallicity dilution, but could still have been important in other respects. For example, a massive injection of gas from a singular direction establishes a high angular momentum in the gas, and may also help to align the disc with the angular momentum of wider scale CGM gas, which ultimately aids in the build-up of an extended stellar disc at higher radii (e.g. see the investigation into Au-18 in Merrow et al. 2024).

To better understand the timeline of disc formation, it will be crucial to obtain stellar age estimates with higher precisions and for wider samples. This will be addressed by upcoming asteroseismology on thousands of stars in ESA's half-sky *PLATO* survey (Rauer et al. 2025), or for a million stars in the further future by the all-sky *Chronos* mission (Michel et al. 2021).

### 4.2 The possible roles of the GSE merger

With the ambiguity surrounding the early formation history of the MW, the GSE merger event stands out as an isolated pillar of certainty. Therefore, it is natural to develop explanations that revolve around this event. The estimated infall time of the GSE is remarkably close to the estimated transition time between the two $\alpha$-sequences, suggesting that it may be responsible in some way.

#### *4.2.1 Gas fuelling and dilution*

The GSE would be an opportune vehicle for the donation of fresh metal-poor gas into the ancient Galaxy, as called for by the 'two-infall' model of Chiappini et al. (1997, 2001). This could reset the gas metallicity to low values whilst also providing the fuel to support the formation of the low-$\alpha$ sequence. The chemical evolution models in Spitoni et al. (2019, 2021, 2023) show that a metallicity dilution of approximately a full dex is able to reproduce the low-$\alpha$ sequence in the MW. Improving stellar age estimates and extensions to existing catalogues have allowed for more detailed chemical evolution models, highlighting some features of the low-$\alpha$ disc that are difficult to explain with a two-infall model (see Dubay et al. 2025).

In our investigation, we have found that singular merger events do not induce a long-lived and significant ($\Delta$[Fe/H] > 0.5) dilution in the gas discs of their hosts. The metal-poor gases donated

by minor mergers are rapidly mixed into the surrounding CGM with low impact on the overall metallicity, whereas major mergers increasingly approach parity with the metallicity of their host and so rarely yield a dilution greater than $\sim 0.3$ dex. In fact, we find some realizations with GSE-like mergers that exhibit no metallicity overlap between their two $\alpha$-sequences, or do not have multiple sequences at all. Furthermore, the gas from most mergers forms only a minor component of subsequent *in-situ* disc stars (with massive major mergers of mass ratio $> 1:2$ being an exception).

This is not to suggest that mergers are entirely incapable of inducing dilution. For instance, Grand et al. (2018) demonstrated that dilution occurs at the edge of the gas disc around the time of a late gas-rich merger in the high-resolution version of Au-23. In the version of Au-23 used in this study, we observe that dilution begins prior to the arrival of the merger gas, suggesting that the merger's gravitational influence may draw metal-poor gas from the CGM onto the disc's outskirts (which is then supplemented by the merger gas itself), thereby indirectly causing dilution. None the less, this dilution is insufficient to account for the metallicity overlap observed in the MW's chemical sequences, and the properties of this merger event are inconsistent with those expected for the GSE.

We show that the greatest metal dilution occurs when metal-poor merger gases mix with the highly enriched central gas of the host, in agreement with Buck (2020) who find that merger-induced metal dilution in the NIHAO simulations was maximized in the inner 3 kpc. Simulations have also shown that gravitational torques from satellites can stimulate the inflow of metal-poor CGM gases into galactic centres (Di Matteo et al. 2007; Bustamante et al. 2018; Moreno et al. 2021; Faria et al. 2025), contributing to a central starburst and further metallicity dilution. Yet, observations of the MW indicate that the metallicity overlap between the two $\alpha$-sequences is least pronounced from within $\sim 6$ kpc (e.g. Queiroz et al. 2020). Therefore, whilst these effects may have occurred in the MW, they were clearly not the primary drivers in the development of the chemical sequences. Then, it is unlikely that the metallicity overlap in the chemical sequences of the MW was caused by a highly radial infall such as the GSE, if it was caused by a merger event at all.

Contrary to our interpretations, Buck (2020) argue that metallicity dilution from gas-rich minor mergers (of less than $1:10$ mass ratio) explains the overlapping metallicities of bimodal chemical sequences. However, we show that the metal dilution following those mergers is insufficient. The metallicity overlap is as great as 0.5 dex (see their fig. 11). Yet, from inspection of their fig. 7, the metal dilution following gas-rich mergers is typically far less than half of this, and is short-lived in most cases. Then, this would be consistent with the findings presented in this work. In fact, our result is general and should hold for any reasonable galaxy formation model in a Lambda-cold dark matter Universe. In Appendix B, we demonstrate this further using simple arguments involving empirical relations (e.g. stellar halo mass relation; Moster, Naab & White 2013) and observations (e.g. the stellar mass–metallicity relation; Kirby et al. 2013).

#### 4.2.2 *Influence on the star formation rate*

In Section 3.2, we showed that AURIGA galaxies with particularly pronounced bimodal chemistry exhibit a lull in their star formation following merger-driven, centrally concentrated starbursts. Indeed, this is a behaviour also seen among other cosmological simulations like SIMBA (Rodríguez Montero et al. 2019), ILLUSTRISTNG (Hani et al. 2020), VINTERGATAN (Segovia Otero, Renaud & Agertz 2022), and FIRE-2 (Li et al. 2025). In AURIGA, the enhanced SFR has either fully or partially depleted the star-forming gas, ultimately leading to a temporary suppression of the star formation. This is reflected in observational data from SDSS (the Sloan Digital Sky Survey) finding evidence that starbursts correlate with features of merger interactions (Luo, Yang & Zhang 2014). This is a possible scenario for the MW, where the two $\alpha$-sequences are divided by an underpopulated gap, and there is a clear star formation quiescence at $10-8$ Gyr ago (e.g. Haywood et al. 2016, 2018; Spitoni et al. 2024, and see also fig. 7 in Snaith et al. 2015).

Our finding is consistent with Beane (2025), where idealized simulations show that a brief star formation hiatus of a few hundred Myr could arise due to gas starvation following a GSE-like merger (though with great sensitivity to the infall inclination of said merger). The hiatus does not necessarily need to owe to the influence of a merger, and in a follow-up work they propose AGN-induced quenching coincident with the formation of a stellar bar as another possible mechanism (Beane et al. 2025, and see also Haywood et al. 2018). Additionally, Grand et al. (2018) demonstrate than the time evolution in the size of the gas disc can lead to the formation of a chemical bimodality over the pertinent radii. Disc-galaxy evolution models in Noguchi (2018) also predict that a star formation hiatus of approximately 2 Gyr can create a bimodality. Finally, simulations in Lian et al. (2020) predict a starburst followed by rapid quenching could be responsible for forming the thick disc. These are compatible with observations of the MW from APOGEE, which finds evidence of a $\sim 2$ Gyr quenching period roughly 9 Gyr ago (see Haywood et al. 2016).

### 4.3 Fuelling from the CGM

We find that, in most cases, the plurality of gas that forms the disc comes from a smooth accretion from the CGM and/or mini mergers. This is a universal feature for galaxies that grow from the cooling of CGM gas. The exceptions are realizations that experience late-time massive major mergers with mass ratios exceeding $1:2$, but these realizations normally have unimodal chemistry, and are not considered a plausible scenario for the MW.

This is a result that is backed up by idealized simulations in Khoperskov et al. (2021), in which the low-$\alpha$ sequence forms from a mixture of metal-poor CGM and enriched gases that were ejected by stellar feedback during the formation of the high-$\alpha$ sequence. Furthermore, they find that the formation of two chemical sequences is not dependent on any one particular merger history or on radial migration.

Other fully cosmological simulations also highlight the importance of the CGM. Analysis of the FIRE-2 cosmological simulations in Parul et al. (2025) showed that bimodal sequences arise when metal-poor gas accretes onto the outskirts of the disc, owing to both merger donation and the CGM. However, their realization with a high fraction of merger-donated gases (36 per cent) is only weakly bimodal, and they conclude that merger-donated gases do not necessarily lead to the development of a distinct chemical sequence.

Investigations into the higher resolution versions of AURIGA in Grand et al. (2018) show that late-time minor mergers can penetrate the hot gas halo and deliver cold metal-poor gas directly onto the galaxy. While our own analysis shows that the merger gas neither dominates the low-$\alpha$ sequence nor reduces the metallicity by more than $\sim 0.3$ dex, it none the less plays an important role in triggering the growth of the gas disc. The merger may also incite the condensation of surrounding metal-poor CGM gas onto the disc, leading to an indirect dilution effect.

In the VINTERGATAN cosmological simulations, a disc with bimodal chemical sequences forms in the presence of a GSE-like merger (Agertz et al. 2021; Renaud et al. 2021). In particular, their low-$\alpha$ sequence exhibits a 'reset' in its metallicity. A detailed investigation revealed that the gas donated by the GSE-like merger was unimportant for both the formation of the low-$\alpha$ sequence and the metallicity dilution. Instead, they owe to an accreting metal-poor and cold gas filament that is initially misaligned with respect to the gas disc. The GSE-like merger is none the less essential because its interaction with this accreting gas filament acts to excite star formation within it. Similarly, we find that mergers are important for triggering star formation, even if their merger mass ratios are less than 1 : 100. This is a result also echoed in Parul et al. (2025).

### 4.4 Our interpretation for the MW history

Based on our research and what we have learned from our analysis of the AURIGA simulation suite, which is broadly consistent across different analytical and empirical models, we can formulate an interpretation for the history of our Galaxy.

The MW began to form early, with an inside–out and top–down formation of the stellar disc. These early stars formed during a period of intense SFR from a turbulent interstellar medium. Whilst the bulk of the high-$\alpha$ sequence formed at this time, there may have been subdominant populations of low-$\alpha$ stars forming in dense gas clumps. The proto-galaxy would have undergone a series of high mass-ratio mergers at this time, perhaps including the speculative *kraken* (i.e. Kruijssen et al. 2019; Massari, Koppelman & Helmi 2019). These mergers contribute gaseous fuel that would have been important for developing the early gas disc and initiating star formation in the high-$\alpha$ sequence, but would have been relatively unimportant in terms of their direct chemical impact (see Orkney et al. 2022 and arguments in Section 3.3).

At around 10 Gyr ago, the GSE merger accretes onto the proto-MW. Following reasoning in the literature (e.g. Kruijssen et al. 2019, 2020; Lane, Bovy & Mackereth 2023), this was a minor merger with a mass ratio $< 1:4$. This event initiates a combination of physical processes (inward funnelling of CGM gas, donation of fresh cold gas, and gas shocks) which conspire to induce a central star formation burst. This burst bolsters the number of stars in the high-$\alpha$ sequence, but also acts to deplete the star-forming gas, and is followed by a reduction of the SFR that lasts $\mathcal{O}1$ Gyr. The GSE has other implications on the fledgling MW disc, kinematically heating its stars and shaping what is now the thick disc (e.g. see Grand et al. 2020).

Assuming a variety of masses and metallicities for the GSE progenitor galaxy and contemporary proto-MW, we show in Appendix B that the GSE merger would have been incapable of driving a substantial metallicity 'reset' between the metal-rich end of the high-$\alpha$ sequence and the metal-poor end of the low-$\alpha$ sequence. Instead, continuous feeding from the CGM is responsible for maintaining a metal-poor zone at the height and radial extremities of the gas disc (see Fig. 11). The metal-poor ends of the low-$\alpha$ sequence is comprised from stars that form out of these gases.

With a relatively quiescent history thereafter, the MW disc becomes thinner and more stable, and the SFR gradually declines as the gas density in the centre dwindles. This is shown in the decline of the local SFR from around 6 Gyr ago in Ruiz-Lara et al. (2020), and around 2.5 Gyr ago in Fernández-Alvar et al. (2025), though with significant fluctuations in both cases. The outcome is the development of a radial gradient in the [$\alpha$/Fe] ratio, which grows from the inside–out as the enrichment from SNe Ia starts to dominate over the enrichment from SNe II. As we show in Section 3.3.4, this gradient can be imprinted on the low-$\alpha$ sequence as a downward 'slant' (a negative relation between [Fe/H] and [Mg/Fe], and see also Mason et al. 2024).

Several key elements are still missing from this interpretation. First, there is a lack of high-quality observational analogues of the MW at redshifts $z > 2$. Current surveys continue to face challenges in resolving galaxies at the necessary mass scales at such early epochs. However, gravitational lensing techniques (Jones et al. 2023) offer a promising avenue for overcoming these limitations by magnifying faint and distant systems, though further progress is needed to fully exploit their potential. It is also crucial to identify observational evidence of dense, star-forming gas clumps and the specific environments in which they emerge. Finally, improved age statistics for stellar populations in the MW is essential, though upcoming missions like *PLATO* (Rauer et al. 2025) may provide some relief on this front.

Our investigation is also relevant to M31, for which early spectroscopic results suggest a single chemical sequence similar to the MW high-$\alpha$ disc (Escala et al. 2020; Nidever et al. 2024). Inspection of fig. 4 in Nidever et al. (2024) reveals a narrow chemical track that spans from high-$\alpha$, low-metallicity to low-$\alpha$, high-metallicity stars. Compared to our analysis, this distribution closely resembles realizations that form an early thin gaseous disc (e.g. Au-2 and Au-24). This appears to conflict with observational evidence from dust geometry indicating that the dominant disc component of M31 has a thickness comparable to the MW thick disc (Dalcanton et al. 2023). However, these kinematics may be shaped by a possible major merger 2–4,Gyr ago (see arguments in Hammer et al. 2018; Bhattacharya et al. 2024; Tsakonas et al. 2025).

A recurring theme is that decrypting the earliest phases of our Galaxy's formation will increasingly rely on constraints from high-redshift observations. The coming decade of MW exploration will not be solely a Galactic Archaeology enterprise, but very likely an extragalactic effort.

## 5 CONCLUSIONS

We have performed an investigation into disc formation in the AURIGA simulations of MW-mass galaxies. Our primary focus is the emergence of chemical sequences in the [Fe/H] versus [Mg/Fe] chemical abundance plane, including bimodalities as is seen in the MW for disc stars around the Solar neighbourhood and in the Galactic centre. Whilst none of the simulations are an ideal match for the MW, we have isolated the formation mechanisms behind various properties which can be used to constrain or probe ideas about our Galaxy. Our key results are as follows:

(i) Half of the thirty AURIGA simulations exhibit bimodal distribution in their [Mg/Fe] distribution, with a further two having multimodal distributions. Regardless of the presence of a bimodality, high- and low-$\alpha$ populations usually adhere to the same typical bulk properties – those being hotter kinematics (with higher $\sigma_z$ by an average of 33 km s$^{-1}$), slower rotational velocities (by an average of 32 km s$^{-1}$), and older ages for the high-$\alpha$ stars (by an average of 4.6 Gyr).

(ii) There is no clear correlation between simulations with pronounced bimodal distributions and those with GSE-like debris in their stellar haloes. This shows that the formation of two chemical sequences is not dependent on a GSE-like merger.

(iii) Singular merger events, in particular those that are GSE-like, do not dilute the metallicity of the CGM by more than $\approx 0.3$ dex,

and usually by even less. The gas donated by GSE-like mergers forms only a minority of subsequent disc stars. The only exceptions are in the cases of very massive mergers (merger mass ratios of $> 1:2$, and total masses of $> 2 \times 10^{10}$), though these are ruled out for the MW. Therefore, the GSE gas was not responsible for the overlapping metallicities of the high- and low-$\alpha$ sequences, and we have reinforced this with empirical estimates. Other merger events such as *kraken* will have been just as unimportant, even if they were of higher mass ratio (see also arguments in Orkney et al. 2022).

(iv) Instead, the plurality of gas fuel for the formation of low-$\alpha$ sequences comes from a smooth accretion of the CGM. Sequences with low-[Fe/H] stars come from metal-poor gas that is freshly accreted onto the extremities of the gas disc. In contrast, gas discs that are thin form fewer low-[Fe/H] stars because the enriched and metal-poor gases mix together to form an enriched radial chemical gradient.

(v) The break between two chemical sequences is driven by a temporary quenching or reduction of the SFR. This can be due to multiple reasons, including gas exhaustion, post-merger quenching, or a violent disruption of the gas disc. As such, massive mergers such as the GSE can potentially contribute to the development of chemical bimodalities due to their merger-induced starbursts and the following quiescence. Other mechanisms may also be important, such as AGN-induced shutdowns or bar formation (e.g. Beane 2025; Beane et al. 2025), though these may themselves be triggered by a merger event. Disentangling these processes is beyond the scope of this work.

(vi) The distribution of chemical sequences in the [Fe/H] versus [Mg/Fe] property planes can be explained by trends in the SFR. Increasing SFRs lead to positive relations between [Fe/H] and [Mg/Fe], whereas decreasing or steady SFRs lead to the inverse.

(vii) As in prior studies, we find that radial migration is in effect but is not the underlying cause for the development of multiple chemical sequences.

From these findings, we can explain several features observed in the MW bimodality:

(i) The gap between the high- and low-$\alpha$ sequences is produced by a hiatus in the star formation, possibly due to the accretion of the GSE merger.

(ii) The slanted-shaped locus in the low-$\alpha$ sequence is due to the decaying SFR over the past $\sim 6$ Gyr.

(iii) The overlap in the metallicities of high- and low-$\alpha$ stars in the Solar neighbourhood arises due to a fraction of low-$\alpha$ stars forming in recently accreted metal-poor gases from the CGM.

(iv) The lack of a metal overlap in the inner Galaxy implies that the low-$\alpha$ sequence formed from a thin gas disc at those radii. This suggests the gas disc was thin but strongly flared during much of the low-$\alpha$ disc formation.

Our results highlight the importance of obtaining high-quality observational data for MW analogues at higher redshifts ($z \geq 2$). The *JWST* has recently begun probing such galaxies (e.g. Mowla et al. 2024; Tan et al. 2024a, b), and surveys like JADES (the JWST Advanced Deep Extragalactic Survey) may already be detecting MW analogues as early as $z \approx 8.2$ (Rusta et al. 2024).

Early spectroscopic studies of M31 suggest that its chemistry may not be bi-modal (Escala et al. 2020; Nidever et al. 2024). As observations of other Local Volume galaxies continue to improve, understanding the formation of chemical abundance patterns beyond the MW will become increasingly important. Our investigation revealed a broad diversity in the chemical evolution of MW-mass disc galaxies, which are highly sensitive to factors such as overall mass assembly, star formation history, and even the specifics of individual merger events. To better constrain the physical and dynamical processes that set chemical trends in MW-like galaxies, detailed extragalactic observations within the Local Volume, as well as the CGM of high-redshift galaxies, will prove essential in the era of Extremely Large Telescopes (see Hammer et al. 2021; Padovani & Cirasuolo 2023).

## ACKNOWLEDGEMENTS

We thank the anonymous referee for their thoughtful comments, which have greatly contributed to this manuscript. This work has used resources from the MareNostrum 4 supercomputer at the Barcelona Supercomputing Center (BSC). Much of our analysis was performed on the Virgo supercomputer at the Max Planck Computing and Data Facility (MPCDF). Further analysis and simulation work was performed on the NYX supercomputer at the Universität de Barcelona (ICCUB). CL and MO acknowledge funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 852839). CL also acknowledges funding from the Agence Nationale de la Recherche (ANR project ANR-24-CPJ1-0160-01). RJJG is supported by an STFC Ernest Rutherford Fellowship (ST/W003643/1).

## DATA AVAILABILITY

The AURIGA simulation data are publicly available to download via the Globus platform as described in section 4 of Grand et al. (2017). Any novel post-processed data are available upon request.

## REFERENCES

Abadi M. G., Navarro J. F., Steinmetz M., Eke V. R., 2003, ApJ, 591, 499
Agertz O. et al., 2021, MNRAS, 503, 5826
Amarante J. A. S., Beraldo e Silva L., Debattista V. P., Smith M. C., 2020, ApJ, 891, L30
Anders F. et al., 2014, A&A, 564, A115
Anglés-Alcázar D., Faucher-Giguère C.-A., Kereš D., Hopkins P. F., Quataert E., Murray N., 2017, MNRAS, 470, 4698
Anguiano B. et al., 2020, AJ, 160, 43
Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Beane A., 2025, ApJ, 982, 106
Beane A., Johnson J. W., Semenov V. A., Hernquist L., Chandra V., Conroy C., 2025, ApJ, 985, 221
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
Bensby T., Feltzing S., Lundström I., Ilyin I., 2005, A&A, 433, 185
Bensby T., Alves-Brito A., Oey M. S., Yong D., Meléndez J., 2011, ApJ, 735, L46
Beraldo e Silva L., Debattista V. P., Nidever D., Amarante J. A. S., Garver B., 2021, MNRAS, 502, 260
Bhattacharya S. et al., 2024, in Tabatabaei F., Barbuy B., Ting Y.-S., eds, IAU Symp. Vol. 377, Early Disk-Galaxy Formation from JWST to the Milky Way, Indian Statistical Institute (ISI), Kolkata, India, p. 123
Bhattacharyya A., 1946, Sankhyā: Indian J. Stat. (1933-1960), 7, 401
Bird J. C., Kazantzidis S., Weinberg D. H., Guedes J., Callegari S., Mayer L., Madau P., 2013, ApJ, 773, 43
Bonaca A. et al., 2020, ApJ, 897, L18
Borbolato L. et al., 2025, preprint (arXiv:2504.00135)
Bournaud F., Elmegreen B. G., Martig M., 2009, ApJ, 707, L1
Bovy J., Rix H.-W., Hogg D. W., 2012, ApJ, 751, 131

Bovy J., Rix H.-W., Schlafly E. F., Nidever D. L., Holtzman J. A., Shetrone M., Beers T. C., 2016, ApJ, 823, 30
Brook C. B. et al., 2012, MNRAS, 426, 690
Buck T., 2020, MNRAS, 491, 5435
Buck T., Obreja A., Ratcliffe B., Lu Y., Minchev I., Macciò A. V., 2023, MNRAS, 523, 1565
Bustamante S., Sparre M., Springel V., Grand R. J. J., 2018, MNRAS, 479, 3381
Carr C., Johnston K. V., Laporte C. F. P., Ness M. K., 2022, MNRAS, 516, 5067
Cerqui V., Haywood M., Snaith O., Di Matteo P., Casamiquela L., 2025, A&A, 699, A277
Chabrier G., 2003, PASP, 115, 763
Chandra V. et al., 2024, ApJ, 972, 112
Chen B., Hayden M. R., Sharma S., Bland-Hawthorn J., Kobayashi C., Karakas A. I., 2023, MNRAS, 523, 3791
Chiappini C., Matteucci F., Gratton R., 1997, ApJ, 477, 765
Chiappini C., Matteucci F., Romano D., 2001, ApJ, 554, 1044
Ciucă I. et al., 2024, MNRAS, 528, L122
Clarke A. J. et al., 2019, MNRAS, 484, 3476
Dalcanton J. J. et al., 2023, AJ, 166, 80
De Silva G. M. et al., 2015, MNRAS, 449, 2604
Deason A. J., Belokurov V., Sanders J. L., 2019, MNRAS, 490, 3426
Dekel A., Sari R., Ceverino D., 2009, ApJ, 703, 785
Deng L.-C. et al., 2012, Res. Astron. Astrophys., 12, 735
Deng M., Du C., Yang Y., Liao J., Ye D., 2024, ApJ, 975, 28
Di Matteo P., Combes F., Melchior A. L., Semelin B., 2007, A&A, 468, 61
Dillamore A. M., Belokurov V., Font A. S., McCarthy I. G., 2022, MNRAS, 513, 1867
Dodge B. C., Slone O., Lisanti M., Cohen T., 2023, MNRAS, 518, 2870
Dubay L. O., Johnson J. A., Johnson J. W., 2024, ApJ, 973, 55
Dubay L. O., Johnson J. A., Johnson J. W., Roberts J. D., 2025, preprint (arXiv:2508.00988)
Elmegreen B. G., Elmegreen D. M., 2005, ApJ, 627, 632
Elmegreen D. M., Elmegreen B. G., Ravindranath S., Coe D. A., 2007, ApJ, 658, 763
Escala I., Gilbert K. M., Kirby E. N., Wojno J., Cunningham E. C., Guhathakurta P., 2020, ApJ, 889, 177
Faria L., Patton D. R., Courteau S., Ellison S., Brown W., 2025, MNRAS, 537, 915
Fattahi A. et al., 2019, MNRAS, 484, 4471
Fernández-Alvar E. et al., 2025, preprint (arXiv:2503.19536)
Fragkoudi F. et al., 2020, MNRAS, 494, 5936
Frankel N., Rix H.-W., Ting Y.-S., Ness M., Hogg D. W., 2018, ApJ, 865, 96
Frankel N., Sanders J., Ting Y.-S., Rix H.-W., 2020, ApJ, 896, 15
Franx M., van Dokkum P. G., Förster Schreiber N. M., Wuyts S., Labbé I., Toft S., 2008, ApJ, 688, 770
Fuhrmann K., 1998, A&A, 338, 161
Fuhrmann K., 2011, MNRAS, 414, 2893
Gaia Collaboration, 2016, A&A, 595, A1
Gaia Collaboration, 2021, A&A, 649, A1
Genel S., Vogelsberger M., Nelson D., Sijacki D., Springel V., Hernquist L., 2013, MNRAS, 435, 1426
Genel S. et al., 2014, MNRAS, 445, 175
Gent M. R., Eitner P., Serenelli A., Friske J. K. S., Koposov S. E., Laporte C. F. P., Buck T., Bergemann M., 2024, A&A, 683, A74
Gilmore G. et al., 2012, The Messenger, 147, 25
Gómez F. A. et al., 2017, MNRAS, 472, 3722
Grand R. J. J., Kawata D., Cropper M., 2015, MNRAS, 447, 4018
Grand R. J. J. et al., 2017, MNRAS, 467, 179
Grand R. J. J. et al., 2018, MNRAS, 474, 3629
Grand R. J. J. et al., 2020, MNRAS, 497, 1603
Grand R. J. J., Fragkoudi F., Gómez F. A., Jenkins A., Marinacci F., Pakmor R., Springel V., 2024, MNRAS, 532, 1814
Grisoni V., Spitoni E., Matteucci F., Recio-Blanco A., de Laverny P., Hayden M., Mikolaitis Ŝ., Worley C. C., 2017, MNRAS, 472, 3637
Hammer F., Yang Y. B., Wang J. L., Ibata R., Flores H., Puech M., 2018, MNRAS, 475, 2754
Hammer F. et al., 2021, The Messenger, 182, 33
Hani M. H., Gosain H., Ellison S. L., Patton D. R., Torrey P., 2020, MNRAS, 493, 3716
Hayden M. R. et al., 2015, ApJ, 808, 132
Hayden M. R., Recio-Blanco A., de Laverny P., Mikolaitis S., Worley C. C., 2017, A&A, 608, L1
Haywood M., Di Matteo P., Lehnert M. D., Katz D., Gómez A., 2013, A&A, 560, A109
Haywood M., Lehnert M. D., Di Matteo P., Snaith O., Schultheis M., Katz D., Gómez A., 2016, A&A, 589, A66
Haywood M., Di Matteo P., Lehnert M., Snaith O., Fragkoudi F., Khoperskov S., 2018, A&A, 618, A78
Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, Nature, 563, 85
House E. L. et al., 2011, MNRAS, 415, 2652
Imig J. et al., 2023, ApJ, 954, 124
Jenkins A., 2013, MNRAS, 434, 2094
Johnson J. W. et al., 2021, MNRAS, 508, 4484
Jones T. et al., 2023, ApJ, 951, L17
Jönsson H. et al., 2018, AJ, 156, 126
Jurić M. et al., 2008, ApJ, 673, 864
Khoperskov S., Haywood M., Snaith O., Di Matteo P., Lehnert M., Vasiliev E., Naroenkov S., Berczik P., 2021, MNRAS, 501, 5176
Khoperskov S. et al., 2025, A&A, 700, A89
Kirby E. N., Cohen J. G., Guhathakurta P., Cheng L., Bullock J. S., Gallazzi A., 2013, ApJ, 779, 102
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
Lane J. M. M., Bovy J., Mackereth J. T., 2023, MNRAS, 526, 1209
Li F., Rahman M., Murray N., Kereš D., Wetzel A., Faucher-Giguère C.-A., Hopkins P. F., Moreno J., 2025, ApJ, 979, 7
Lian J., Luo L., 2024, ApJ, 960, L10
Lian J. et al., 2020, MNRAS, 497, 3557
Limberg G., Souza S. O., Pérez-Villegas A., Rossi S., Perottoni H. D., Santucci R. M., 2022, ApJ, 935, 109
Lindegren L. et al., 2018, A&A, 616, A2
Linden S. T. et al., 2017, ApJ, 842, 49
Livermore R. C. et al., 2012, MNRAS, 427, 688
Loebman S. R., Roškar R., Debattista V. P., Ivezić Ž., Quinn T. R., Wadsley J., 2011, ApJ, 737, 8
Loebman S. R., Debattista V. P., Nidever D. L., Hayden M. R., Holtzman J. A., Clarke A. J., Roškar R., Valluri M., 2016, ApJ, 818, L6
Luo W., Yang X., Zhang Y., 2014, ApJ, 789, L16
Mackereth J. T. et al., 2017, MNRAS, 471, 3057
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
Majewski S. R. et al., 2017, AJ, 154, 94
Marinacci F., Pakmor R., Springel V., 2014, MNRAS, 437, 1750
Mason A. C., Crain R. A., Schiavon R. P., Weinberg D. H., Pfeffer J., Schaye J., Schaller M., Theuns T., 2024, MNRAS, 533, 184
Massari D., Koppelman H. H., Helmi A., 2019, A&A, 630, L4
McWilliam A., 1997, ARA&A, 35, 503
Merrow A., Grand R. J. J., Fragkoudi F., Martig M., 2024, MNRAS, 531, 1520
Michel E. et al., 2021, Exp. Astron., 51, 945
Miglio A. et al., 2021, A&A, 645, A85
Minchev I., Famaey B., 2010, ApJ, 722, 112
Minchev I., Famaey B., Combes F., Di Matteo P., Mouhcine M., Wozniak H., 2011, A&A, 527, A147
Minchev I., Chiappini C., Martig M., 2013, A&A, 558, A9
Minchev I., Chiappini C., Martig M., 2014a, A&A, 572, A92
Minchev I. et al., 2014b, ApJ, 781, L20
Monachesi A. et al., 2019, MNRAS, 485, 2589
Moreno J. et al., 2021, MNRAS, 503, 3113

Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121
Mould J., 2005, AJ, 129, 698
Mowla L. et al., 2024, Nature, 636, 332
Naidu R. P. et al., 2021, ApJ, 923, 92
Nelson D. et al., 2019, Comput. Astrophys. Cosmol., 6, 2
Nidever D. L. et al., 2014, ApJ, 796, 38
Nidever D. L. et al., 2024, in Tabatabaei F., Barbuy B., Ting Y.-S., eds, IAU Symp. Vol. 377, Early Disk-Galaxy Formation from JWST to the Milky Way. Cambridge Univ. Press, Cambridge, p. 115
Noguchi M., 2018, Nature, 559, 585
Orkney M. D. A., Laporte C. F. P., 2025, preprint (arXiv:2509.09576)
Orkney M. D. A. et al., 2022, MNRAS, 517, L138
Orkney M. D. A. et al., 2023, MNRAS, 525, 683
Padovani P., Cirasuolo M., 2023, Contemp. Phys., 64, 47
Palla M., Matteucci F., Spitoni E., Vincenzo F., Grisoni V., 2020, MNRAS, 498, 1710
Papovich C. et al., 2015, ApJ, 803, 26
Parul H., Bailin J., Loebman S. R., Wetzel A., Barry M., Bhattarai B., 2025, MNRAS, 537, 1571
Patel S. G. et al., 2013a, ApJ, 766, 15
Patel S. G. et al., 2013b, ApJ, 778, 115
Patil A., Bovy J., 2024, in American Astronomical Society Meeting Abstracts, p. 129.03
Planck Collaboration XVI, 2014, A&A, 571, A16
Popping G., Caputi K. I., Somerville R. S., Trager S. C., 2012, MNRAS, 425, 2386
Prantzos N. et al., 2023, MNRAS, 523, 2126
Queiroz A. B. A. et al., 2020, A&A, 638, A76
Queiroz A. B. A. et al., 2023, A&A, 673, A155
Rauer H. et al., 2025, Exp. Astron., 59, 26
Renaud F., Agertz O., Andersson E. P., Read J. I., Ryde N., Bensby T., Rey M. P., Feuillet D. K., 2021, MNRAS, 503, 5868
Richard S., Brook C. B., Martel H., Kawata D., Gibson B. K., Sanchez-Blazquez P., 2010, MNRAS, 402, 1489
Robin A. C., Haywood M., Creze M., Ojha D. K., Bienayme O., 1996, A&A, 305, 125
Rockosi C. M. et al., 2022, ApJS, 259, 60
Rodríguez Montero F., Davé R., Wild V., Anglés-Alcázar D., Narayanan D., 2019, MNRAS, 490, 2139
Roškar R., Debattista V. P., Quinn T. R., Stinson G. S., Wadsley J., 2008, ApJ, 684, L79
Ruchti G. R. et al., 2011, ApJ, 737, 9
Ruiz-Lara T., Gallart C., Bernard E. J., Cassisi S., 2020, Nat. Astron., 4, 965
Rusta E., Salvadori S., Gelli V., Koutsouridou I., Marconi A., 2024, ApJ, 974, L35
Sahlholdt C. L., Feltzing S., Feuillet D. K., 2022, MNRAS, 510, 4669
Sanderson R. E. et al., 2020, ApJS, 246, 6
Santiago B. X. et al., 2016, A&A, 585, A42
Schönrich R., Binney J., 2009a, MNRAS, 396, 203
Schönrich R., Binney J., 2009b, MNRAS, 399, 1145
Scott N., van de Sande J., Sharma S., Bland-Hawthorn J., Freeman K., Gerhard O., Hayden M. R., McDermid R., 2021, ApJ, 913, L11
Segovia Otero Á., Renaud F., Agertz O., 2022, MNRAS, 516, 2272
Sellwood J. A., Binney J. J., 2002, MNRAS, 336, 785
Sharma S., Hayden M. R., Bland-Hawthorn J., 2021, MNRAS, 507, 5882
Snaith O., Haywood M., Di Matteo P., Lehnert M. D., Combes F., Katz D., Gómez A., 2015, A&A, 578, A87
Snaith O. N., Bailin J., Gibson B. K., Bell E. F., Stinson G., Valluri M., Wadsley J., Couchman H., 2016, MNRAS, 456, 3119
Soderblom D. R., 2010, ARA&A, 48, 581
Spitoni E., Silva Aguirre V., Matteucci F., Calura F., Grisoni V., 2019, A&A, 623, A60
Spitoni E. et al., 2021, A&A, 647, A73
Spitoni E. et al., 2023, A&A, 670, A109
Spitoni E., Matteucci F., Gratton R., Ratcliffe B., Minchev I., Cescutti G., 2024, A&A, 690, A208
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Springel V. et al., 2005, Nature, 435, 629
Steinmetz M. et al., 2006, AJ, 132, 1645
Tacconi L. J. et al., 2013, ApJ, 768, 74
Tan V. Y. Y. et al., 2024a, preprint (arXiv:2412.07829)
Tan V. Y. Y., Muzzin A., Marchesini D., Sok V., Sarrouh G. T. E., Marsan Z. C., 2024b, ApJ, 964, 177
Tinsley B. M., 1979, ApJ, 229, 1046
Tinsley B. M., 1980, Fund. Cosmic Phys., 5, 287
Tsakonas C., Arnaboldi M., Bhattacharya S., Hammer F., Yang Y., Gerhard O., Wyse R. F. G., Hatzidimitriou D., 2025, A&A, 699, A56
Tsukui T., Wisnioski E., Bland-Hawthorn J., Freeman K., 2025, MNRAS, 540, 3493
van der Wel A. et al., 2014, ApJ, 788, 28
van Dokkum P. G. et al., 2013, ApJ, 771, L35
Vincenzo F., Spitoni E., Calura F., Matteucci F., Silva Aguirre V., Miglio A., Cescutti G., 2019, MNRAS, 487, L47
Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031
Willott C. J. et al., 2022, PASP, 134, 025002
Wuyts S. et al., 2012, ApJ, 753, 114
Xiang M., Rix H.-W., 2022, Nature, 603, 599
Yanny B. et al., 2009, AJ, 137, 4377
Yoachim P., Dalcanton J. J., 2006, AJ, 131, 226

## SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

**suppl_data**

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

## APPENDIX A: FULL STAR FORMATION HISTORY PLOT

We include star formation histories for the full AURIGA suite in Fig. A1.

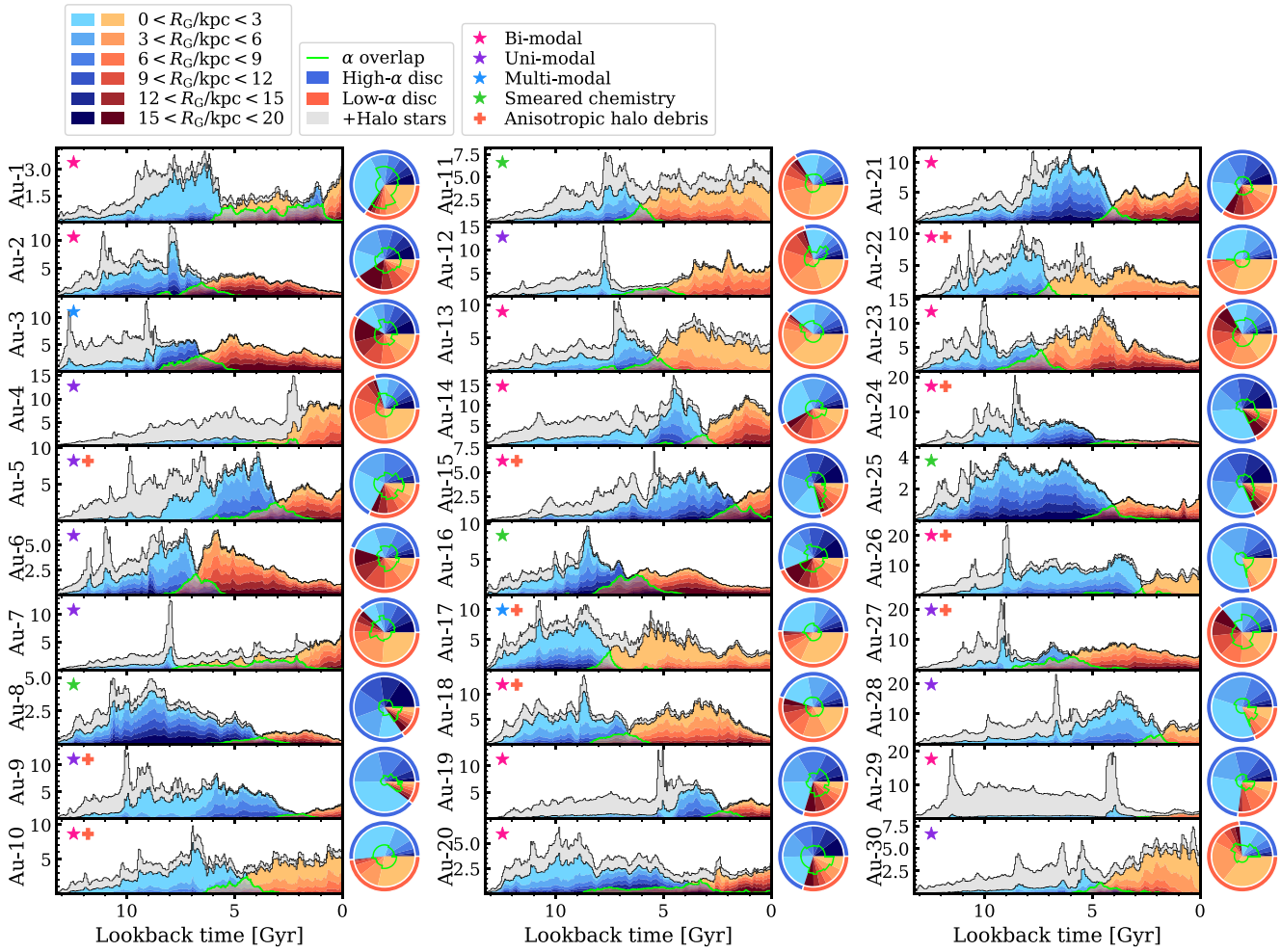


**Figure A1.** The same as in Fig. 4, but now for the entire AURIGA suite.

## APPENDIX B: ESTIMATING THE GAS DILUTION IN THE MW DUE TO THE GSE

We extend our analysis by estimating the metallicity dilution resulting from the GSE infall event. Our model parameters are described in Table B1, including the stellar, gas and halo mass, as well as the gas metallicity for both the proto-MW and GSE. We now describe our methods for deriving these model parameters.

We assume that the GSE infall occurred 10 Gyr ago, which corresponds to a redshift of $z \approx 2$. The stellar mass of the proto-MW at this time can be estimated following comparisons with MW analogue galaxies in observational surveys (i.e. Dokkum et al. 2013; Patel et al. 2013b; Papovich et al. 2015; Tan et al. 2024a). We opt to employ the relation with $1\sigma$ scatter shown in fig. 2 of Tan et al. (2024a), based upon a selection of MW analogues out to $z = 5$ from the CANUCS survey (The Canadian NIRISS Unbiased Cluster Survey, Willott et al. 2022). The gas mass is determined assuming a gas fraction of 0.4–0.6 (loosely based on measurements of star-forming massive galaxies at $z = 2$ Tacconi et al. 2013), where $f_{\rm gas} = M_{\rm gas}/(M_{\rm gas} + M_\star)$. The corresponding halo mass is found following the redshift-dependent stellar mass–halo mass relation of Moster et al. (2013, hereafter Moster13 SMHM).

We estimate the gas metallicity by considering that the stars born at the metal-rich end of the MW high-$\alpha$ sequence will reflect the ambient gas metallicity before the GSE infall. From visual inspection of the figures in Queiroz et al. (2020), we take [Fe/H] $= 0.00 \pm 0.05$ as a rough guideline. Then, the amount of dilution required to reach the metal-poor end of the low-$\alpha$ sequence would be around $-0.5$ dex. This is far less than what is presented in the chemical evolution models of Spitoni et al. (2019, 2021, 2023), but we consider it as a conservative minimum required dilution.

An alternative estimate of the metallicity can be obtained using the stellar mass–metallicity relation from Kirby et al. (2013, given by their equation 4, hereafter Kirby13 MZR), which gives [Fe/H](proto-MW) $= -0.58^{+0.27}_{-0.16}$. Although this relation is calibrated on galaxy observations at $z = 0$, it none the less captures the spread and median of observed stellar metallicities within the high-$\alpha$ sequence. However, it is not suitable for our purposes, as we are interested in the chemical evolution stage near the end of the high-$\alpha$ sequence where metallicities approach [Fe/H] $\approx 0$.

For the properties of the GSE progenitor galaxy, we construct three models based upon literature values. We assume the gas metallicity matches the stellar metallicity. Average stellar enrichment lags behind the gas enrichment, and so this choice should be

**Table B1.** The model properties for the proto-MW and GSE merger at a time of 10 Gyr ago or $z \approx 2$.

| Model | $M_\star$ [$M_\odot$] | $M_{\rm gas}$ [$M_\odot$] | $M_{\rm DM}$ [$M_\odot$] | Gas [Fe/H] |
|---|---|---|---|---|
| MW | $4.92^{+7.97}_{-3.37} \times 10^9$ | $4.92^{+8.28}_{-3.36} \times 10^9$ | $4.92^{+4.97}_{-2.96} \times 10^{11}$ | $0.00 \pm 0.05$ |
| GSE 1 | $3.69^{+31.5}_{-0.33} \times 10^7$ | $1.47^{+13.3}_{-1.30} \times 10^8$ | $3.37^{+8.57}_{-2.89} \times 10^{10}$ | $-1.22 \pm 0.23$[a] |
| GSE 2 | $1.45^{+0.92}_{-0.51} \times 10^8$[b] | $5.80^{+5.59}_{-2.38} \times 10^8$ | $7.07^{+6.00}_{-3.59} \times 10^{10}$ | $-1.04^{+0.20}_{-0.18}$ |
| GSE 3 | $6 \times 10^8$[c] | $2.40^{+1.56}_{-0.76} \times 10^9$ | $1.53^{+1.08}_{-0.65} \times 10^{11}$ | $-1.6$[c] |

*Notes*. Sources are as follows: (a) Limberg et al. (2022), (b) Lane et al. (2023), and (c) Helmi et al. (2018). The methods used to derive these model parameters are described in the main text of Appendix B.

considered as a low estimate of the gas metallicity. We calculate the gas mass assuming high gas fractions of 0.7–0.9, reflecting the high gas fractions expected for ancient dwarf galaxies (i.e. Popping et al. 2012). Once again, the halo masses are approximated from the Moster13 SMHM. Our models are described as follows:

(i) *GSE 1*. We take the GSE stellar metallicity value of [Fe/H] $= -1.22 \pm 0.23$ from Limberg et al. (2022) and find the corresponding stellar mass from the Kirby13 MZR. We note that Limberg et al. (2022) also quote their own stellar mass estimate of $\sim 1.3 \times 10^9$ $M_\odot$, but we determine that this is likely too high because the entire MW stellar halo is only $\sim 1.4 \times 10^9$ $M_\odot$ (Deason, Belokurov & Sanders 2019).

(ii) *GSE 2*. We take the GSE stellar mass value of $M_\star = 1.45^{+0.92}_{-0.51} \times 10^8$ $M_\odot$ from Lane et al. (2023) and find the corresponding stellar metallicity from the inverted Kirby13 MZR.

(iii) *GSE 3*. Here, we take the direct GSE stellar mass and metallicity values reported in Helmi et al. (2018) as $M_\star = 6 \times 10^8$ $M_\odot$ and [Fe/H] $= -1.6$, and note that these values are significantly offset from the Kirby13 MZR. This model can be considered as a particularly high-mass and metal-poor estimate for the GSE, thereby maximizing any resulting dilution.

The uncertainties on the parameters in Table B1 are either taken from their respective sources, or are approximated as the $\pm\sigma$ percentiles from a Monte Carlo method using $10^6$ samples. Symmetric uncertainties are sampled from a normal distribution, whereas the upper and lower bounds for asymmetric uncertainties are calculated independently with a two-piece normal distribution provided by the TWOPIECE PYTHON package. The distributions in $f_{\rm gas}$ are sampled from a uniform distribution. The Kirby13 MZR and Moster13 SMHM have a further 0.17 and 0.15 dex scatter, respectively, which we include in addition to the parametric uncertainties of each relation. The GSE parameters from Helmi et al. (2018) do not include uncertainties, and so we caution the reader that the true uncertainties will be greater than those shown here.

We calculate the gas dilution by averaging the metallicities of the GSE model and the proto-MW, weighting by their respective gas masses. It is conceivable that the GSE gas did not mix homogeneously with the proto-MW gas, and so we choose to calculate the dilution over a range of proto-MW gas masses in Fig. B1. There, the vertical black line indicates our estimate for the true gas mass of the proto-MW from Table B1, with associated uncertainty represented as a shaded band. This figure shows that the most likely outcome is for a dilution magnitude of $< 0.2$, very similar to the maximum dilution magnitudes that we estimated for AURIGA in panels (b) and (c) of Fig. 7, but far from the required 0.5 dex needed in the MW. The dilution magnitude remains relatively low even in cases where the GSE gas has a greater mass than the proto-MW gas, and this is due to the logarithmic nature of the [Fe/H] ratio.

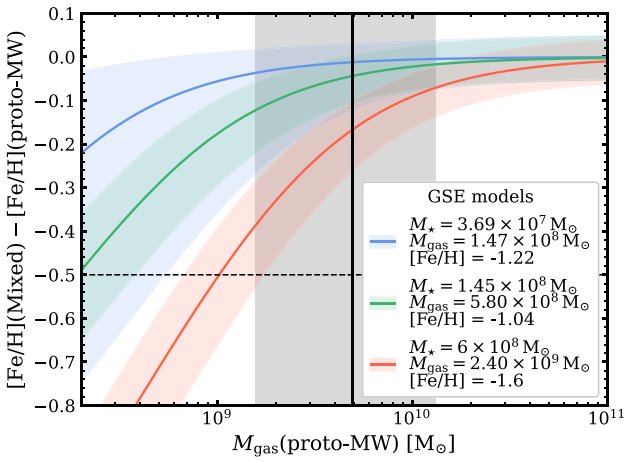


**Figure B1.** The metallicity dilution resulting from three different GSE merger models (see the legend and further description in the main text) mixing with a range of gas masses at a metallicity of $0.00 \pm 0.05$ (approximately the metal-rich end of the MW high-$\alpha$ sequence). The shaded bands represent the $1\sigma$ uncertainty in each case. The vertical solid line is an estimate of the gas mass in the proto-MW at the time of the GSE merger (10 Gyr ago) where once again the shaded band represents the $1\sigma$ uncertainty. A dashed horizontal line marks a dilution of 0.5 dex, which we consider the minimum dilution required to meet the requirements of the two-infall model.

When the uncertainties are accounted for, the dilution magnitude can exceed 0.5 dex for the GSE 3 model. However, this is assuming a very high GSE gas mass and a very low proto-MW gas mass simultaneously, which would be disfavoured by current estimates of the total merger mass ratio (of $\leq 1:4$). Also, a higher GSE stellar mass with a fixed halo mass would favour a lower gas fraction. These two effects are not considered by our uncertainty estimates. For the other two models, greater dilution magnitudes can only be achieved assuming that the GSE gas mixes with an infinitesimal volume of proto-MW gas, which is not realistic.

The GSE may have deposited its gas around the edge of the proto-MW gas disc, where the gas densities are lower and higher dilutions can be achieved more easily. An example of this scenario is described in AURIGA by Grand et al. (2018), for a late-time minor merger in the higher resolution version of Au-23. This merger has a wide pericentre passage of $> 40$ kpc, which enables its gas to be gradually stripped away and accreted onto the edge of the gas disc. This is a less likely possibility for the GSE, which must have had a highly radial infall trajectory with a low pericentre.

Our results are consistent with the study by Palla et al. (2020). They model the gas contribution from the GSE, calibrated to

match the abundances reported in Helmi et al. (2018). The results indicate that the merger supplied approximately 15 per cent of the total gas during the second-infall epoch, remarkably close to the fractions shown in our Fig. 8 in the case of a 1 : 8 mass ratio merger. They also find a minimal metal dilution, no matter the assumed accretion time or the chemical composition of the merger gas.

This paper has been typeset from a TEX/LATEX file prepared by the author.