CrossMark

Tracing Birth Properties of Stars with Abundance Clustering

Bridget L. Ratcliffe[1] , Melissa K. Ness[2,3] , Tobias Buck[4] , Kathryn V. Johnston[2,3] , Bodhisattva Sen[1] ,
Leandro Beraldo e Silva[5] , and Victor P. Debattista[5,6]

[1] Department of Statistics, Columbia University, 1255 Amsterdam Avenue, New York, NY 10027, USA; blr2147@columbia.edu
[2] Department of Astronomy, Columbia University, 550 West 120th Street, New York, NY 10027, USA
[3] Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA
[4] Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany
[5] Jeremiah Horrocks Institute, University of Central Lancashire, Preston, PR1 2HE, UK
[6] Institute of Space Sciences & Astronomy, University of Malta, Msida, MSD 2080, Malta

Abstract

To understand the formation and evolution of the Milky Way disk, we must connect its current properties to its past. We explore hydrodynamical cosmological simulations to investigate how the chemical abundances of stars might be linked to their origins. Using hierarchical clustering of abundance measurements in two Milky Way–like simulations with distributed and steady star formation histories, we find that groups of chemically similar stars comprise different groups in birth place (R_{birth}) and time (age). Simulating observational abundance errors (0.05 dex), we find that to trace distinct groups of (R_{birth}, age) requires a large vector of abundances. Using 15 element abundances (Fe, O, Mg, S, Si, C, P, Mn, Ne, Al, N, V, Ba, Cr, Co), up to ≈10 groups can be defined with ≈25% overlap in (R_{birth}, age). We build a simple model to show that in the context of these simulations, it is possible to infer a star's age and R_{birth} from abundances with precisions of ±0.06 Gyr and ±1.17 kpc, respectively. We find that abundance clustering is ineffective for a third simulation, where low-α stars form distributed in the disk and early high-α stars form more rapidly in clumps that sink toward the Galactic center as their constituent stars evolve to enrich the interstellar medium. However, this formation path leads to large age dispersions across the [α/Fe]–[Fe/H] plane, which is inconsistent with the Milky Way's observed properties. We conclude that abundance clustering is a promising approach toward charting the history of our Galaxy.

Unified Astronomy Thesaurus concepts: Stellar populations (1622); Clustering (1908); Stellar abundances (1577); Astrostatistics techniques (1886); Chemical abundances (224); Milky Way Galaxy (1054); Hydrodynamical simulations (767)

1. Introduction

With large spectroscopic surveys, we have access to precise individual chemical element abundance measurements for 10^5–10^6 stars. The GALactic Archaeology with HERMES survey (GALAH; Buder et al. 2018) provides stellar parameters and up to 23 abundances for 342,682 stars, and the Gaia–European Southern Observatory (ESO) survey (Gilmore et al. 2012) measures detailed abundances for 12 elements in about 10,000 field stars. Another example of the current depth of observational data is the 16th data release of the Apache Point Observatory Galactic Evolution Experiment (APOGEE) survey, which contains information for 437,485 unique stars and more than 20 abundances (Ahumada et al. 2020; Jönsson et al. 2020).

For the Milky Way, we can use large spectroscopic surveys to catalog an ensemble of measurements. These include precise stellar metallicities and abundances ([Fe/H], [X/Fe]) and imprecise ages, as well as current-day positions and orbital parameters. These numbers can be used to work toward the reconstruction of the initial state of the Milky Way. While the chemical abundances of stars are birth properties, stars move and their orbits evolve over time (e.g., Sellwood & Binney 2002; Roškar et al. 2008; Schönrich & Binney 2009a; Minchev & Famaey 2010; Hayden et al. 2018) and their dynamical

properties change (e.g., Roškar et al. 2012). Chemical tagging utilizes the unchanging chemical abundances to identify star formation groups (Freeman & Bland-Hawthorn 2002). This is in theory possible as birth clusters up to $10^5 M_\odot$ are anticipated to be chemically homogeneous (Bland-Hawthorn et al. 2010). Chemical tagging has great promise (e.g., Hogg et al. 2016; Martell et al. 2016); however, it has been shown to be difficult due to the need for extremely large sample sizes (Ting et al. 2015) and high-precision data (Lindegren & Feltzing 2013).

In paper I of this series (Ratcliffe et al. 2020), we examined the distribution of clusters defined in a 19-dimensional chemical abundance space for 27,000 red clump APOGEE Data Release 14 (DR14) stars in the Milky Way's disk (Bovy et al. 2014). Using a nonparametric agglomerative hierarchical clustering method, we determined that the groups defined in abundance space are spatially separated as a function of age.

Yet, to reconstruct the disk in the past, we need to know where the stars were born, which we cannot measure directly from data. Recently, we have gained access to some of the largest and highest-resolution samples of zoom-in Milky Way–like simulations, e.g., NIHAO-UHD (Buck et al. 2020), AURGIA (Grand et al. 2017), and Feedback In Realistic Environments-2 (FIRE-2; Garrison-Kimmel et al. 2018). We also now have access to more abundance information in simulations, with g8.26e11 from the Numerical Investigation of a Hundred Astronomical Objects (NIHAO)-UHD suite providing information for 15 abundances and Ananke from Sanderson et al. (2020) having information for 11 abundances. Thus, now in simulations we have access to the full set of properties to

trace the formation and evolution of disks (e.g., [Fe/H], [X/Fe], age, and their origin as indicated by their birth radii within the Galactic disk, R_{birth}). This enables us to use simulations to investigate and understand the dependencies and relationships between these properties in disk galaxies, under particular initial conditions and evolutionary events.

Some recent work has examined R_{birth} in simulations in an attempt to better understand the Milky Way's formation (e.g., Loebman et al. 2011; Minchev et al. 2012a, 2012b). With the use of a series of high-resolution smooth-particle hydrodynamics simulations of isolated galaxy formations, Roškar et al. (2008) showed that radial migration is possible on short timescales. In agreement with Grand et al. (2016), Bird et al. (2021) argued for inside-out and upside-down formation, in addition to showing a correlation between R_{birth} and birth kinematics. Johnson et al. (2021) find that in their hybrid hydrodynamical simulation there is a relationship between age, abundances, and R_{birth} in the solar neighborhood (consistent with the expectations of earlier works: Matteucci & Francois 1989; Friedli et al. 1994; Schönrich & Binney 2009b; Minchev et al. 2018; Hemler et al. 2021), and that the low-α sequence represents a superposition of populations achieved by radial migration rather than an evolutionary sequence (see also Buck 2020).

In this paper, we use simulations to explore the physical meaning of groups of stars defined only in ([Fe/H], [X/Fe]) space in the observational data. We do not wish to trace back individual birth groups such as done in chemical tagging proposed by Freeman & Bland-Hawthorn (2002), but rather focus on the general birth properties of chemically similar stars in n abundance dimensions. That is, to test if abundance clustering allows the extraction of information about stellar birth properties. The questions we wish to answer are (i) do abundances link to birth associations, and if so does it rely on star formation processes; (ii) how does the presence of observational errors and sample size effect results; (iii) are results dependent on the clustering methods used; and (iv) is there a relationship between stellar birth properties and their abundances? Milky Way analog simulations are a good tool to investigate the questions we pose and qualitatively represent the formation processes of our Galaxy. Both observations and hydrodynamical simulations of Milky Way analogues show that from about $z = 1$, stellar disks form inside-out, with ongoing enrichment and star formation across the disk until late times (e.g., Chiappini et al. 1997; Muñoz-Mateos et al. 2007; Roškar et al. 2008; Bovy et al. 2012; Stinson et al. 2013; Minchev et al. 2015; Sanderson et al. 2020) (see Haywood et al. 2013, for argument against inside-out formation).

The numerical relationship between R_{birth}, abundances, and age in the Milky Way has been investigated before. Minchev et al. (2018) proposed a largely model-independent method to infer stellar birth radii for observational data by using age and [Fe/H] to project stars to their birth radius, along gradients of age–[Fe/H] . Further, Frankel et al. (2018) constructed a full physical model of the age–metallicity distribution given stellar radii, and inverted this relation to find birth radii for the low-α sequence of APOGEE data. Ness et al. (2019) also suggest that [Fe/H], age, and high- or low-α sequence membership is all that is needed to infer a star's R_{birth}. Using abundances, age, and the metallicity profile of the interstellar medium at the time of the star's formation, Feltzing et al. (2020) tested the effects of radial migration on red giant branch stars by quantifying the

fraction of stars that have been subject to blurring and churning. While it is possible to infer R_{birth} given some modeling assumptions, one can never directly measure the birth radius of an individual star in the Milky Way. Using simulations, we can study the link between chemical composition at birth ([Fe/H], [X/Fe]), birth time (t_{birth}, or age) and birth location (R_{birth}).

This paper is organized as follows. In Section 2 we discuss the two simulations upon which this paper focuses. Clustering methods used in this work are described in Section 3. Section 4 explores how chemically similar stars separate into distinct groups in the age–R_{birth} plane, which means they occupy different spaces in birth time and place. Section 5 explores how these results change under sampling and observational errors. Our last results section, Section 6, quantifies the relationship between ([Fe/H], [X/Fe]) and age, and ([Fe/H], [X/Fe], age) and R_{birth} using simple second-order polynomial regressions. Finally, Sections 7 and 8 present the discussion and conclusions of this analysis.
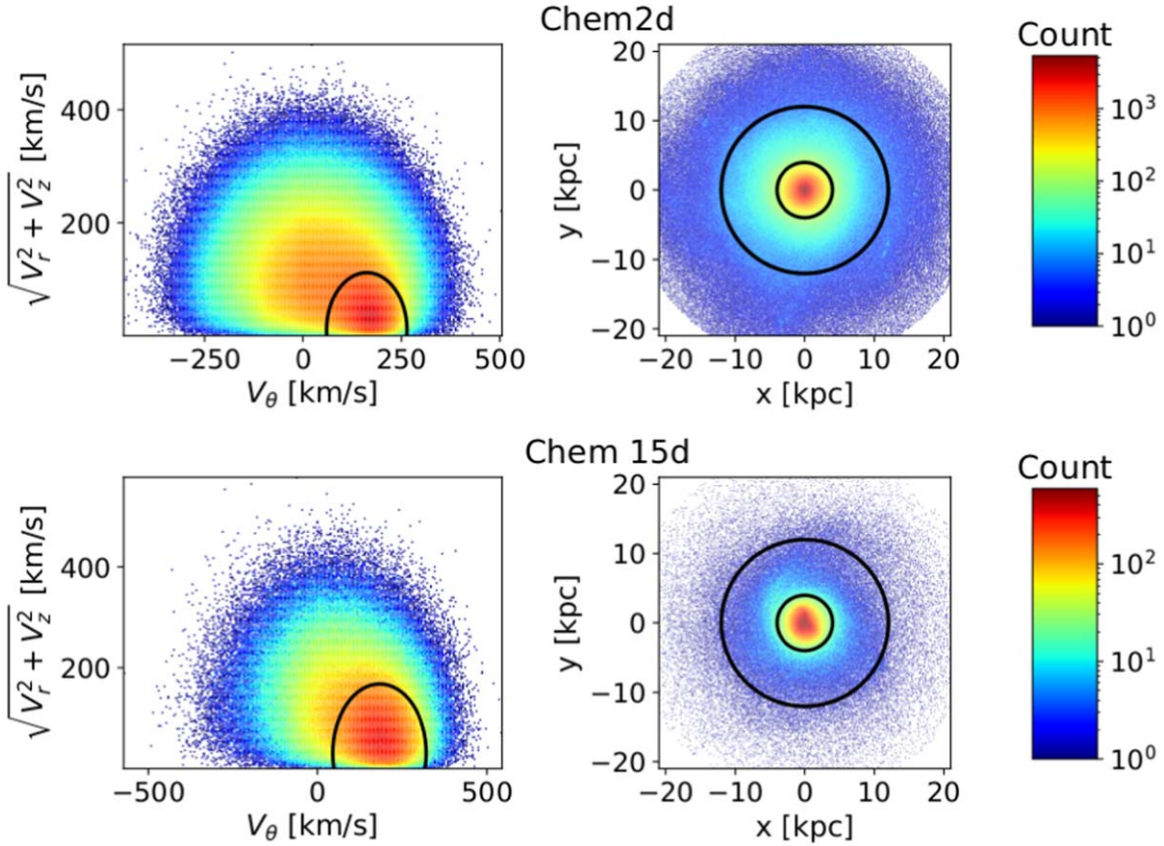
2. Simulations

In this paper, we focus on two simulations, one with only two chemical abundances available (g7.55e11, referred to as Chem2d) and another with 15 chemical abundances (g8.26e11, referred to as Chem15d). We include both the high- and low-dimensional simulations to investigate how stellar birth information is able to be captured in a few versus large vector of abundances.

The Chem2d and Chem15d simulations are taken from the NIHAO simulation suite (Wang et al. 2015), and are part of the NIHAO-UHD suite (Buck et al. 2020, 2019a). The simulations were performed with the smooth-particle hydrodynamics solver GASOLINE2 (Wadsley et al. 2017). They are both spiral disk galaxies with bulges, and were bulge dominated until redshift $z \geqslant 1$, with prominent stellar disks forming about 7–8 Gyr ago (Buck et al. 2020). Chem2d has a stellar particle mass of $0.093 \times 10^5 \, M_\odot$, while Chem15d has a stellar particle mass of $1.06 \times 10^5 \, M_\odot$. For more details, see Buck et al. (2020) for g7.55e11 and Buck et al. (2021) for g8.26e11.[7]

There are two main differences between the simulations: their resolution and chemical enrichment prescription. As seen from the particle masses, Chem2d has a higher mass resolution while Chem15d is at fiducial resolution. However, NIHAO galaxies are numerically well converged, so for the purpose of this work the resolution difference should not matter. The other difference is a detail in the numerics. Chem15d has an updated chemical enrichment prescription as described in Buck et al. (2021), which allows us to trace the 15 elements investigated here. This does not affect the global properties of the galaxy such as stellar mass, star formation history, or disk size much.

One other difference is that the two galaxies have slightly different formation histories as they are two different realizations of Lambda cold dark matter (ΛCDM) initial conditions. This mainly affects the accretion history and the final stellar mass or disk size. However, what is important for this work is that the stellar disk properties of these simulations, such as stellar mass, size, and rotation, agree with observations of the Milky Way and local galaxies (Buck et al. 2020). Furthermore,

---

Figure 1. Selection cuts of the (top) Chem2d, and (bottom) Chem15d simulations in the (left) velocity and (right) spatial planes. The black circle in the Toomre diagram marks where we defined the separation for the kinematically different disk and halo stars. Spatially, we define disk stars to have $|z| \leqslant 0.5$ kpc and $4 \leqslant R_{\mathrm{GAL}} \leqslant 12$ kpc, shown in the right of the figure.

the age and R_{birth} distribution in the $[\alpha/\mathrm{Fe}]$–$[\mathrm{Fe}/\mathrm{H}]$ abundance plane is very similar to that observed in the Milky Way (e.g., Lu et al. 2021; Minchev et al. 2018).

Chem2d has abundance information for $[\mathrm{Fe}/\mathrm{H}]$ and $[\mathrm{O}/\mathrm{Fe}]$ while Chem15d has the abundances of 15 elements from five different families, five iron peak elements (Fe, V, Cr, Mn, Co), two light elements (C, N), two light odd-Z elements (Al, P), five α-elements (O, Mg, S, Si, Ne), and one s-process element (Ba). Both simulations show a bimodality in the $[\alpha/\mathrm{Fe}]$–$[\mathrm{Fe}/\mathrm{H}]$ plane. The high- and low-α sequences in both simulations are a consequence of a gas-rich merger; the high-α sequence evolves first in the early galaxy, while the low-α sequence forms after the gas-rich merger dilutes the interstellar medium's metallicity (Buck 2020).

2.1. Selection Cuts

We focus our analysis on the present-day disk. We first select stars that overlap in space with the disk by imposing limits of $|z| \leqslant 0.5$ kpc and 4 kpc $\leqslant R_{\mathrm{GAL}} \leqslant 12$ kpc, though our results are consistent for other spatial cuts. We then determine stars that are current disk members using three-dimensional velocity space. We model a two-component Gaussian mixture model in (v_θ, v_r, v_z), similar to the approach taken by Buck et al. (2019b) and Obreja et al. (2018a) to model simulations in kinematic space using Galactic Structure Finder (GSF; Obreja et al. 2018b), and define disk stars to have a Mahalanobis distance less than 2 from the center of the corresponding Gaussian distribution. Figure 1 shows these selection cuts in the equivalent Toomre diagram and $x - y$ plane for both

simulations. Finally, we do an additional cut of $R_{\mathrm{birth}} \leqslant 15$ kpc to ensure we are not looking at infalling debris.

Additionally, since the abundances cover different ranges, we make quality cuts on our abundance data in the scaled ($[\mathrm{Fe}/\mathrm{H}]$, $[\mathrm{X}/\mathrm{Fe}]$) space, where the transformed abundances have mean 0 and a standard deviation of 1. Since the goal of this paper is to focus on global properties between abundances, age, and R_{birth}, we remove outliers by only selecting stars that have scaled abundances between –4 and 4. Our final sample sizes are 229,045 and 44,359 particles for Chem2d and Chem15d, respectively.

2.2. Birth Properties in the Abundance Plane

Figure 2 shows the simulation data in the $[\alpha/\mathrm{Fe}]$–$[\mathrm{Fe}/\mathrm{H}]$ plane after the selection cuts discussed above. Due to the formation history, both simulations have obvious trends in R_{birth} and age (middle and right columns). For a given value of $[\alpha/\mathrm{Fe}]$, R_{birth} increases as $[\mathrm{Fe}/\mathrm{H}]$ decreases. Similarly, for a given value of $[\mathrm{Fe}/\mathrm{H}]$, age increases as $[\alpha/\mathrm{Fe}]$ increases. As discussed in Buck (2020), the horizontal age gradient and the diagonal radius separation in $[\alpha/\mathrm{Fe}]$–$[\mathrm{Fe}/\mathrm{H}]$ are simply a reflection of star formation in the disk happening at different radii, where metallicity decreases with increasing radii.

The left column of Figure 2 shows the density structure of the two simulations. For Chem2d (top-left), there are density ridges which follow along different age tracks, whereas there is no noticeable substructure in Chem15d (bottom-left), presumably due to its lower mass resolution. The left panel of Figure 2 also shows that the footprint of Chem15d is different
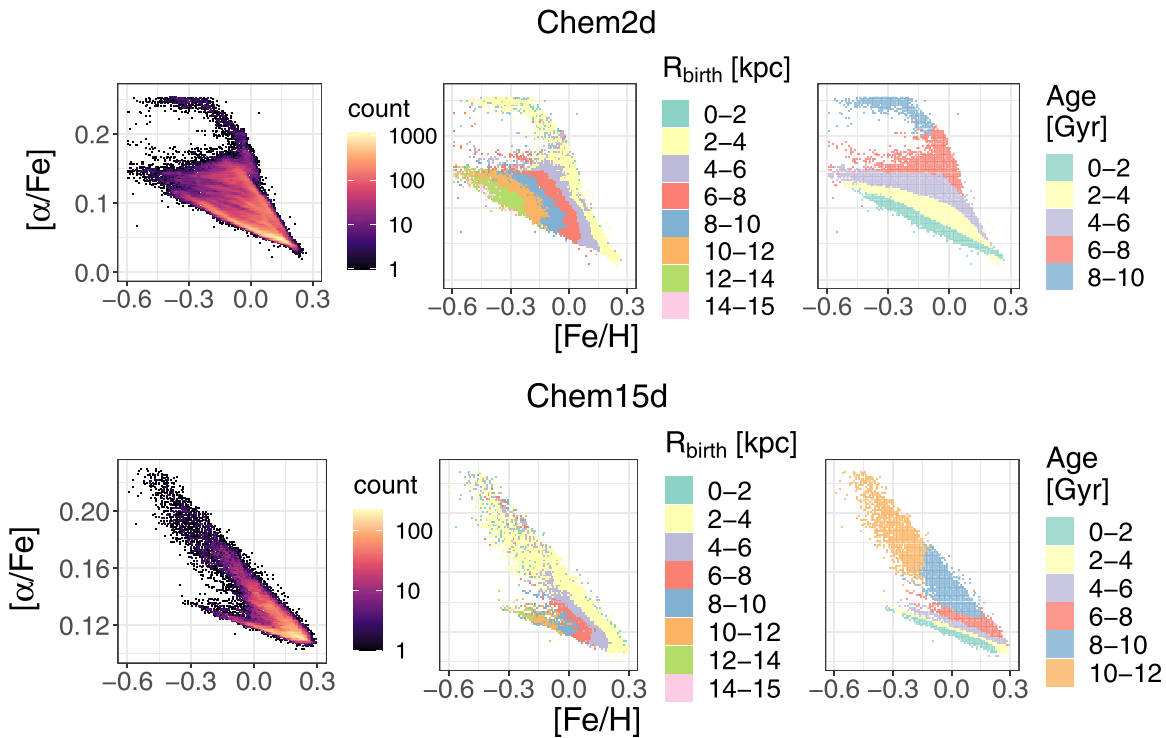
Figure 2. The [α/Fe]–[Fe/H] plane of (top) Chem2d , and (bottom) Chem15d colored by (left) density, (middle) birth radius, and (right) age. The two simulations have a different footprint in this abundance plane (discussed in Section 2), but both clearly have R_{birth} and age trends due to their formation history.

than Chem2d. Most noticeably, the spread in this plane primarily captures high-α stars for Chem15d. The structural differences between the [α/Fe]–[Fe/H] planes of Chem2d and Chem15d are due to slight differences in their formation history (discussed above) and the different set of chemical yields for chemical enrichment (see Buck et al. 2021 for discussion on the impact of yield tables and tracks in this abundance plane).

3. Clustering Methods

We focus on two different clustering methods: agglomerative hierarchical clustering using Ward's minimum variance criterion (Ward 1963) and EnLink (Sharma & Johnston 2009). While both are nonparametric approaches, hierarchical clustering has the advantage of being simpler with only one tuning parameter, the distance metric. On the other hand, EnLink needs two input parameters, but is able to fit complex structures since it has a locally adaptive distance metric.

Unlike other clustering methods, hierarchical clustering and EnLink are nonparametric and thus do not force clusters to fit specific distributions. Additionally, other clustering methods, such as K-means (Hartigan & Wong 1979), require prior knowledge for how many clusters comprise the data, whereas the two methods focused on in this work do not. Particularly in the high-dimensional space of Chem15, where we cannot visualize all 15 abundance dimensions at once, choosing the wrong number of clusters could give rise to misleading results for a method requiring the number of clusters beforehand.

3.1. Hierarchical Clustering: Tree-based Clustering with a Fixed-distance Metric

Following the same methodology Ratcliffe et al. (2020) used with observational red clump DR14 APOGEE data, we use

agglomerative hierarchical clustering using Ward's minimum variance criterion (Ward 1963) as one of the ways to combine the most chemically similar stars. Specifically, we use the Ward2 algorithm described in Kaufman & Rousseeuw (2009) and Murtagh & Legendre (2014). We conceptually describe the algorithm here, and refer the reader to Ratcliffe et al. (2020) for a more in-depth explanation.

The algorithm begins with each star as its own cluster, and at each step we combine the pair of clusters that leads to a minimum increase in total within-cluster variance until only one large cluster containing all the stars remains. The output is a tree showing the combination of groups at each step, called a dendrogram. Thus, the user decides the number of clusters to separate the data into after seeing the linking structure of the data.

3.2. EnLink: Density-based Clustering with an Adaptive Metric

EnLink is a nonparametric hierarchical clustering algorithm built on a locally adaptive distance metric and hence able to identify complex structures in the data. The full data set is first divided via a binary-partitioning algorithm which uses an entropy criterion to preferentially bisect dimensions that contain maximum information ("EnBid"; Sharma & Steinmetz 2006). This approach allows a nonparametric definition of "local" regions in the data set from which the adaptive metric, with flexible scales and orientations, can then be derived (see Sharma & Johnston 2009 for full details). It is this metric that defines the distance between particles subsequently.

EnLink partners the adaptive distance metric machinery with IsoDen (Pfitzner et al. 1997), which is a density-based clustering algorithm. Conceptually, clusters can be thought of as regions around high-density peaks that are separated from one another by lower-density regions. Thus, as we lower the

isodensity contours when examining a high-density region, we stay within the cluster until we encounter a lower-density region that connects to another high-density region. Then, as the isodensity contour continues to lower, a new group encompassing both clusters is formed. Continuing in this fashion forms a hierarchy of density-based parent–child clusters.

EnLink has two user-specified parameters: the number of nearest neighbors used in calculating density (k_{den}) and the threshold significance level when comparing the high and low density levels of parent–child clusters (S_{th}). Since the goal of our analysis is not to find the best clustering but rather to investigate the stability of our results, we choose to vary k_{den} from 30 to 1000, and S_{th} is such that the expected number of groups due to Poisson noise is 0.5, 1, and 2. We did not observe major differences in our results.

4. Results I: Abundance Clusters form Groups of (R_{birth}, Age)

In this section, we investigate the birth properties of groups defined in a two-dimensional (Chem2d; Section 4.1) and 15-dimensional (Chem15d; Section 4.2) chemical abundance space. Since the goal of this work is to explore the ability of using abundances to get to birth properties, we measure this effectiveness by quantifying the separation (conversely overlap) between groups in (age, R_{birth}) for two to 15 clusters determined using both Chem2d and Chem15d.

4.1. Two Abundances Tag Distinct Ages and R_{birth}

As mentioned in Section 3.1, hierarchical clustering produces a dendrogram showing how stars in the abundance space combine, starting from each star being its own group to one cluster containing every star. After the linking structure is determined, the user then specifies how many groups to separate the sample into. Walking down the tree, and thus increasing the number of clusters, corresponds to one group being separated to form two new clusters at each step. For a given number of k groups defined in the two- or 15-dimensional abundance space, we determine the contour level that contains 50% of the stars within each abundance group after projecting into the age–R_{birth} plane. We then calculate the percent that each 50% contour level overlaps with the other $k-1$ group's 50% contour levels by laying down a fine grid and comparing the number of points in just the ith group to the number of points that fall in more than just the ith group. For the ith group, the overlap percentage is defined as the percent of area that is common between the 50% contour region of the ith group and the 50% contour regions of the other $k-1$ groups.

The top-left panel of Figure 3 shows the median of these overlap percentages and standard deviation of the k overlap percentages as a function of the number of groups found in Chem2d using hierarchical clustering. We see that groups defined solely using two abundances show consistently low overlap in birth time and space for up to 10 groups at the 50% contour level (and up to seven groups at the 75% contour level, which is not shown).

The middle and right panels of the top row in Figure 3 show the seven groups in the [α/Fe]–[Fe/H] and R_{birth}–age planes at the 50% contour level. We can see that the groups found using two abundances separate diagonally, both as a function of age

and R_{birth}. We believe that this primarily is a consequence of formation history, as the low-α stars have gradients in abundances, age, and R_{birth}.

4.1.1. Using EnLink to Leverage Density Structure in Chem2d

In the two-dimensional abundance space of Chem2d, we can see streaks of higher-density regions along the different age bins (see top-left of Figure 2). Hierarchical clustering does not leverage the density of the simulation in abundance space and thus is unable to capture the streak formations (see top-middle of Figure 3). Therefore, to include this structure as information in assignment of cluster groups, we employ a density-based clustering method with an adaptive distance metric to use the ridge-like structure in association of groups.

As discussed in Section 3.2, EnLink has the number of nearest neighbors used to determine the density at a point and the maximum number of spurious clusters created by noise as input parameters. We find that the group separation in the age–R_{birth} plane is fairly stable when we focus on allowing either 0.5 or 1 spurious clusters and 30 to 1000 nearest neighbors, with the majority of EnLink groups having a median overlap percent of below 25% at the 75% contour level, and near 0% at the 50% contour level.

Figure 4 shows the 50% contours in the abundance and age–R_{birth} planes for seven groups and parameter settings 260 nearest neighbors and a maximum of one spurious cluster. We can clearly see that EnLink captures the streaks of overdensity in the [α/Fe]–[Fe/H] plane better than hierarchical clustering, and in doing so we see better separation in the age–R_{birth} plane, with a median overlap percentage of only 1%. Since EnLink traces the higher-density ridges, and the ridges trace different age bins, the group separation in the age–R_{birth} plane no longer follows the diagonal trends given by hierarchical clustering. The right panel of Figure 4 shows that groups which follow along the same age track (e.g., the middle-aged blue, orange, and green groups), have similar ages and are separated as a function of birth radius. This causes the blue and green groups to flip from the groups found using hierarchical clustering, but this is an inconsequential difference between the two clustering methods. For an easier way to directly compare the results of hierarchical clustering and EnLink in Chem2d, please see Figure 18 in the Appendix.

Each EnLink group primarily lives in a unique birth time and place, though there is some overlap with the middle-aged groups, possibly due to not fine tuning the algorithm. Overall, we conclude that for the high-resolution Chem2d simulation, using a density-based method with an adaptive metric is desirable for the best results.

4.2. Additional Nucleosynthetic Families and Abundances Provide More Information about Birth Properties

So far we have demonstrated that just two abundances ([α/Fe] and [Fe/H]) can trace separate ages and birth radii. We now investigate how additional abundances, including additional nucleosynthetic families, strengthen this result. The list of abundances and their families is given in Section 2. Due to the difficult problem of estimating density in a high-dimensional space, in addition to the issue of tuning the algorithm, we do not use EnLink to cluster in 15 dimensions, and instead focus on hierarchical clustering.
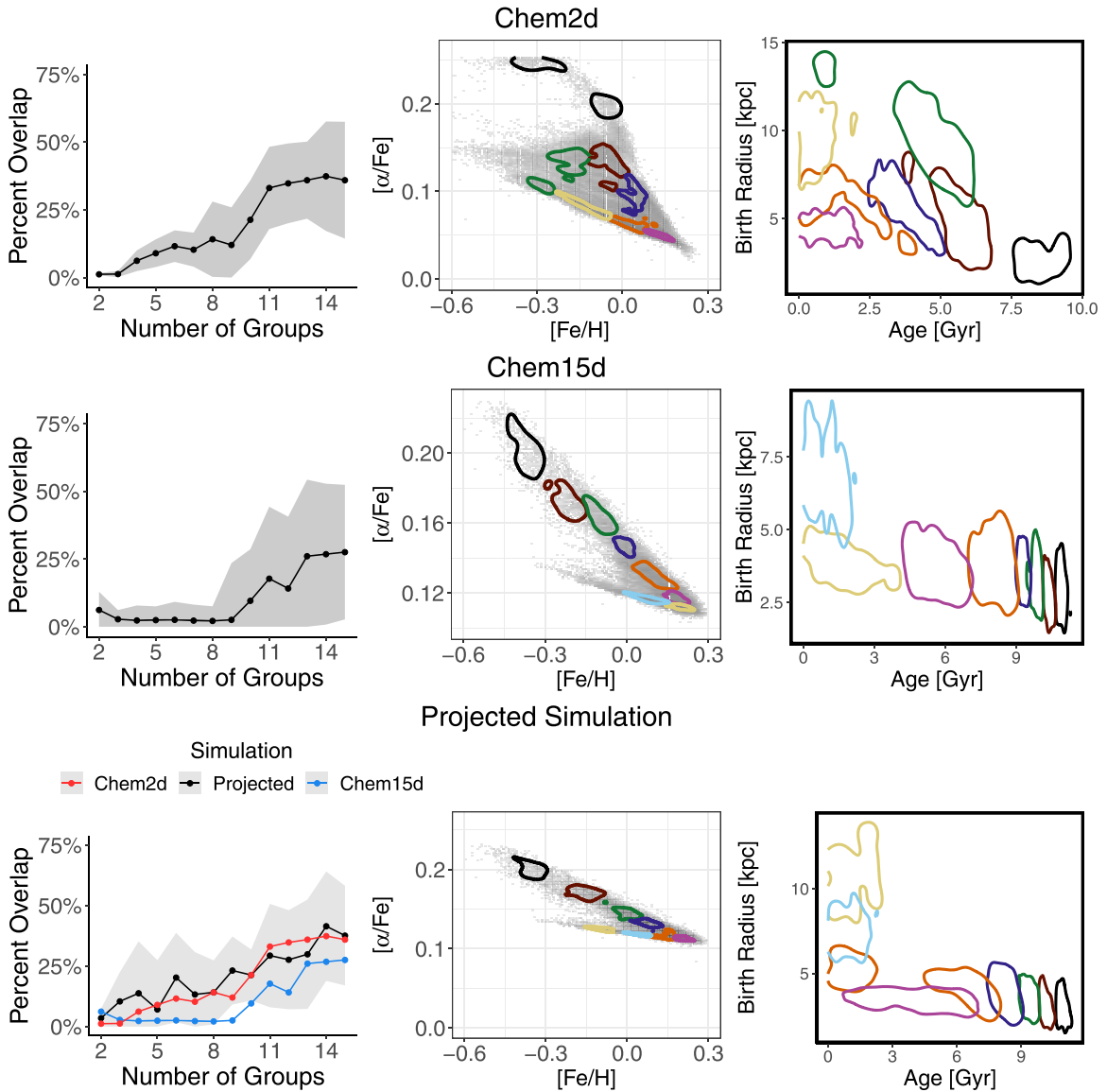
Figure 3. Left: the median overlap percentage for groups in the age–R_{birth} plane as a function of the number of groups determined by hierarchical clustering for (top) Chem2d, and (middle) Chem15d. Each point is determined by calculating the median percent each group overlaps with the other groups at the 50% contour level in the age–R_{birth} plane. The gray ribbon represents one standard deviation about the median overlap percentages. The 50% contour lines of groups projected into the (middle) $[\alpha/\mathrm{Fe}]$–$[\mathrm{Fe/H}]$ plane, and (right) age–R_{birth} plane. There are seven and eight groups for Chem2d and Chem15d, respectively. The bottom row shows the results for clustering in only the $[\alpha/\mathrm{Fe}]$–$[\mathrm{Fe/H}]$ plane of Chem15d (labeled "Projected") in comparison to clustering in the full 15-dimensional chemical space of Chem15d and Chem2d.

As shown in the middle-left panel of Figure 3, groups defined in the chemical space of 15 abundances show more separation in the age–R_{birth} plane than the groups defined in the two-dimensional abundance space of Chem2d. These exhibit separation for 13 groups at the 50% contour level, and eight groups at the 75% contour level (not shown). As shown in the middle and right columns of the middle row of Figure 3, the groups comprised of older stars (which trace the high-α stars) primarily show separation as a function of age, whereas the younger low-α stars show separation in both R_{birth} and age. This shows that high-α stars were all born near the Galactic center, whereas the low-α stars were born at different radii and times throughout the galaxy.

Comparing these results to Chem2d (top-left of Figure 3), we see that groups defined in the 15-dimensional abundance space of Chem15d using hierarchical clustering overlap less in age and R_{birth}. In particular, for up to nine groups, the 15-dimensional groups are predominantly distinct in birth time and space at the 50% contour level, whereas in two dimensions the groups have some overlap even for as few as four groups.

4.3. Comparing Groups in Two Dimensions and 15 Dimensions

In the previous section (Section 4.2) we showed that groups defined in the 15-dimensional abundance space of Chem15d showed more separation in age and R_{birth} than groups defined in the two-dimensional abundance space of Chem2d. Here we show the results of clustering in just the $[\alpha/\mathrm{Fe}]$–$[\mathrm{Fe/H}]$ plane of Chem15d are consistent to those of Chem2d.
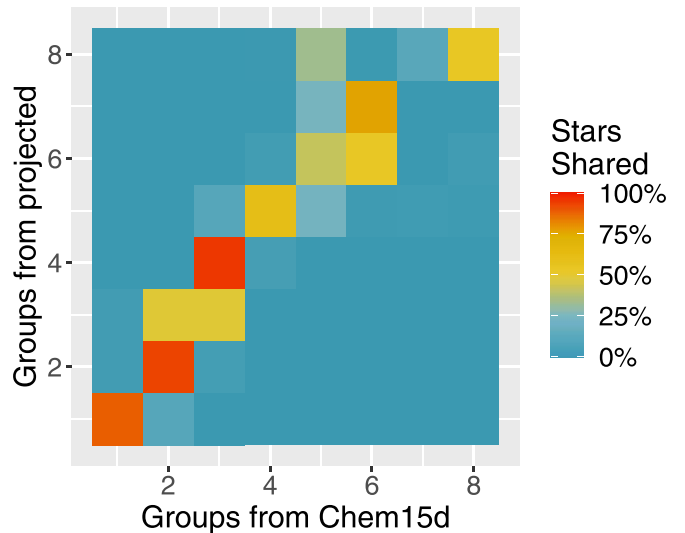
The bottom-left of Figure 3 shows that the amount that abundance groups overlap in the age–R_{birth} plane is similar between Chem2d (red) and the $[\alpha/\mathrm{Fe}]$–$[\mathrm{Fe/H}]$ plane of Chem15d (black, labeled "projected"). As the simulations are
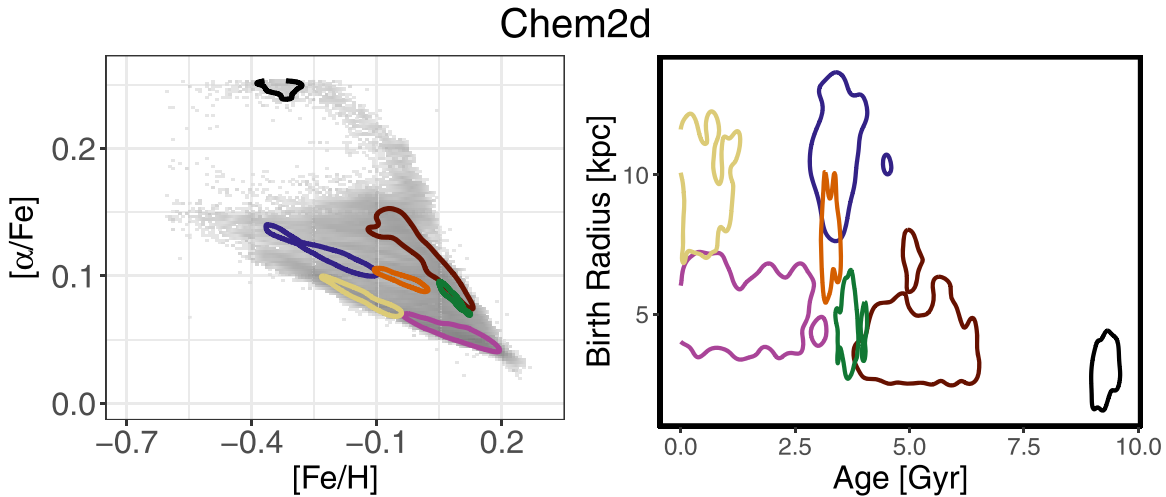
Figure 4. Seven groups found using the density-based nonparametric method EnLink in the two-dimensional abundance space of Chem2d projected into the [α/Fe]–[Fe/H] and age–$R_{\rm birth}$ planes. All stars are assigned to a group, and each contour captures 50% of the stars within the group. The groups now follow the density ridges discussed in Section 2.2 and show more separation in the age–$R_{\rm birth}$ plane than those groups found using hierarchical clustering. The median amount each group overlaps is 1%.

split into more groups, both consistently show less distinction in (age, $R_{\rm birth}$). On the other hand, the groups defined in the full 15-dimensional abundance space of Chem15d retain separate ages and birth radii for up to nine groups. This shows that more separate birth information is retained in abundance groups when more abundances are included, and therefore we claim abundance groups being more separate in age and $R_{\rm birth}$ is due to additional abundance information and not an artifact of different simulation history or resolution.

Figure 5 compares eight groups defined in the [α/Fe]–[Fe/H] plane of Chem15d to eight groups defined in the full 15-dimensional abundance space of Chem15d. The groups are arranged in order of age, with group 1 being the oldest and group 8 being the youngest. The oldest groups share the most stars between the two simulations, whereas the middle-aged and youngest groups are more muddled.

4.4. A Grid in [α/Fe]–[Fe/H] Separates Birth Properties Less Effectively

So far we focused on how chemically similar stars differ in birth time and space using clustering methods. Now we give motivation for why use of a clustering method is needed for grouping stars.

Figure 6 shows the median and standard deviation of the percent group overlap for both Chem2d and Chem15d when separating the stars using different grouping methods. We compare hierarchical clustering with laying down a Cartesian grid in the [α/Fe]–[Fe/H] plane. The number of [α/Fe]–[Fe/H] bins are chosen to produce four, seven/eight (for Chem2d/Chem15d respectively), and 12 groups. We also show the results of EnLink for Chem2d. The left panel of Figure 6 shows that when only two dimensions of abundance information are known, we do not gain any more knowledge of birth properties from using a simple clustering method than if we were to create bins by laying down a Cartesian grid across the [α/Fe]–[Fe/H] plane. The groups retain more separate birth properties when leveraging density with an adaptive distance metric, however the errors are higher than that of binning and hierarchical clustering.

Figure 5. Comparison of the number of stars shared between eight groups defined in all 15 abundances versus just ([α/Fe], [Fe/H]) of Chem15d. The groups are arranged in order of age, with group 1 being the oldest and group 8 being the youngest. The percent of stars shared between groups is determined by counting the number of stars found in the projected group that are in the 15-dimensional group, and then dividing by the number of stars in the projected group.

The right panel of Figure 6 reveals that when higher-dimensional abundance information is available, there is a noticeable difference between only looking at stars separated using a grid in [α/Fe]–[Fe/H] versus using a clustering method in the full abundance space. Not only does clustering in 15 dimensions produce less overlap in the age–$R_{\rm birth}$ plane, but based on the smaller standard deviation, groups defined using hierarchical clustering produce consistently small overlap between groups whereas groups defined by binning in the [α/Fe]–[Fe/H] plane produce a large range of amount of groups that overlap in birth time and place.

Figure 7 compares the mean and standard deviation of (age, $R_{\rm birth}$) of each group determined using two different methods: hierarchical clustering (left) and binning in the [α/Fe]–[Fe/H]
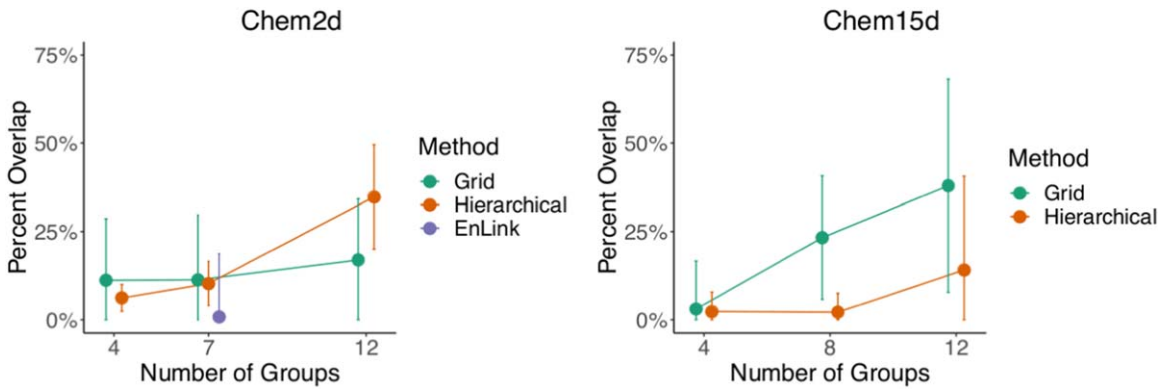
Figure 6. Each point and error bar represents the median and standard deviation of the amount that groups overlap with one another in the age–R_{birth} plane. Comparing three different ways of combining stars in chemical space (a grid laid out in the [α/Fe]–[Fe/H] plane, hierarchical clustering, and EnLink) shows that if we only have two abundances available (left), then groups of chemically similar stars determined via hierarchical clustering and gridding on average show similar separation in the age–R_{birth} plane. Leveraging the density in the abundance plane allows for even better separation in birth time and space. When we include more abundances and nucleosynthetic channels (right), we find that hierarchical clustering done in 15 dimensions yields more distinct groups in age–R_{birth} than gridding in the visual two dimensions. Note that EnLink was inconclusive in 15 dimensions due to the curse of dimensionality and also only reported seven groups for Chem2d, and therefore no EnLink result is shown for Chem15d or for four and 12 groups for Chem2d.
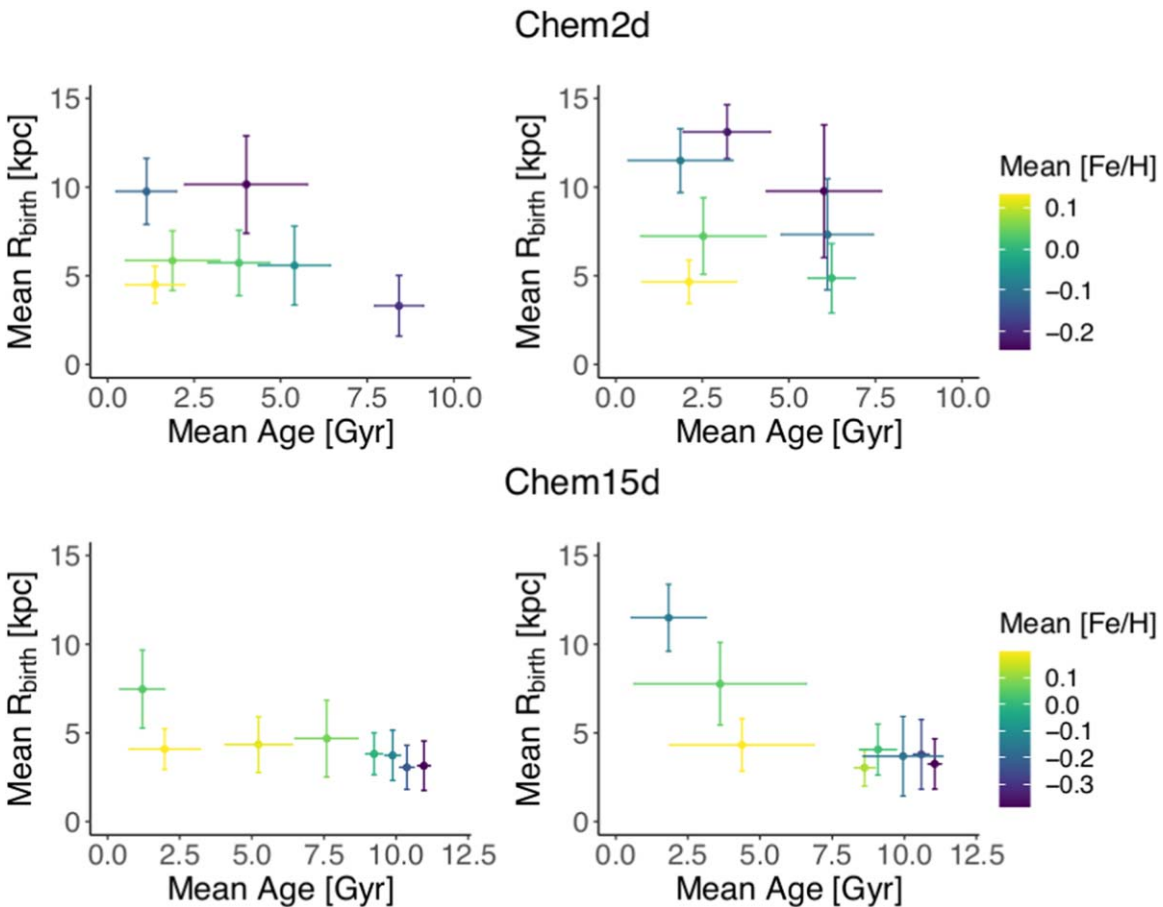

Figure 7. Seven and eight groups separated using (left) hierarchical clustering (right) and gridding in the [α/Fe]–[Fe/H] plane for (top) Chem2d and (bottom) Chem15d, respectively, projected into the R_{birth}–age plane. Each point represents the mean (age, R_{birth}) for each group, colored by the mean metallicity. Error bars shown are 1σ standard deviations. Groups defined using hierarchical clustering show a metalicity gradient for a given R_{birth}, suggesting that the groups are physically meaningful. Groups defined in mono-[α/Fe]–[Fe/H] bins do not show a metalicity gradient, and have larger dispersions in age.

plane (right) to create a Cartesian grid. This figure demonstrates that groups defined using hierarchical clustering preserve physical interpretation and agree with expected trends, namely we see a clear metallicity gradient as a function of age for a given R_{birth} when stars are separated using hierarchical clustering. The trend is not as obvious to see when stars are separated in the [α/Fe]–[Fe/H] plane with a grid. Additionally, we see that groups defined using a clustering method tend to have less overlap, and the overlap they do have is consistent among nearly all the groups. However, for the groups defined

Chem2d

Chem15d
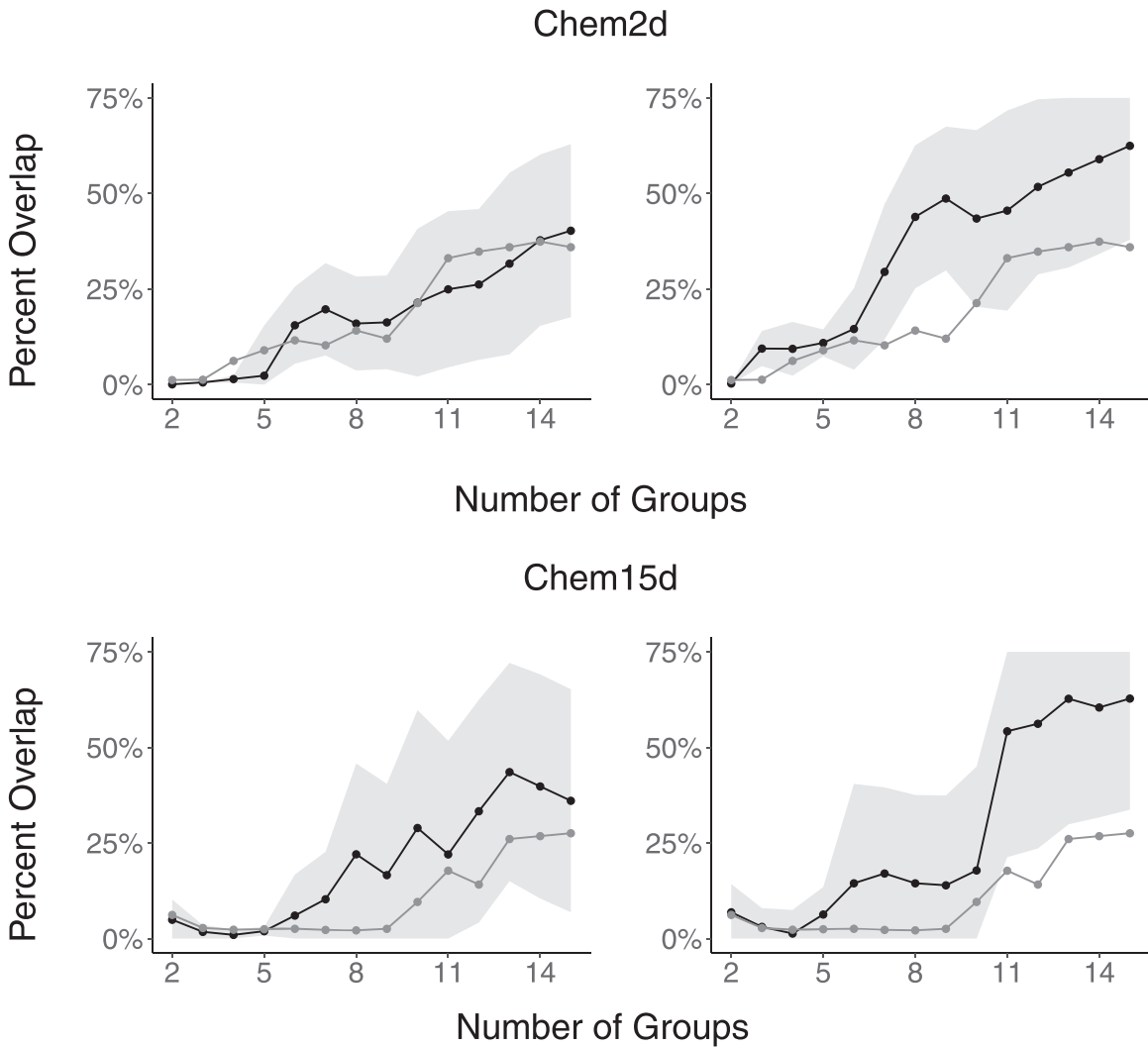
Figure 8. The black points and line correspond to the median percent that groups defined in an error-convolved (top) two-dimensional and (bottom) 15-dimensional abundance space overlap in the age–R_{birth} plane. For each star, new abundance measurements are drawn from a normal distribution with the true abundance value as the mean and a standard deviation of (left) 0.02 dex and (right) 0.05 dex. The gray ribbon captures one standard deviation about the median percent that groups overlap in the age–R_{birth} plane. When simulating errors of 0.02 dex, groups still stay separated in the age–R_{birth} plane and the separation is comparable to that given by groups found with no abundance error (gray points and line). Adding an error of 0.05 dex in each dimension affects our ability of finding separate birth information slightly more in Chem2d than Chem15d.

via a grid, the overlap between groups is irregular, with some groups being mainly separate and others completely overlapping multiple groups.

Additionally, particularly for Chem15d, the age dispersion for each group when defined using hierarchical clustering is much smaller than when groups are separated using a grid. This again shows that the groups found using hierarchical clustering represent different physical groups in time (i.e., the groups have differing properties in birth age and location). Establishing a connection between those groups and star formation episodes or satellite passages triggering star formation can be done in future work.

5. Results II: Clustering in Observational Data

Section 4 showed that under certain formation conditions and no observational limitations that abundances trace stellar birth information R_{birth} and age. In this section we examine some of the consequences of observational limitations.

5.1. Incorporating Measurement Uncertainty

Current-day element abundance measurements are reported with uncertainties of $\approx$0.02–0.05 dex (Ahumada et al. 2020; Jönsson et al. 2020). Consequently, we examine how the clustering changes once we incorporate errors in the chemical abundances and how this impacts our ability to trace birth properties with observational data.

For each star in both Chem2d and Chem15d, we redraw a new set of element abundances, each from a Gaussian distribution where the standard deviation is representative of the measurement uncertainty. We test two precision regimes: $\sigma_{\text{err}} = 0.02$ and 0.05 dex. The left column of Figure 8 shows the overlap in the age–R_{birth} plane of groups defined in Chem2d (top) and Chem15d (bottom) when the abundances are modified with an equivalent of 0.02 dex error in each abundance direction (black points and line). With the current best observational error of $\sigma_{\text{err}} = 0.02$ dex, the groups defined in both the two-dimensional and 15-dimensional abundance space retain separate birth properties, similar to the overlap
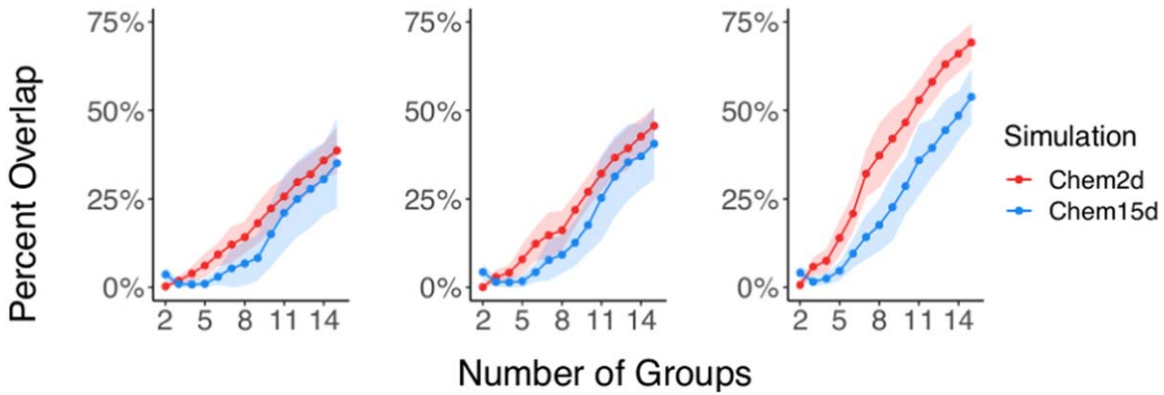
Figure 9. The mean percent of 50 subsampling replications that groups defined in a 30,000 stellar sample of an error-convolved two-dimensional (red) and 15-dimensional (blue) abundance space overlap in the age–R_{birth} plane. The errors added are equivalent to the observational best case scenario (middle, 0.02 dex) and average observational error (right, 0.05 dex) in each dimension. Left has no error added. The ribbon shows one standard deviation in percent overlap for the 50 Monte Carlo simulations.

found when the simulations have no error added (gray points and line). When abundances are redrawn with observational errors of $\sigma_{err} = 0.05$ dex for each data point (right column), we find that the majority of groups from the higher-dimensional simulation retain more separate birth information compared to those found in Chem2d.

5.2. Modifying the Sample Size

While current large surveys have captured many millions of stars, we have so far examined only about 30,000 stars with precise abundance measurements, within a narrow evolutionary state (e.g., APOGEE DR14 RC catalog; Bovy et al. 2014). To test more generally how useful chemical abundances are for linking to birth properties, we need to examine the impact of sample size.

So far in this work, we have been working with ∼229,000 and ∼44,000 star particles for the Chem2d and Chem15d simulations, respectively. Now we examine how the groups defined in abundance space, both with and without the addition of errors, change in the age–R_{birth} plane for a random subsample of 30,000 stars throughout the whole disk. Figure 9 shows the percent overlap for both the subsampled Chem2d (red) and subsampled Chem15d (blue) with no error (left), the equivalent of 0.02 dex error (middle) and 0.05 dex error (right) in each dimension. In order to test the consistency of these results, we subsample with 50 replications. The mean and standard deviation of the median percent overlap are shown as a point and ribbon.

We can see from Figure 9 that as error increases, the amount the Chem2d groups overlap in the age–R_{birth} plane also increases. This indicates that for the sample size and error used in paper I of this series (Ratcliffe et al. 2020), only two abundance dimensions are not enough to recover separate birth time and place groups observationally for more than ≈6 groups.

On the other hand, the groups in the subsampled Chem15d are less affected by error than the subsampled Chem2d . We can see that as error increases, the amount each group overlaps with other groups in the age–R_{birth} plane stays more consistent than the groups in the subsampled Chem2d. While the groups from both simulations trace less birth information in the presence of subsampling, we see that the recovery of separate

birth place and time groups is still possible with the inclusion of more abundances and nucleosynthetic families, especially for a smaller number (≤≈10) of groups.

6. Results III: Modeling How Observable Stellar Properties Relate to Birth Properties

In the previous sections, we showed that groups in abundance space occupy different regions in birth time and space. Now we quantify how chemical abundances relate to age and R_{birth}.

The left panel of Figure 10 shows the [Fe/H] running mean of Chem2d across R_{birth}, separated and colored by age bins. In order to visually examine if an age–abundance–R_{birth} relationship exists, we examine running [X/Fe] means of solar metallicity stars (which are taken to be stars within the gray band) as a function of R_{birth}, divided into age bins. The right panel of Figure 10 shows that for Chem2d, there is a strong age–[α/Fe] relation in solar metallicity stars that is approximately quadratic for older stars and linear for younger stars. Thus, given a fixed [Fe/H], we anticipate that ages can be determined from abundances. Similarly for Chem15d shown in Figure 11, each age group has its own unique polynomial trend in [X/Fe]–R_{birth}.

This visual analysis done in Figures 10 and 11 leads to the conclusion that ages can be determined from abundances alone. To quantify this relationship, we use a simple second-order polynomial to estimate age from ([X/Fe], [Fe/H]). The model for Chem2d is

$$
\begin{aligned}
age = {} & a_1[\text{Fe/H}]^2 + a_2[\alpha/\text{Fe}]^2 \\
& + a_3[\text{Fe/H}] \times [\alpha/\text{Fe}] + a_4[\text{Fe/H}] \\
& + a_5[\alpha/\text{Fe}] + a_0,
\end{aligned}
$$

where the a_i's are the coefficients determined using a training set that is 75% the size of the data set. The model for Chem15d is similar, however with the inclusion of more abundances. The left column of Figure 12 shows the inferred age from the polynomial regression versus the true age of stars in the simulation test set. Even with just two abundances (top row), we are able to estimate age within ±0.52 Gyr. With the addition of more abundance information (bottom row), we find
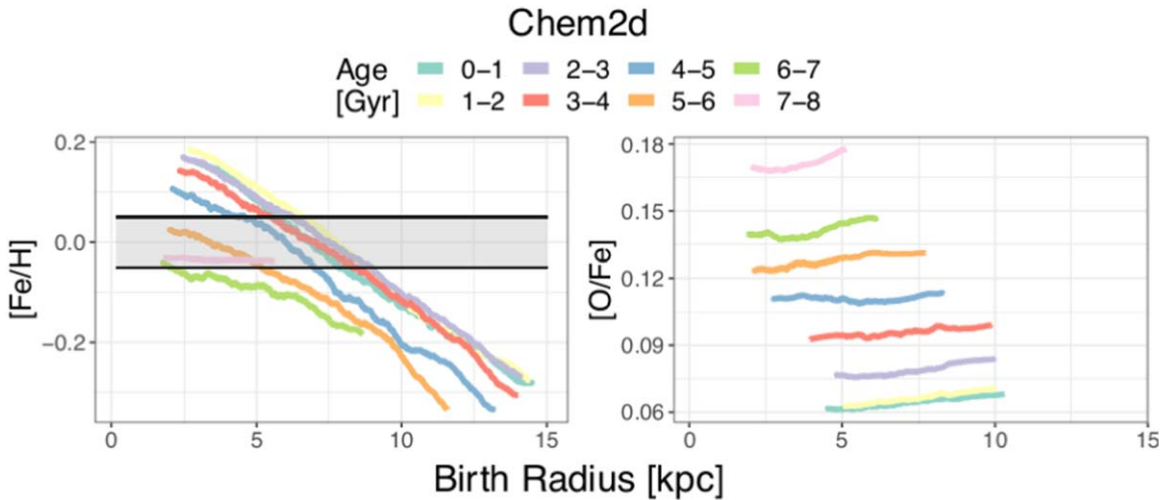
Figure 10. Left: the running mean of [Fe/H] across $R_{\rm birth}$ colored by the associated 1 Gyr age bin for Chem2d. The black lines and gray area mark off the solar metallicity stars, which we consider to be ± 0.05 dex in [Fe/H]. Right: the running mean of [O/Fe] of the solar metallicity stars across $R_{\rm birth}$ colored by age bin selected from within the horizontal lines at left. We see that for a given bin of metallicity, stars of different ages separate out and form either approximately quadratic or linear relations.

that we are able to accurately estimate age from 15 abundances to within ± 0.06 Gyr.

We also wish to test how well we can quantify the relationship between age, $R_{\rm birth}$, and abundances. Given the low $R_{\rm birth}$ dispersion in the [Fe/H]–age plane shown in Figure 13, we use a second-order polynomial model to estimate $R_{\rm birth}$ from ([Fe/H], age). The model for both simulations is

$$r_{\rm birth} = a_1 [{\rm Fe/H}]^2 + a_2 \, {\rm age}^2 + a_3 [{\rm Fe/H}] \times {\rm age}$$
$$+ \, a_4 \, {\rm age} + a_5 [{\rm Fe/H}] + a_0,$$

where the a_i's are the coefficients determined using a training set that is 75% the size of the data set. The right panels of Figure 12 reveal that we can predict the test set birth radii to within ± 1.24 kpc for Chem2d (top) and within ± 1.17 kpc for Chem15d (bottom). The inclusion of the additional abundances increases the accuracy by 0.06 kpc for Chem15d and only 0.01 kpc for Chem2d. We again see that additional abundance information helps inform more about stellar birth properties. However, the difference is not as drastic as it was in estimating age.

While we do not fit for the best model, we believe that this simple second-order polynomial relationship between age, abundances, and $R_{\rm birth}$ cannot be drastically improved upon. For a given value of [Fe/H], [X/Fe], and age, we find that the intrinsic dispersion in $R_{\rm birth}$ is ≈ 1.1 kpc for Chem15d and ≈ 1.2 kpc for Chem2d. Thus, ages and abundances alone will not be able to estimate $R_{\rm birth}$ more accurately. This could be due to asymmetries causing abundance distributions to not lie in perfect annuli about the Galactic center, reducing the tightness of the relationship between $R_{\rm birth}$ and abundances.

With the inclusion of 0.05 dex abundance error, we find that our age estimates decrease in accuracy to about ± 0.76 Gyr for Chem15d. With the addition of 0.05 dex abundance error and age error of 30%, the $R_{\rm birth}$ accuracy of Chem15d decreases to ± 1.31 kpc. This shows that according to our model, estimating $R_{\rm birth}$ from abundances and age is less sensitive to noise than estimating age from abundances.

7. Discussion: Implications for Future Applications to the Milky Way

Paper I of this series (Ratcliffe et al. 2020) used hierarchical clustering in the 19-dimensional abundance space of 30,000 red clump stars in the Milky Way cataloged by APOGEE DR14. In that work, we found that up to six groups have statistically significant different mean ages and distinct spatial distributions. The goal of this work is aimed to interpret those results and determine if groups observed in chemical space correspond to physically meaningful groups. With the use of simulations, we are able to test the potential and current ability of linking current stellar properties ([Fe/H], [X/Fe]) to their birth properties ($R_{\rm birth}$, age). We wish to emphasize that the stellar groups found by clustering algorithms in this work represent stars that are chemically similar in abundance space which can be linked to different ($R_{\rm birth}$, age); they do not represent individual stellar birth clusters.

7.1. Empirical Context

The simulations used in this work (g7.55e11 with two abundances and g8.26e11 with 15 abundances) are Milky Way analogs from the NIHAO-UHD suite. Both simulations were bulge-dominated systems up to redshift $z \geqslant 1$ with prominent stellar disks forming 7–8 Gyr ago. The formation of the α sequences was due to a gas-rich merger, with the high-α sequence forming during the early galaxy and the low-α sequence forming after the merger. The main differences between these simulations are a slightly different formation history (sampling valid formation histories of Milky Way–like galaxies) as well as an updated chemical enrichment prescription for the Chem15d model galaxy. The modifications made to chemical enrichment prescriptions are described in Buck et al. (2021) and enabled us to follow 15 different elements while at the same time leaving global galaxy properties such as star formation history, stellar mass, and disk size unaffected. We believe that these simulations are representative of the Milky Way due to their formation history and age gradient in the abundance plane.

Even though simulations provide us with particles and not individual stars, they allow us to examine the relationship
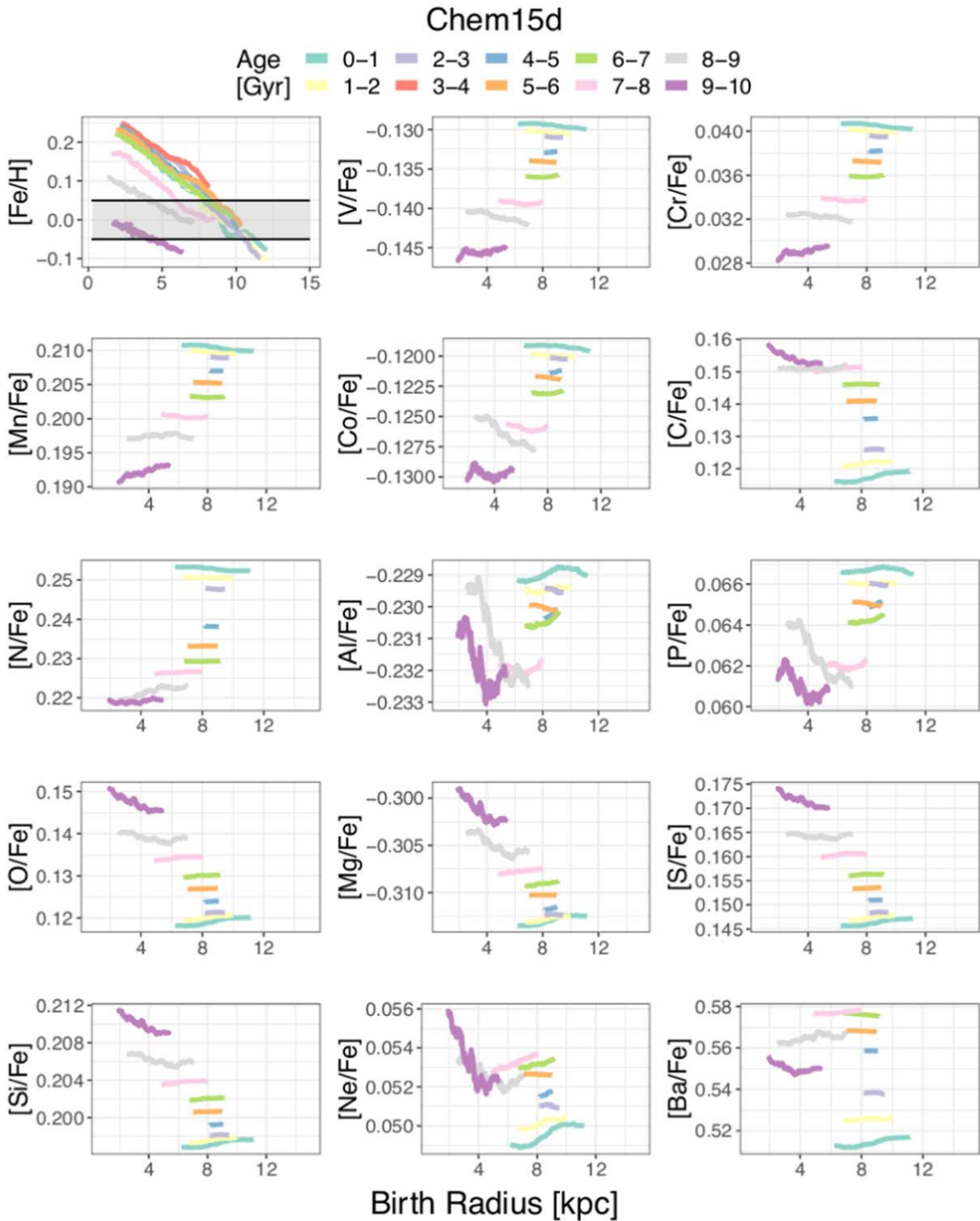
Figure 11. Top-left: the [Fe/H] running mean across R_{birth} colored by age bin for Chem15d. The black lines and gray area mark off the solar metallicity stars, which we consider to be ± 0.05 dex in [Fe/H]. All other plots show the running mean of [X/Fe] of the solar metallicity stars across R_{birth} colored by age. Similar to Chem2d shown in Figure 10, solar metallicity stars of different ages separate into approximately quadratic or linear curves.

between observable chemical properties, age, and birth location. Thus we focus on disk particles in our work.

7.2. Clustering Approaches

In this work, we focused on three ways of grouping stars: hierarchical clustering, EnLink, and binning in the $[\alpha/\text{Fe}]$–[Fe/H] plane. Section 4.4 shows that binning in just ($[\text{Fe/H}]$, $[\alpha/\text{Fe}]$) does not effectively link abundance information to birth

properties, especially when there is higher-dimensional abundance data available.

Of the clustering methods we explore, hierarchical clustering is advantageous for observational work. Section 4.1.1 shows that leveraging density with an adaptive distance metric in Chem2d is the best way for chemical groups to correspond to distinct groups in the age–R_{birth} plane. However, due to the curse of dimensionality, leveraging density in a 15-dimensional space is difficult and unrealistic. We attempted to run EnLink
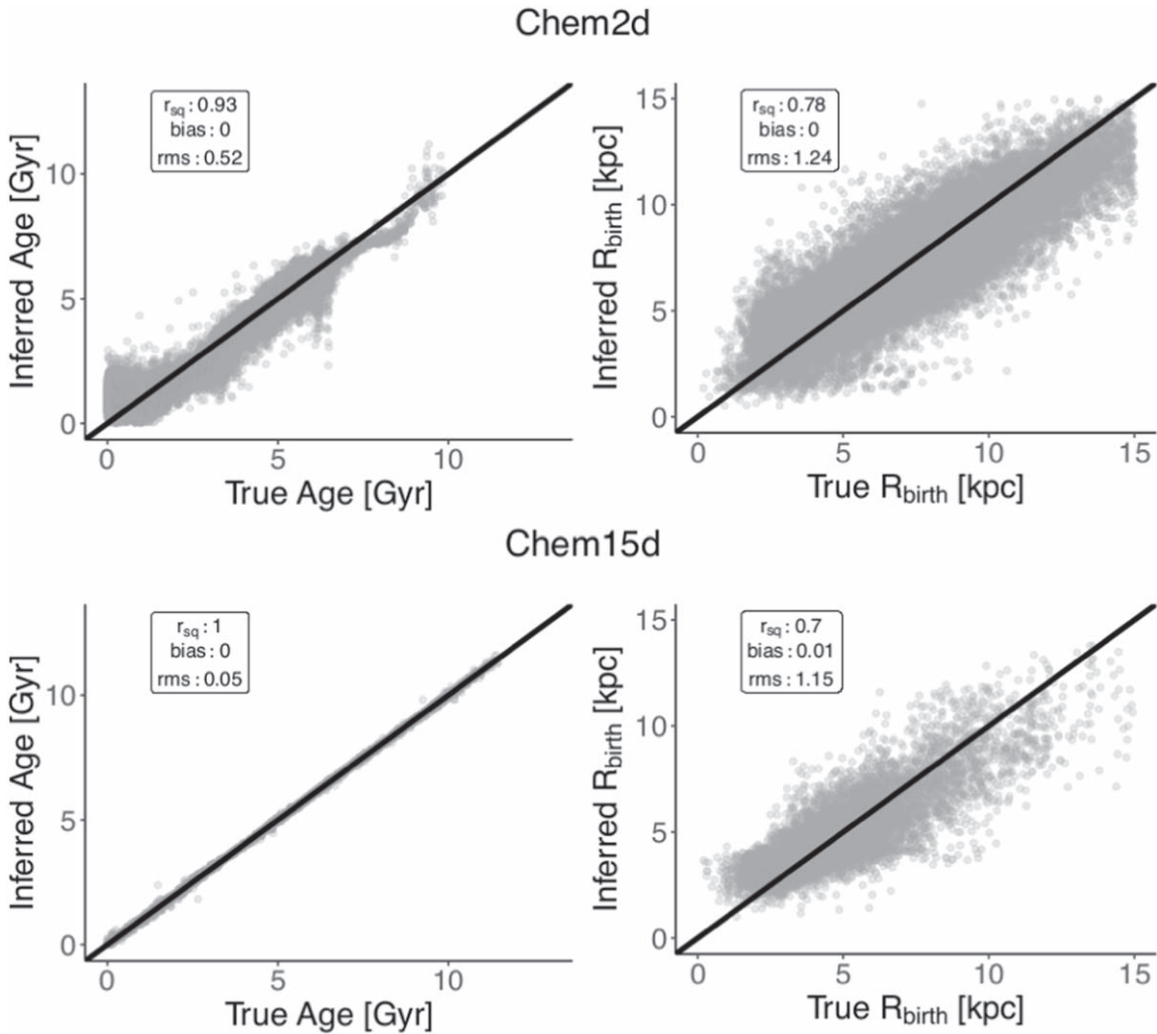
Chem2d

Chem15d

Figure 12. Left: the inferred age of (top) Chem2d and (bottom) Chem15d using a second-order polynomial in ([Fe/H], [X/Fe]) plotted against the true age of the star. With the additional abundance information provided in the Chem15d simulation, we are able to accurately and precisely estimate age, showing that abundances are chemical clocks. Right: inferred R_{birth} of (top) Chem2d and (bottom) Chem15d using a second-order polynomial in ([Fe/H], age) plotted against the true R_{birth} of the stars. With just [Fe/H] and age, we can infer a star's R_{birth} to within just over 1 kpc.

in the 15-dimensional chemical abundance space of Chem15d; however, the clustering results were inconclusive and did not define many stars to a cluster. Additionally, for EnLink (or any other density-based clustering method) to be used correctly on Milky Way catalogs, the survey-selection function would need to be accounted for, as the selection function would possibly alter the distribution of stars in abundance space. Furthermore, we found that EnLink performed poorly when the density structure in the two-dimensional abundance plane of Chem2d vanished after the addition of observational uncertainty.

7.3. Likelihood of Success: Comparison to Other Simulations

In this paper, we have focused on the relationship between abundances and birth properties of stars when the chemical bimodality is caused by a merger and successive dilution of the interstellar medium. We also consider the relationship when a galaxy is formed by clumpy star formation[8]. We examine the

---

[8] Note, the galaxies simulated within the NIHAO project also go through a clumpy phase (Buck et al. 2017) in agreement with observed clumpy galaxies at high redshift. However, for the NIHAO feedback scheme those clumps are agglomerations of young stars and only appear in stellar light not in stellar mass.

N-body and smooth-particle hydrodynamics simulation of the formation of an isolated galaxy outlined in Beraldo e Silva et al. (2021). Star-forming clumps at high redshift start forming low-α stars, then quickly self-enrich in α-elements due to their high-star-formation-rate density and produce a high-α sequence while a low-α sequence is produced by radially distributed star formation. After about 4 Gyr, the clumps become less efficient, and the high-α sequence stops growing (Clarke et al. 2019). For more detailed information of the simulation, see Beraldo e Silva et al. (2021) and Fiteni et al. (2021).

Figure 14 shows the [α/Fe]–[Fe/H] plane for a 230,000 particle subsample of simulation M2_c_nb, which undergoes clumpy star formation. Similar to Chem2d and Chem15d, there is a linear trend between the abundances and R_{birth}, where R_{birth} decreases as [Fe/H] increases. Age, however, does not appear to have a simple relationship between the two abundances. For instance, age decreases as [α/Fe] decreases for solar metallicity stars at higher values of [α/Fe], but the relationship is reversed for lower [α/Fe].

We find that in this simulation, age and [Fe/H] are able to predict R_{birth} within ± 0.72 kpc, about 40% better than the
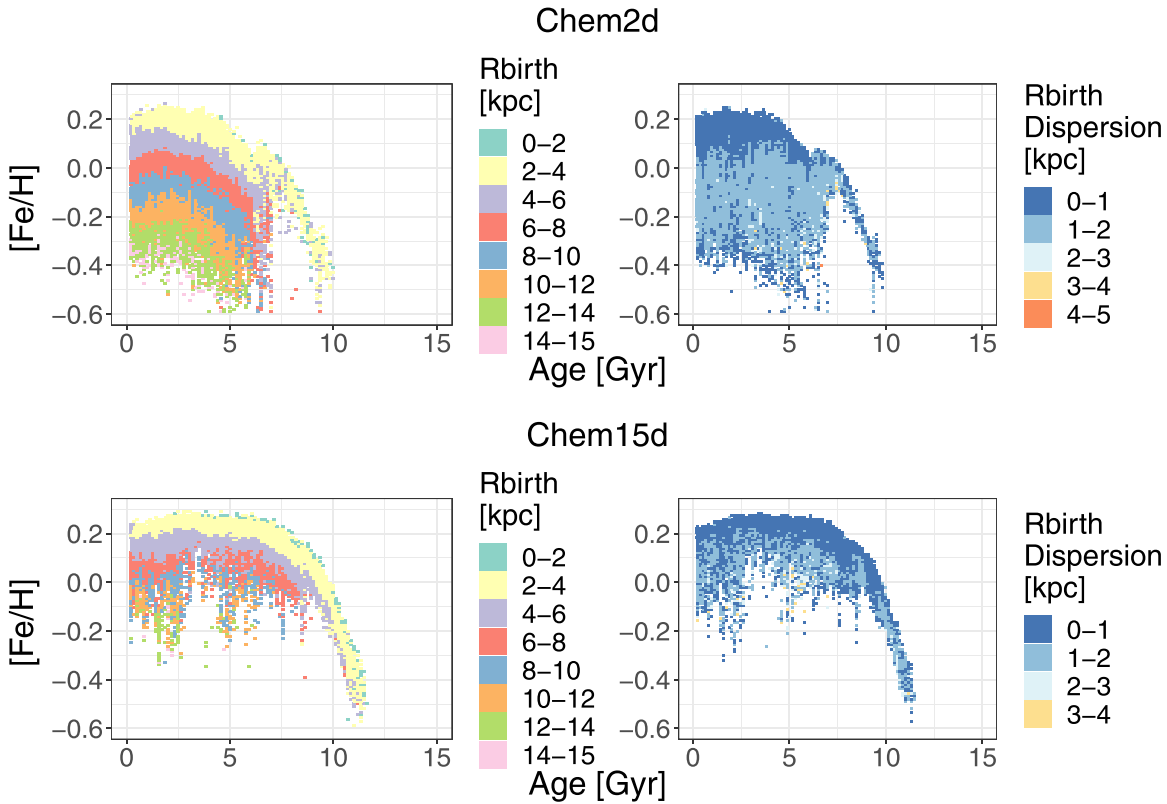
Chem2d

Chem15d

Figure 13. The [Fe/H]–age plane colored by (left) R_{birth} and (right) R_{birth} dispersion for (top) Chem2d and (bottom) Chem15d. The low dispersion in this plane indicates that [Fe/H] and age alone can determine R_{birth} accurately.
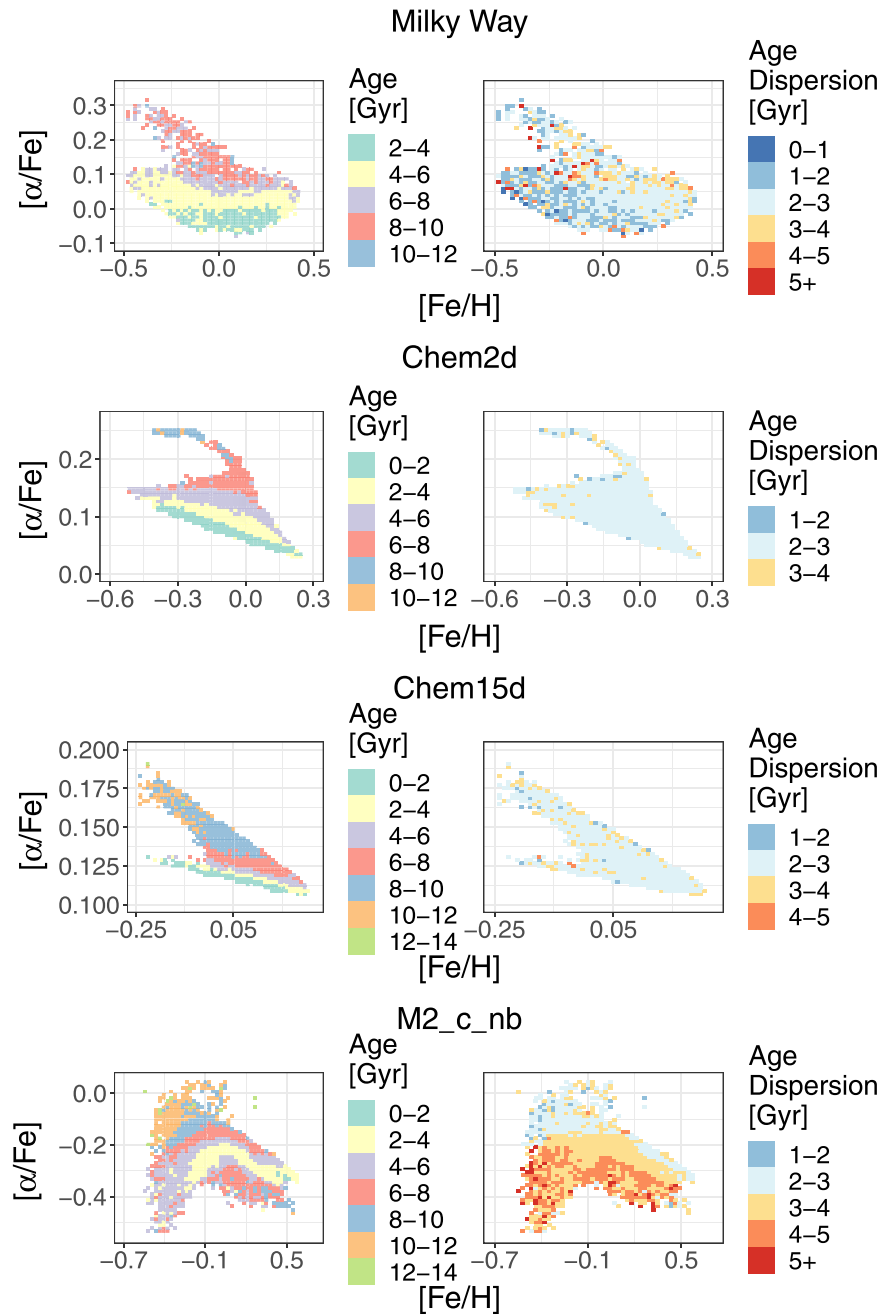
M2_c_nb
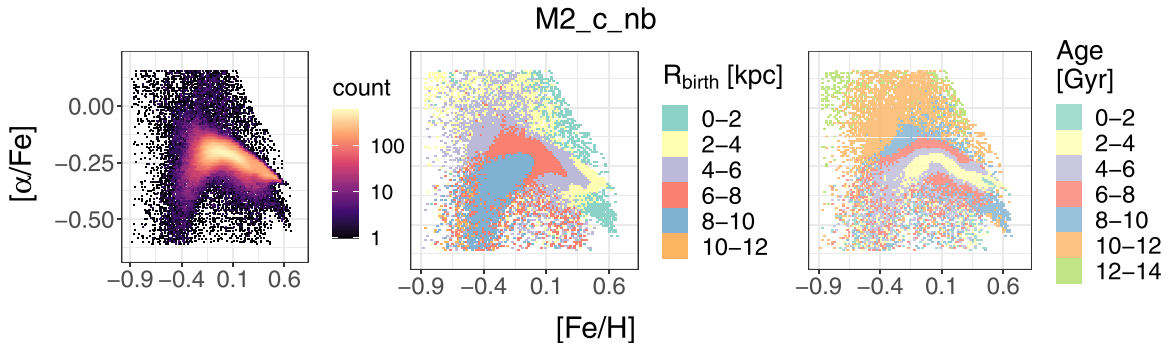
Figure 14. The [α/Fe]-[Fe/H] plane colored by (left) density, (middle) R_{birth}, (right) age for the Beraldo e Silva et al. (2021) simulation using clumpy star formation. The formation history produces quadratic age and R_{birth} distributions in the abundance plane, and causes the simple relationship between age, R_{birth}, and clustered abundances to disappear.

precision for the simulations focused on in this work. Again, we also see that the addition of other abundances does not notably change our ability to estimate R_{birth}, where the accuracy only increases by 0.01 kpc when [α/Fe] is included in the regression. This shows that the formation history in all three simulations sets an underlying relationship with age, [Fe/H], and R_{birth}, where if we know the metallicity and age of a star, we can determine where it was born.

However, while [Fe/H] and age are a link to R_{birth}, in this particular simulation the star formation history gives rise to a more complex relationship between the abundances ([Fe/H], [α/Fe]) and age (Figure 14). Therefore, chemically similar groups of stars identified using hierarchical clustering no longer correspond to separate groups in the age–R_{birth} plane in this scenario.

In order to determine the ability to extend our conclusions to the Milky Way, we compare the Milky Way's age and age dispersion in the [α/Fe]–[Fe/H] plane to the three simulations with simulated "observational" ages by redrawing from a normal distribution with their true age as the mean and a standard deviation of $\approx$2.6 Gyr, the median uncertainty of low-α stars from the Lu et al. (2021) catalog. When simulating observational ages for Chem2d and Chem15d, the dispersion in age across [α/Fe]–[Fe/H] is uniformly low (see middle rows of Figure 15), while M2_c_nb has an increase in age dispersion as [α/Fe] decreases (see bottom row of Figure 15). The Milky Way (shown in the top row of the same figure using ages and abundances from Lu et al. 2021) has a consistently small dispersion in age of about 2–3 Gyr, with the dispersion being slightly smaller for low ([Fe/H], [α/Fe]), the reverse of

Figure 15. Top: the Milky Way in the [α/Fe]–[Fe/H] plane using Lu et al.'s (2021) ages and abundances. The right panel shows the standard deviation of ages within each bin. The ages for (second from top) Chem2d, (third from top) Chem15d and (bottom) M2_c_nb are redrawn from a normal distribution with their true age as the mean and a standard deviation equivalent to the median uncertainty of low-α stars from the Lu et al. (2021) catalog. Only bins containing at least 10 stars are shown in each plot above.

M2_c_nb's dispersion trends. On the other hand, similar to the Chem2d and Chem15d simulations focused on in this paper, the age of the stars in the Milky Way increases as [α/Fe] increases for a given value of [Fe/H].

The question of which simulation most closely matches the Milky Way's star formation history still requires further investigation. The selection cuts described in Section 2.1 produce different density trends in the [α/Fe]–[Fe/H] plane for the three simulations and observational data, where some samples have both high- and low-α stars (e.g., Chem15d), while others primarily consist of stars with lower values of

[α/Fe] (e.g., M2_c_nb). There is room for exploration into how the Milky Way compares to the different [α/Fe]–[Fe/H] trends each simulation produces with different selection cuts; however, we find that our results are consistent under different cuts to capture disk stars.

This exploration shows that resolving distinct birth properties of chemically similar stars requires a small dispersion in age and R_{birth} trends in the [α/Fe]–[Fe/H] plane. Since the Milky Way has been shown to have age trends in the [α/Fe]–[Fe/H] abundance plane with small dispersion, we believe the conclusions we have drawn in this paper are relevant to the

Milky Way. We conclude that the six groups found in our previous work (Ratcliffe et al. 2020) using hierarchical clustering in the 19-dimensional APOGEE red clump sample are expected to have (R_{birth}, age) distributions that differ from each other and reflect a continuous evolution of the disk.

7.4. Limitations and Future Work

There are some limitations to our approach. In Section 5.2 we discuss the effect of subsampling data with observational errors. However, we do not take into account the complexity of survey-selection functions. Additionally, in Section 5 we explore how errors affect our clustering results by adding 0.02 dex or 0.05 dex error to every abundance. Realistically, though, some abundances are measured more accurately than others.

While in the previous section we argued that the main conclusions of this work could be extended to the Milky Way and that the groups found in Ratcliffe et al. (2020) are expected to occupy different regions in (age, R_{birth}), we do not extend the numerical relationship between birth properties and abundances (Section 6) to estimate age and R_{birth} for stars in the Milky Way. This extension would require the assumption that the regression coefficients of the Milky Way are exactly the same as those used for the simulations. However, abundance values between simulations and the Milky Way differ as well as spatial coverage and orbital and structural properties. Our analysis in Section 7.3 shows that a quadratic regression model will probably work for the Milky Way, but determining the coefficients that are appropriate for the Milky Way is beyond the scope of this work.

Section 7.2 discussed how density-based clustering failed when a structure vanished due to measurement error or in a high-dimensional abundance space. For future work, combining hierarchical clustering with an adaptive distance metric would be interesting to explore. Partnering an adaptive distance metric with hierarchical clustering would avoid the problems of estimating density and determining the best distance metric in a high-dimensional-abundance space, and could potentially provide even more striking results.

As simulation resolution continues to increase, in the future it would also be useful to explore if satellite debris could be picked up by abundance clustering and complement clustering analysis done in action space (such as in Wu et al. 2022). In the simulations we use in this study, only ≈20 stellar particles have $R_{birth} \geqslant 20$ kpc. With future data sets and simulations in mind, testing to see if accreted material differ in abundance space could be useful to determine accreted debris in the Milky Way.

8. Summary and Conclusions

Our main results are summarized below:

1. We find with just [Fe/H] and [α/Fe] alone we can trace separate R_{birth}–age groups, with the separation being more distinct when we include more abundances, as demonstrated by our 15-element simulation g8.26e11, where we find nearly completely separate groups for up to ≈10 groups (Figure 3). Considering current-day observational uncertainty and sampling constraints, higher-dimensional abundance data is necessary to trace birth properties from chemical abundance data. Groups from the subsampled Chem2d with 0.05 dex error had substantial overlap with other groups in the age–R_{birth} plane. On the other hand, groups found in the subsampled Chem15d with 0.05 dex error had almost no overlap in age and R_{birth} when finding six or fewer groups (Section 5; Figure 9).

2. The groups found in this paper presumably trace not only separate areas in the age–R_{birth} plane, but also link to different underlying physical properties. Stars separated by hierarchical clustering preserved a clear metallicity gradient as a function of age for a given R_{birth}, whereas groups defined by binning in the [α/Fe]–[Fe/H] plane lost the [Fe/H] gradient and increased their age dispersion (Figure 7).

3. Chemical clusters of high- and low-α stars separate differently in the age–R_{birth} plane. Groups defined with low-α stars separate both as a function of age and R_{birth}, showing low-α stars are born throughout the galaxy at different radii. High-α stars are older (>7 Gyr) and are born near the Galactic center, but separate as a function of age (see Figures 3 and 4).

4. Using a simple second-order polynomial regression, we quantify the relationship between observable abundance labels and birth property outputs (Section 6). We model age as a function of ([Fe/H], [X/Fe]), and can infer a star's age to a precision of ±0.52 Gyr for Chem2d and ±0.06 Gyr for Chem15d. We also model R_{birth} as a function of ([Fe/H], age), and infer it to a precision of ±1.24 kpc and ±1.17 kpc for Chem2d and Chem15d respectively.

5. The ability to reconstruct stellar groups born in different times and places from their abundances is determined by the formation history of the galaxy. When formation conditions lead to age and R_{birth} trends in the abundance plane with small dispersion, we find that there is a simple connection between clustered abundances and separate birth times and places. Under clumpy star formation, however, the simple relationship vanishes (Section 7.3).

6. Our comparison of three simulations implies that the low dispersion of age across the [α/Fe]–[Fe/H] plane of the Milky Way indicates that the Milky Way's star formation history is sufficiently quiet and that clustering in abundance will correspond to birth associations in time and location (Figure 15).

We seek to examine how abundance structure links to birth properties. We find that there is a simple relationship between age and chemical abundances, which agrees with previous work (e.g., Ness et al. 2019). R_{birth} cannot be tested as we can do for age; we never have direct access to this quantity in observations. From our regression, however, we see age and ([Fe/H], [X/Fe]) link us to R_{birth} in the simulations. Indeed, this analytical formalism has been adopted in models of radial migration (e.g., Frankel et al. 2018; Minchev et al. 2019). We examine the R_{birth}–age distribution further using the idea of abundance clustering, in which we seek to see if it links to underlying physical processes.

This work highlights how we might use clustering of high-dimensional-abundance measurements in large surveys to infer groups of different birth place and time, and the impact of measurement uncertainty in working with the observational data.

Appendix
Additional Figures

Here we include additional figures that help readers interpret results. Figures 16 and 17 are abundance–age plots colored by R_{birth} for both Chem2d and Chem15d. These plots are similar to Figures 10 and 11, however the coloring and x-axis are switched. Figure 18 allows the reader to make a direct comparison between the groups found in Chem2d using hierarchical clustering and EnLink.
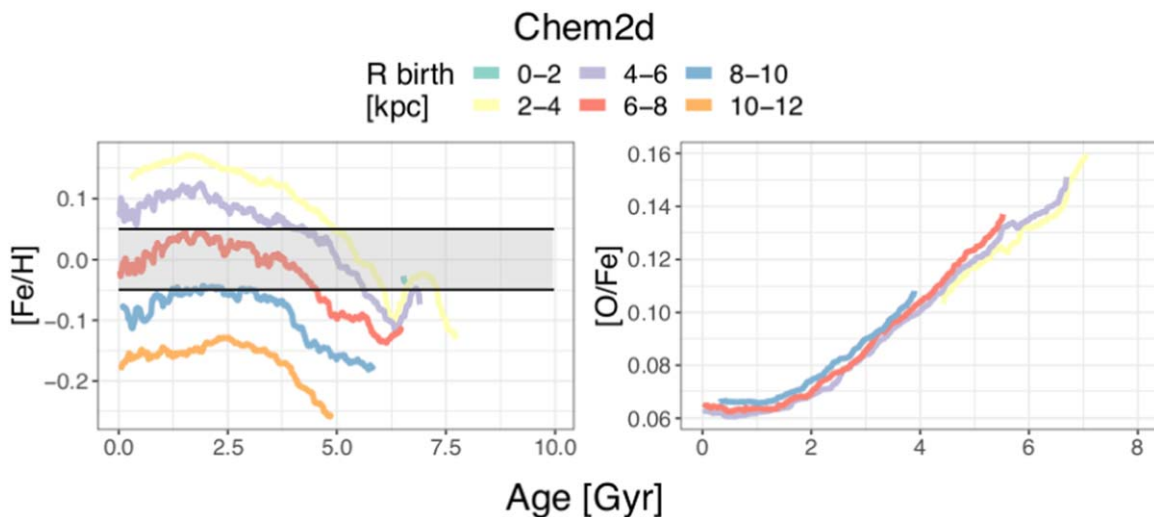
Figure 16. Left: the [Fe/H]–age plane colored by R_{birth} for Chem2d. The black lines and gray area mark off the solar metallicity stars, which we consider to be ±0.05 dex in [Fe/H]. Right: the running mean of [O/Fe] of the solar metallicity stars across age colored by R_{birth} selected from within the horizontal lines at left. For a given bin of metallicity, stars clearly have a polynomial trend in [X/Fe]–age.
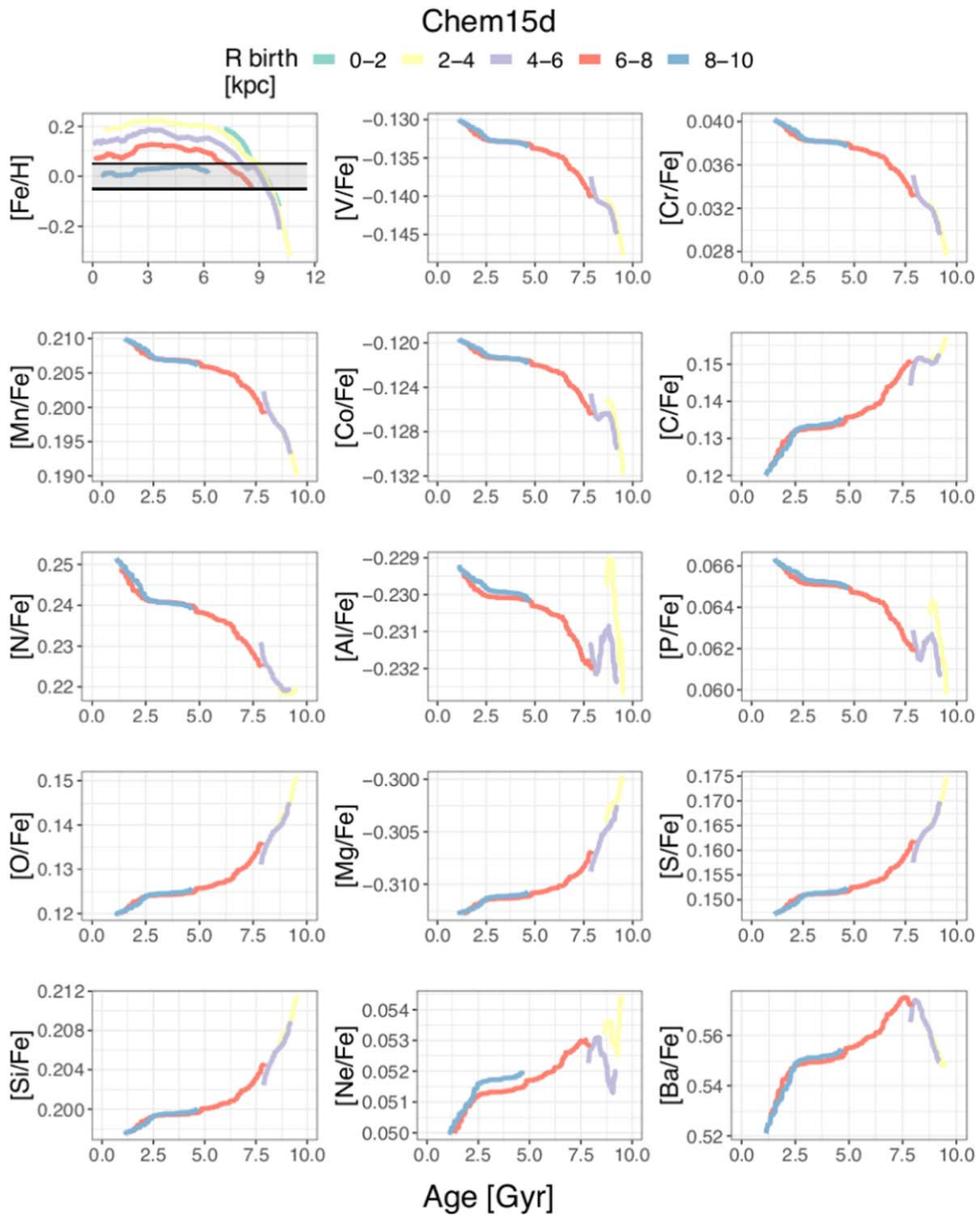
Figure 17. Top-left: the [Fe/H]–age plane colored by R_{birth} for Chem15d. The black lines and gray area mark off the solar metallicity stars, which we consider to be ±0.05 dex in [Fe/H]. All other plots show the running mean of [X/Fe] of the solar metallicity stars accross age colored by R_{birth}. Similar to Chem2d shown in Figure 16, solar metallicity stars of different ages separate into different polynomial curves.
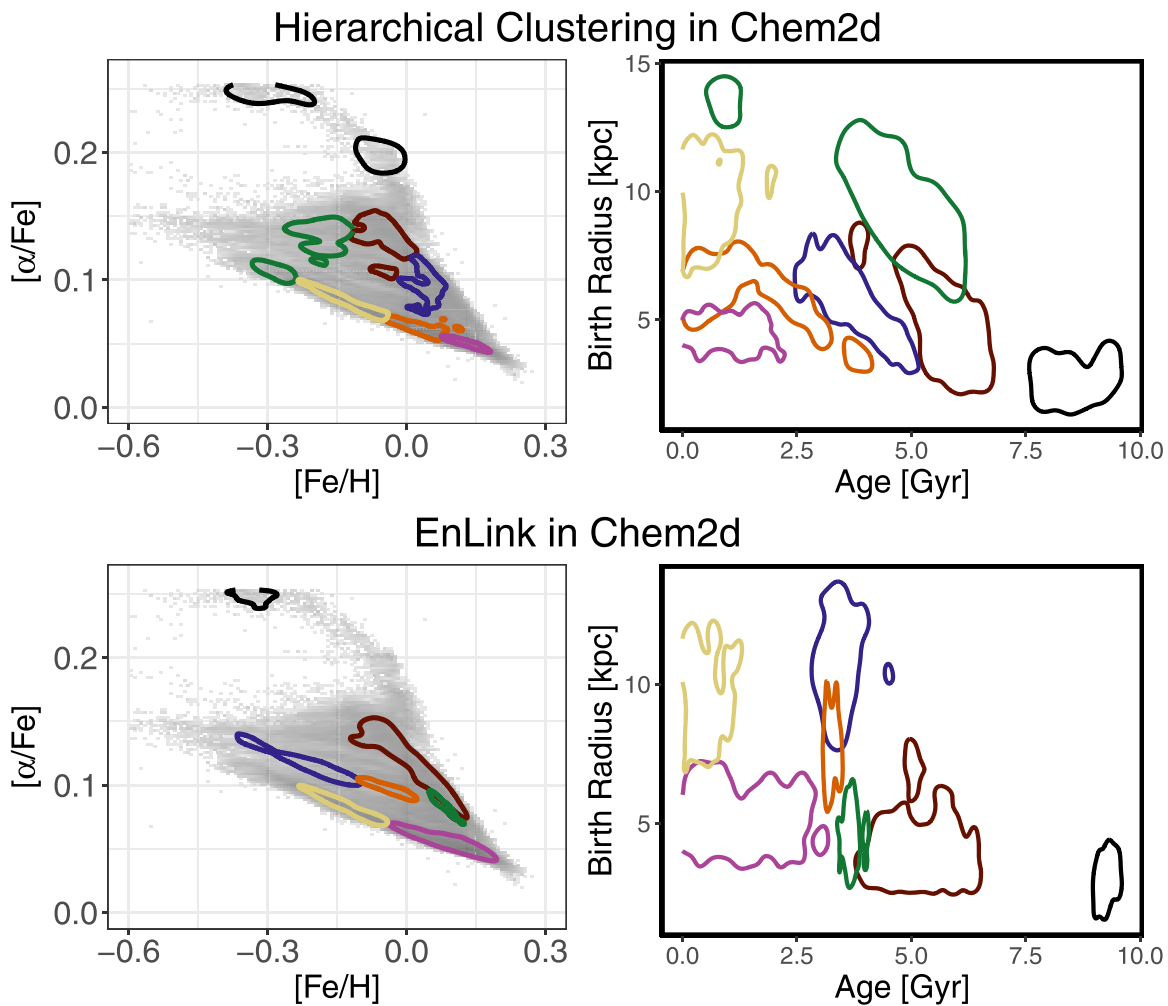
Figure 18. Seven groups found using (top) hierarchical clustering and (bottom) EnLink in the two-dimensional abundance space of Chem2d projected into the (left) [α/Fe]–[Fe/H] and (right) R_{birth}–age planes.

ORCID iDs

Bridget L. Ratcliffe ⓘ https://orcid.org/0000-0003-1124-7378
Melissa K. Ness ⓘ https://orcid.org/0000-0001-5082-6693
Tobias Buck ⓘ https://orcid.org/0000-0003-2027-399X
Kathryn V. Johnston ⓘ https://orcid.org/0000-0001-6244-6727
Bodhisattva Sen ⓘ https://orcid.org/0000-0002-4667-3150
Leandro Beraldo e Silva ⓘ https://orcid.org/0000-0002-0740-1507
Victor P. Debattista ⓘ https://orcid.org/0000-0001-7902-0116

References

Ahumada, R., Prieto, C. A., Almeida, A., et al. 2020, ApJS, 249, 3
Beraldo e Silva, L., Debattista, V. P., Nidever, D., Amarante, J. A. S., & Garver, B. 2021, MNRAS, 502, 260
Bird, J. C., Loebman, S. R., Weinberg, D. H., et al. 2021, MNRAS, 503, 1815
Bland-Hawthorn, J., Krumholz, M. R., & Freeman, K. 2010, ApJ, 713, 166
Bovy, J., Nidever, D. L., Rix, H.-W., et al. 2014, ApJ, 790, 127
Bovy, J., Rix, H.-W., Liu, C., et al. 2012, ApJ, 753, 148
Buck, T. 2020, MNRAS, 491, 5435
Buck, T., Macciò, A. V., Dutton, A. A., Obreja, A., & Frings, J. 2019a, MNRAS, 483, 1314
Buck, T., Macciò, A. V., Obreja, A., et al. 2017, MNRAS, 468, 3628
Buck, T., Ness, M., Obreja, A., Macciò, A. V., & Dutton, A. A. 2019b, ApJ, 874, 67
Buck, T., Obreja, A., Macciò, A. V., et al. 2020, MNRAS, 491, 3461
Buck, T., Rybizki, J., Buder, S., et al. 2021, MNRAS, 508, 3365
Buder, S., Asplund, M., Duong, L., et al. 2018, MNRAS, 478, 4513
Chiappini, C., Matteucci, F., & Gratton, R. 1997, ApJ, 477, 765

Clarke, A. J., Debattista, V. P., Nidever, D. L., et al. 2019, MNRAS, 484, 3476
Feltzing, S., Bowers, J. B., & Agertz, O. 2020, MNRAS, 493, 1419
Fiteni, K., Caruana, J., Amarante, J. A. S., Debattista, V. P., & Beraldo e Silva, L. 2021, MNRAS, 503, 1418
Frankel, N., Rix, H.-W., Ting, Y.-S., Ness, M. K., & Hogg, D. W. 2018, ApJ, 865, 96
Freeman, K., & Bland-Hawthorn, J. 2002, ARA&A, 40, 487
Friedli, D., Benz, W., & Kennicutt, R. 1994, ApJL, 430, L105
Garrison-Kimmel, S., Hopkins, P. F., Wetzel, A., et al. 2018, MNRAS, 481, 4133
Gilmore, G., Randich, S., Asplund, M., et al. 2012, Msngr, 147, 25
Grand, R. J. J., Gómez, F. A., Marinacci, F., et al. 2017, MNRAS, 467, 179
Grand, R. J. J., Springel, V., Gómez, F. A., et al. 2016, MNRAS, 459, 199
Hartigan, J. A., & Wong, M. A. 1979, Journal of the Royal Statistical Society. Series C (Applied Statistics), 28, 100
Hayden, M. R., Recio-Blanco, A., de Laverny, P., et al. 2018, A&A, 609, A79
Haywood, M., Di Matteo, P., Lehnert, M. D., Katz, D., & Gómez, A. 2013, A&A, 560, A109
Hemler, Z. S., Torrey, P., Qi, J., et al. 2021, MNRAS, 506, 3024
Hogg, D. W., Casey, A. R., Ness, M., et al. 2016, ApJ, 833, 262
Johnson, J. W., Weinberg, D. H., Vincenzo, F., et al. 2021, MNRAS, 508, 4484
Jönsson, H., Holtzman, J. A., Allende Prieto, C., et al. 2020, AJ, 160, 120
Kaufman, L., & Rousseeuw, P. J. 2009, Finding Groups in Data: an Introduction to Cluster Analysis (New York: Wiley)
Lindegren, L., & Feltzing, S. 2013, A&A, 553, A94
Loebman, S. R., Roškar, R., Debattista, V. P., et al. 2011, ApJ, 737, 8
Lu, Y., Ness, M., Buck, T., & Zinn, J. 2021, Universal properties of the high- and low-α disk: small intrinsic abundance scatter and migrating stars, arXiv:2102.12003
Martell, S. L., Shetrone, M. D., Lucatello, S., et al. 2016, ApJ, 825, 146

Matteucci, F., & Francois, P. 1989, MNRAS, 239, 885
Minchev, I., Anders, F., Recio-Blanco, A., et al. 2018, MNRAS, 481, 1645
Minchev, I., & Famaey, B. 2010, ApJ, 722, 112
Minchev, I., Famaey, B., Quillen, A. C., & Dehnen, W. 2012a, European Physical Journal Web of Conferences, 19, 07002
Minchev, I., Famaey, B., Quillen, A. C., et al. 2012b, A&A, 548, A127
Minchev, I., Martig, M., Streich, D., et al. 2015, ApJL, 804, L9
Minchev, I., Matijevic, G., Hogg, D. W., et al. 2019, MNRAS, 487, 3946
Muñoz-Mateos, J. C., Gil de Paz, A., Boissier, S., et al. 2007, ApJ, 658, 1006
Murtagh, F., & Legendre, P. 2014, Journal of Classification, 31, 274
Ness, M. K., Johnston, K. V., Blancato, K., et al. 2019, ApJ, 883, 177
Obreja, A., Macciò, A. V., Moster, B., et al. 2018a, MNRAS, 477, 4915
Obreja, A., Macciò, A. V., Moster, B., et al. 2018b, gsf: galactic structure finder, Astrophysics Source Code Library, ascl:1806.008
Pfitzner, D. W., Salmon, J. K., & Sterling, T. 1997, Data Mining and Knowledge Discovery, 1, 419
Pontzen, A., Roškar, R., Stinson, G. S., et al. 2013, pynbody: Astrophysics Simulation Analysis for Python, Astrophysics Source Code Library, ascl:1305.002

Ratcliffe, B. L., Ness, M. K., Johnston, K. V., & Sen, B. 2020, ApJ, 900, 165
Roškar, R., Debattista, V. P., Quinn, T. R., Stinson, G. S., & Wadsley, J. 2008, ApJL, 684, L79
Roškar, R., Debattista, V. P., Quinn, T. R., & Wadsley, J. 2012, MNRAS, 426, 2089
Roškar, R., Debattista, V. P., Stinson, G. S., et al. 2008, ApJL, 675, L65
Sanderson, R. E., Wetzel, A., Loebman, S., et al. 2020, ApJS, 246, 6
Schönrich, R., & Binney, J. 2009a, MNRAS, 396, 203
Schönrich, R., & Binney, J. 2009b, MNRAS, 399, 1145
Sellwood, J. A., & Binney, J. J. 2002, MNRAS, 336, 785
Sharma, S., & Johnston, K. V. 2009, ApJ, 703, 1061
Sharma, S., & Steinmetz, M. 2006, MNRAS, 373, 1293
Stinson, G. S., Brook, C., Macciò, A. V., et al. 2013, MNRAS, 428, 129
Ting, Y.-S., Conroy, C., & Goodman, A. 2015, ApJ, 807, 104
Wadsley, J. W., Keller, B. W., & Quinn, T. R. 2017, MNRAS, 471, 2357
Wang, L., Dutton, A. A., Stinson, G. S., et al. 2015, MNRAS, 454, 83
Ward, J. H., Jr. 1963, Journal of the American Statistical Association, 58, 236
Wu, Y., Valluri, M., Panithanpaisal, N., et al. 2022, MNRAS, 509, 5882