



In the Galactic Disk, Stellar [Fe/H] and Age Predict Orbits and Precise [X/Fe]

M. K. Ness^{1,2} , K. V. Johnston¹, K. Blancato¹, H-W. Rix³ , A. Beane⁴ , J. C. Bird⁵, and K. Hawkins⁶

¹Department of Astronomy, Columbia University, Pupin Physics Laboratories, New York, NY 10027, USA; melissa.ness@columbia.edu

²Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA

³Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

⁴Department of Physics & Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA

⁵Department of Physics and Astronomy, Vanderbilt University, 6301 Stevenson Center, Nashville, TN 37235, USA

⁶Department of Astronomy, The University of Texas at Austin, 2515 Speedway Boulevard, Austin, TX 78712, USA

Received 2019 July 23; revised 2019 August 23; accepted 2019 August 23; published 2019 October 3

Abstract

We explore the structure of the element abundance–age–orbit distribution of the stars in the Milky Way’s low- α disk, by (re-)deriving precise [Fe/H], [X/Fe], and ages, along with orbits, for red clump stars from the APOGEE survey. There has been a long-standing theoretical expectation and observational evidence that metallicity ([Fe/H]) and age are informative about a star’s orbit, e.g., about its angular momentum and the corresponding mean Galactocentric distance or its vertical motion. Indeed, our analysis of the APOGEE data confirms that [Fe/H] or age alone can predict the stars’ orbits far less well than a combination of the two. Remarkably, we find and show explicitly that, for known [Fe/H] and age, the other abundances [X/Fe] of Galactic disk stars can be predicted well (on average to 0.02 dex) across a wide range of Galactocentric radii, and therefore provide little additional information, e.g., for predicting their orbit. While the age–abundance space for metal-poor stars and potentially for stars near the Galactic center is rich or complex, for the bulk of the Galaxy’s low- α disk it is simple: [Fe/H] and age contain most information, unless [X/Fe] can be measured to 0.02 or better. Consequently, we do not have the precision with current (and likely near-future) data to assign stars to their individual (coeval) birth clusters, from which the disk is presumably formed. We can, however, place strong constraints on future models of Galactic evolution, chemical enrichment, and mixing.

Unified Astronomy Thesaurus concepts: Milky Way disk (1050); Milky Way evolution (1052); Milky Way formation (1053); Milky Way Galaxy (1054)

1. Introduction

In the pursuit of using stars to understand the processes that have formed and evolved the Milky Way disk, many millions of low- and medium-resolution spectra have been obtained (e.g., Steinmetz et al. 2006; Yanny et al. 2009; Gilmore et al. 2012; Newberg et al. 2012; De Silva et al. 2015; Majewski et al. 2017). The acquisition rate of stellar spectral data-intake will increase into the next decade (e.g., Newberg et al. 2012; Cirasuolo et al. 2014; De Silva et al. 2015; Bonifacio et al. 2016; Kollmeier et al. 2017; Tamura et al. 2018; de Jong et al. 2019). Stellar abundances are expected to encode the temporal enrichment of the disk. An ensemble of abundance measurements, for vast stellar samples, over large spatial extents, is therefore critical to constrain the diversity and characteristics of nucleosynthetic channels. From a theoretical standpoint, it is also expected that, given precise enough chemical abundance measurements and sufficient data, we might also be able to reconstruct the original birth sites of stars in the Galactic disk, that is, to assign them to the individual clusters from which they formed (Bland-Hawthorn et al. 2010; Armillotta et al. 2018a; Krumholz et al. 2019). However, the empirical information content and diagnostic power of the multitude of element abundances being assembled, in terms of understanding the disk’s formation, has to date been poorly constrained (see, however, Ting et al. 2012).

At this point we have established, from the large number of large Milky Way stellar surveys and subsequent studies, a global set of descriptive properties of the Galactic disk. These pertain to gradients in metallicity ([Fe/H]), age and α -element enhancements. Stars in the disk show a vertical abundance gradient (Schlesinger et al. 2014) of about -0.24 dex kpc^{-1} (e.g., Duong et al. 2018)

which increases to -0.45 dex kpc^{-1} in the bulge (e.g., Ness et al. 2013). Radially, the young stars confined to the plane of the Milky Way disk show a weak gradient in their [Fe/H] of -0.075 dex kpc^{-1} (e.g., Genovali et al. 2014; Frankel et al. 2018). This gradient, and the small dispersion around it among young stars, presumably reflects an initial birth relation between the level of chemical enrichment of the star-forming gas and its radii. Such a correlation would be weakened by any subsequent evolutionary processes that have moved stars from their birth places over time (e.g., Sellwood & Binney 2002; Roškar et al. 2008; Minchev et al. 2011). This is directly seen in global models of the age–[Fe/H] relation of the Galactic disk (Frankel et al. 2018), as stars in the Milky Way suffer significant dynamical “memory loss” over time, through radial migration. It may be possible, in part, to reconstruct this memory using precise [Fe/H] and age estimates, as shown by Minchev et al. (2018) using the high-resolution HARPS GTO data. That said, the weak trends seen between orbits and [Fe/H] at all ages indicate there is some level of memory retention linking orbits to chemical compositions (Beane et al. 2018).

In addition to the metallicity gradient seen in the disk, bulk element enhancement trends, like $[\alpha/\text{Fe}]$, are dramatically informative as to location in the disk. The mapping of $[\alpha/\text{Fe}]$ across the disk, from Galactic radii 4–16 kpc, shows that stars with high α -enhancement are concentrated in the inner Galaxy and those with low α -enhancement are concentrated in the outer Galaxy (e.g., Bensby et al. 2012). Both the high and low α -sequences appear near the Sun. The fraction of α -enhanced stars also increases at larger heights from the plane at a given radius (Bovy et al. 2012; Nidever et al. 2014; Hayden et al. 2015). Broadly, the α -enhancement trends seen across the Galaxy

imply a faster rate of star formation in the inner compared to the outer Galaxy. These trends reflect the convolution of many physical processes, including the star formation history, disk growth and heating, and migratory processes during the disk’s evolution. The two α -enhancement sequences, as described visually in the $[\text{Fe}/\text{H}]-[\alpha/\text{Fe}]$ plane, have been demonstrated to be empirically dynamically different (e.g., Gandhi & Ness 2019; Mackereth et al. 2019b) and explained theoretically as having different Galactic formation origins (e.g., Clarke et al. 2019; Mackereth et al. 2019a).

The APOGEE ($\approx 500,000$ stars) survey (Zasowski et al. 2013; Majewski et al. 2017) and its successor, the upcoming (>5 million star) Sloan V Milky Way Mapper survey (Kollmeier et al. 2017) are benchmark Galactic archaeology programs. These are providing the data to trace the chemo-dynamical structure of the Galaxy over a large spatial extent, by observing hundreds of thousands and, from 2020 with Milky Way Mapper, millions of bright giant tracers. The GALAH survey was engineered to be the ultimate “chemical tagging” experiment, observing an ensemble of elements from a larger multitude of nucleosynthetic families, primarily for nearby main-sequence stars, in large enough numbers to reconstruct individual cluster birth sites using abundances measured for individual stars. Critically, from current (and for future) surveys, we now have, from data-driven modeling using benchmark stars with precision asteroseismic ages, access not only to abundances but to ages for large numbers of stars across the Milky Way disk (e.g., Martig et al. 2016; Ness et al. 2016; Ho et al. 2017; Bovy et al. 2019; Leung & Bovy 2019). We can see from the APOGEE survey, for example, evidence for the inside-out formation of the Milky Way from these ages (Ness 2018). Age is the fundamental variable of temporal evolution, and the key tracer around which all other variables can be pivoted in the pursuit of using the ensemble of stellar measurements to understand our Galaxy.

In this paper, we seek to answer the following fundamental questions about the information content in the data: (i) what does each abundance tell us about stellar age, (ii) is there additional information content in each abundance at a given age, and (iii) how do the present-day orbits of stars of different mean abundances vary as a function of age at fixed $[\text{Fe}/\text{H}]$?

We use orbit actions to quantify the stellar orbital properties (Binney & Tremaine 2008). Using *Gaia* parallaxes and proper motions, we can access these actions. The three actions of J_R , J_z , and J_ϕ are each expressions of the radial excursion around the guiding radius, the vertical extent of the orbit above its orbital plane, and its guiding, or average, radius respectively.

We use approximately 20,000 red clump stars identified in the catalog of Ting et al. (2018), with abundances that have been corrected for known systematics, from the APOGEE survey. From this set, we isolate the $\approx 15,000$ stars in the low- α sequence. We examine only the low- α stars as we wish to exclude stars in the high- α sequence with a probable separate origin of formation (see Clarke et al. 2019; Gandhi & Ness 2019; Mackereth et al. 2019a). The APOGEE red clump sample extends across a wide radial extent (4–16 kpc) and from these spectra we measure a multitude of elements. From this set of stars, we derive a set of precision chemical abundances, ages precise to $\lesssim 1.6$ Gyr, and orbital actions precise to $<10\%$. We determine 19 abundances using a data-driven modeling training on ASPCAP measurements (García Pérez et al. 2015; Holtzman et al. 2015, 2018) and correcting for systematics introduced by the varying line-spread function across the spectrograph (Ness 2018).

We have a particularly valuable set of ≈ 1100 local red clump stars in this sample, from the APOKASC catalog (Pinsonneault et al. 2018). These stars have masses and consequently ages determined from their internal oscillations via asteroseismology, from their Kepler power spectra. We use a subset of this sample to investigate the relationship between element abundances $[\text{X}/\text{Fe}]$ and age, and compare our results to very high-precision analyses from $R = 115,000$ spectra of main-sequence and turn-off solar-twin stars (Bedell et al. 2018). We also use the APOKASC stars as a reference set to determine ages for the full red clump low- α sample, using data-driven modeling.

To examine the intrinsic information content in each abundance as a function of age, we measure the intrinsic abundance dispersion around the age–abundance relationships for each element, that is, any additional dispersion in the abundances as a function of age that cannot be explained by the age and abundance measurement errors. Such intrinsic dispersions directly quantify to what level each element may differentiate sites of star formation at a given time, which has been theoretically expected but not empirically demonstrated (e.g., Armillotta et al. 2018b; Krumholz et al. 2019). Furthermore, the intrinsic dispersion indicates the precision at which each element must be measured in order to render that element informative beyond being a label of age for any individual star. We also examine the trends between stellar abundances and orbital action labels. We consider populations of stars, combined using their abundance similarity, to demonstrate that stars separate out dynamically and temporally at fixed $[\text{Fe}/\text{H}]$, and the separation is dissimilar between groups of old and young stars. This population analysis is a practical demonstration that abundances indicate age and orbital properties of the disk architecture.

Finally, we discuss the implications of our findings, which demonstrate the power of abundances to indicate ages and that the majority of information about a star, e.g., linking to orbits, is captured in its age and $[\text{Fe}/\text{H}]$ for the low- α disk. We conclude by citing the prospects for the coming era of multi-million star surveys, where combining measurements of stars will tap into the orbital sub-structure of the disk, opening up opportunities in Galactic archaeology across a breadth of spatial scales.

2. Data

Our data is sourced from the APOGEE DR14 (Majewski et al. 2017) catalog, the second APOKASC data release of stellar ages (Pinsonneault et al. 2018) and the *Gaia* DR2 (Gaia Collaboration 2018) data release. For the APOGEE stars, we use the $\approx 20,000$ spectroscopically confirmed red clump stars (Hawkins et al. 2018). This sample represents the most pristine red clump selection available with an expected contamination rate of 3% (Ting et al. 2018). These stars have precise distances from their fixed red clump absolute magnitudes (Hawkins et al. 2017). Of these stars, we select $\approx 15,000$ low- α sequence stars for our analysis, based on their location in the $[\text{Fe}/\text{H}]-[\alpha/\text{Fe}]$ plane. To determine abundances for our red clump stars, we use *The Cannon* on the DR14 spectra, with the ASPCAP abundances and stellar parameters as the training labels, following the prescription and training set of Ness (2018). The APOGEE survey delivers a multitude of element abundances: we learn labels of $[\text{Fe}/\text{H}]$ plus 18 additional abundance measurements from the APOGEE red clump infrared spectra. These elements are C, N (light proton), Na, Al, S, K (light odd- z), Mg, Si, Ca, Ti, O (α), Fe, V, Mn, Ni, P, Cr, Co (iron-peak), and Rb (s-process). The elements Na and Rb have high uncertainties and are not robustly

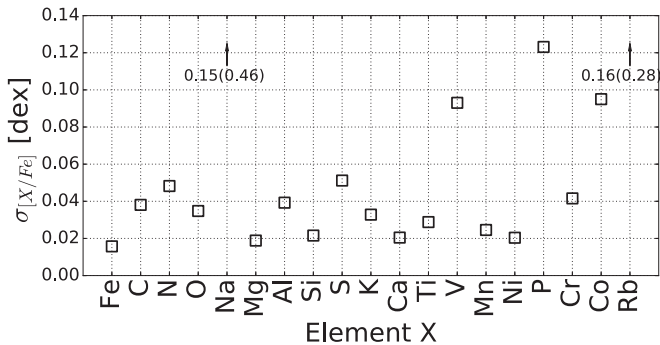


Figure 1. Mean uncertainty for each abundance measurement that we derive from APOGEE spectra using *The Cannon* (Ness et al. 2015). These errors are the rms sum of the cross-validation error of a subset of ≈ 2200 of the stars (that span the narrow red clump parameter space) from the ≈ 4500 star training set of *The Cannon* (Ness et al. 2018), and the formal errors, as reported from the optimization in *The Cannon*. Note the formal errors are typically about 10% of the magnitude of the cross-validation errors. The elements Na and Rb fail for a subset of stars in cross-validation and the bracketed errors indicate what the error is including these failed stars (if we exclude the failures at cross-validation the errors are 0.15 and 0.16 dex respectively, as indicated). Na and Rb are largely excluded from our analysis as a consequence of their cross-validation failure for a fraction of the stars.

recovered in our data-driven inference and we end up excluding these elements from our analysis.

We take care with our abundance derivations to remove the remaining systematic abundance trends that are apparent in the abundance correlations seen with fiber number (as described in Ness 2018). These trends are significant and, if not accounted for, may propagate into problematic systematic effects when abundances are used in concert for analyses. We determine the error on our abundances using a root mean square (rms) sum of the cross-validation error from *The Cannon*’s inference and the formal uncertainty of *The Cannon*’s optimizer. For the cross-validation errors, we generate 20 models, excluding for each model 5% of the data. We subsequently test how well the abundances are recovered for the excluded stars, determining the rms difference of the true (ASPCAP) minus the inferred values from *The Cannon*. Our analysis uses only the red clump stars, so we restrict our cross-validation test to the ≈ 2200 training stars within the red clump parameter space of $T_{\text{eff}} = 4820 \pm 200$ K, $\log g = 2.45 \pm 0.5$ and $[\text{Fe}/\text{H}] = 0 \pm 0.5$ dex. This additional restriction to the overall parameter space of the training set leads to only a marginal or negligible improvement in the rms uncertainty for the element abundances but a marked improvement in the $[\text{Fe}/\text{H}]$ precision, which decreases from 0.022 dex to 0.015 dex. The mean uncertainty of these abundance measurements overall is <0.03 dex (see Figure 1). In our cross-validation analysis, we find that we fail to recover the Na and Rb abundances for a subset of the stars. We therefore exclude these two elements from the majority of our analysis.

We determine spectroscopic ages for the red clump stars following Ness et al. (2016). However, instead of modeling the spectral features to determine age labels, we use the precision abundances directly as our features to learn the relationship between abundances (features) and age (labels). Our training set in this case is comprised of the red clump stars in the second data release of the APOKASC catalog of Pinsonneault et al. (2018), for which there are ≈ 1100 red clump stars in total with asteroseismic ages. We implement a number of quality cuts

described in Section 3.2. This delivers ages for the full red clump sample that are precise to $\lesssim 1.6$ Gyr.

We calculate actions for the APOGEE stars using their radial velocities (from APOGEE), proper motions from *Gaia* DR2, and distances assuming their red clump membership (Ting et al. 2018), using the Galpy code (Bovy 2016). We use solar reference values of 220 km s^{-1} and $R_{\text{GAL}} = 8 \text{ kpc}$ and adopt the potential of MWPotential14. The radial velocity precision for APOGEE is $<300 \text{ m s}^{-1}$. Under the epicyclic approximation, the total error on the actions can be approximated as the rms sum of the fractional error of the parameters’ input to the action calculation and dominated by the distance error, which is typically 5% (errors on our proper motion are typically $<1\%$). This gives us a high-fidelity sample of stars with precision measurements on actions of typically $<8\%$, that cover a large radial extent across the disk, with a multitude of element families with precision abundance measurements of typically <0.03 dex.

In this work we analyze the low- α stars, our selection of which is shown in Figure 2. The left-hand panel shows the asteroseismic sample of stars colored by their asteroseismic stellar ages calculated from their power spectra (Pinsonneault et al. 2018) and the right-hand panel shows the full set of 20,000 APOGEE red clump stars from which we derive ages for the $\approx 15,000$ stars in the low- α sequence across the parameter range of the asteroseismic training set, as described in Section 3.2.

We proceed by first examining the small sample of stars with precise and accurate ages from asteroseismic measurements in the APOKASC catalog. We use ≈ 70 of these stars at solar metallicity, $[\text{Fe}/\text{H}] = 0 \pm 0.05$ dex, to examine the abundance–age trends and the mean intrinsic dispersion of each element. We then use a larger sample of ≈ 600 of the asteroseismic red clump stars across a wide range of $[\text{Fe}/\text{H}]$ to train a data-driven model to label $\approx 15,000$ APOGEE red clump stars with ages from their abundances. The $\approx 15,000$ stars correspond to those within the stellar parameter space of the training set with reference ages that we use to build our model. We use this set of 600 stars with asteroseismic ages to measure the intrinsic dispersion of each element in further detail, modeling age and $[\text{Fe}/\text{H}]$ together. We finally examine the abundance–orbital properties of stars using our sample of $\approx 15,000$ stars across the disk for which we have derived ages.

3. Abundance–Age Correlations in the Low- α Disk

3.1. The Local APOKASC Asteroseismic Sample of Stars at Solar $[\text{Fe}/\text{H}]$

The ≈ 600 stars that meet our analysis criteria span a metallicity range of $-0.5 < [\text{Fe}/\text{H}] < 0.45$ dex (with a median $[\text{Fe}/\text{H}] \approx 0$ dex) and an age range of $0.3 < \tau_{\text{age}} < 11.7$ Gyr (with a median age ≈ 3.5 Gyr). We wish to examine the age–abundance correlations of these stars. The left-hand panel of Figure 2 shows clearly that the mean abundance, in this case $[\text{Mg}/\text{Fe}]$, changes as a function of overall metallicity, $[\text{Fe}/\text{H}]$. Therefore, to proceed with our investigation, we condition on a single $[\text{Fe}/\text{H}]$ value. From the set of ≈ 600 stars with asteroseismic ages that meet our quality cuts, we select the ≈ 70 solar metallicity stars, with $[\text{Fe}/\text{H}] = 0 \pm 0.05$ dex. The individual abundances of these stars versus $[\text{Fe}/\text{H}]$, are shown in Figure 3. The best fitting (second-order) polynomial line to each age–abundance trend is shown as the thick black line in

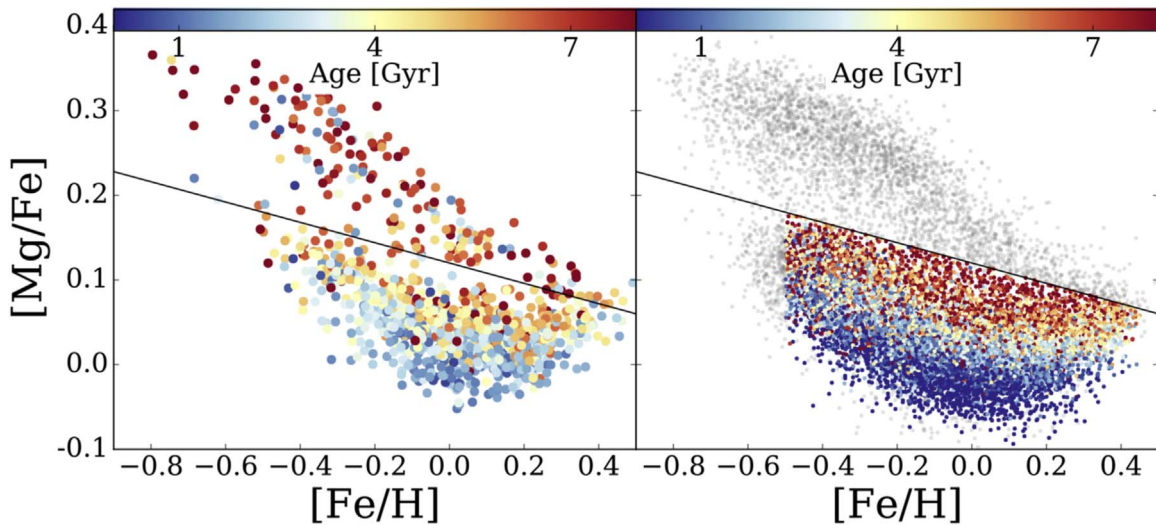


Figure 2. Stellar ages shown in the $[\text{Fe}/\text{H}]$ – $[\text{Mg}/\text{Fe}]$ plane. Left: the ≈ 1100 APOKASC red clump sample of stars with asteroseismic ages (Pinsonneault et al. 2018). Right: the larger sample of APOGEE red clump stars, for which we derive ages, shown using a model built from the low- α asteroseismic stars. The age scale is provided at the top of each panel. The ≈ 800 red clump stars in the APOKASC sample that meet our quality cuts at left are described in Section 3.1 (with 600 stars in the low- α sample). The full 20,000 red clump sample of stars in APOGEE shown in the panel at right have membership determined from Ting et al. (2018). The subset of $\approx 15,000$ stars for which we determine ages in the right-hand panel is described in Section 3.2. The age range for the right-hand panel is broader than for the left as we place no prior on our data-driven age modeling and our full set of red clump derived ages span a larger range than the APOKASC reference sample that we use to build our model (including to negative values). For our analysis, we consider stars in the low- α sequence only, selected as below the (ad hoc) line drawn. This line is intended to exclude stars with a probable separate, different formation history to those of the low- α stars that describe the chemical abundance space of the subsequent inside-out formed Milky Way disk.

each of the sub-panels. This line is fit using a least-squares minimization between the line and individual points. For reference, the slope of the best-fitting straight line is given in the top of each sub-panel, to indicate the relative magnitude of the gradients for each element (as all y-axes are scaled on a per-element basis). The two shaded gray lines around the line of best fit represent the 1σ dispersion of the data (taking 10 stars per bin) around this line.

Figure 3 shows that the elements C and N, which are associated with mass-dependent evolutionary surface abundance changes in red giants (Masseron & Gilmore 2015; Martig et al. 2016), have the steepest age–abundance gradients. The elements O and Mg, which more likely reflect temporal chemical enrichment of the disk (e.g., Bensby et al. 2017), are the second most correlated elements with age. Notably, the α -elements show a significant variation in their slopes. This is presumably indicative of the different chemical enrichment pathways and subsequent mass-dependent yield variations, even within a single nucleosynthetic family of elements (e.g., Blancato et al. 2019; Weinberg et al. 2019).

For these sets of stars shown in Figure 3, we calculate the intrinsic dispersion of each element around the best-fitting line describing the age–abundance relation. This intrinsic dispersion is the amount of scatter that is not explained by the measurement error due to the abundance and age uncertainty. The total dispersion around the abundance–age relation is simply the rms sum of the dispersion expected given the measurement error and the underlying intrinsic dispersion for that element around a given age:

$$\sigma_{\text{TOTAL}}^2 = \sigma_{\text{MEASUREMENT ERROR}}^2 + \sigma_{\text{INTRINSIC}}^2.$$

We therefore calculate the intrinsic dispersion quantity $\sigma_{\text{INTRINSIC}}$ at a given age, for each element, by taking the rms difference of the mean rms dispersion around the line of best fit to the age–abundance trend and the mean measurement error. The latter, which contributes to the scatter around the

best-fitting age–abundance line, is calculated as the rms combination of the mean measurement errors on the abundances and the propagated age error that contributes to additional scatter around the best-fit line to the age–abundance relation. This (mean) quantity is shown in the top right-hand corner of each sub-panel in Figure 3. The contribution of the age error to the uncertainty is determined by drawing a new age for every star from its Gaussian distribution, given its age and age error, and calculating the subsequent total rms dispersion away from the already determined second-order polynomial best fit through the data, for this new distribution of ages. This is repeated 100 times. The mean of the calculated dispersion around the best-fit line from these draws is taken as the contribution of the age uncertainty to the dispersion accounted for by the age measurement error around the line of best fit. This value is typically <0.01 dex and decreases with a decreasing age–abundance correlation. The intrinsic dispersion is then determined as the rms difference of the 1σ dispersion measured around the line of best fit to the age–abundance relation (shown in the bottom right-hand corner of each sub-panel) and the 1σ dispersion of the mean joint measurement errors from the age and abundance uncertainties.

Due to the small sample size of the solar metallicity stars, a single element abundance outlier can be sufficient to impact the measured intrinsic dispersion. We found that a single star with an anomalous abundance significantly increased the intrinsic dispersion for two elements S and Co, and our results shown in Figure 4 are for this single star removed, for each element. We allowed for a 3.5σ clipping for all elements to test the sensitivity of the intrinsic dispersion of all elements to any outliers, and find that this removed only a single star for S and Co (and the intrinsic dispersion results are robust for the other elements upon removal of the largest outlier for each). Therefore, the intrinsic dispersion is reported using 69 stars for Co and S and the full set of 70 stars for the remaining elements.

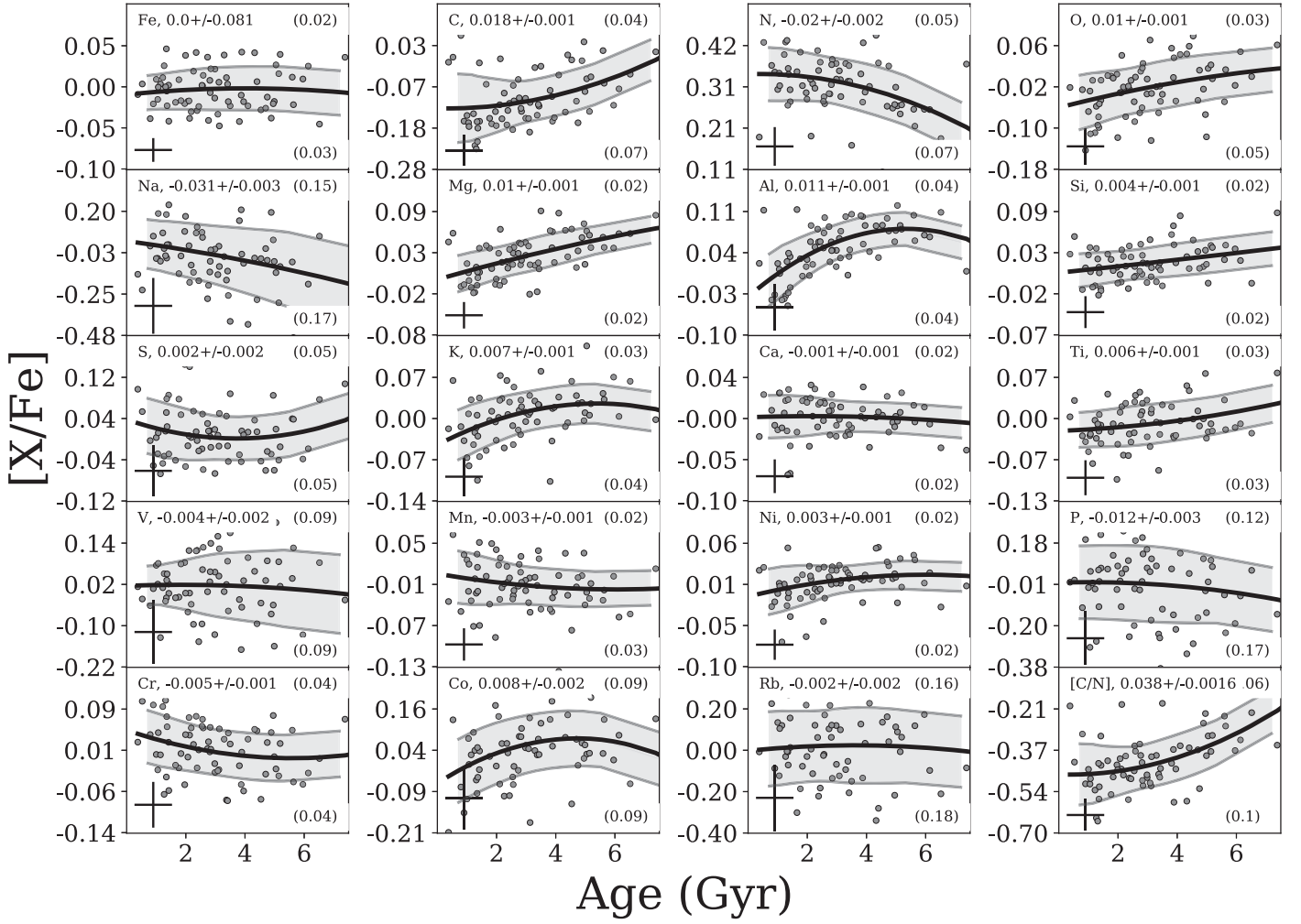


Figure 3. Summary of the small scatter around the age–abundance relations at fixed age and $[\text{Fe}/\text{H}]$. Shown are the ≈ 70 stars with asteroseismic ages in low- α sequence with $[\text{Fe}/\text{H}] = 0 \pm 0.05$ dex and signal-to-noise ratio > 150 , showing the age–abundance plane for 19 derived element abundances (we also show $[\text{C}/\text{N}]$ as this is used as an empirical age indicator, i.e., Martig et al. 2016). The best-fit second-order polynomial line is drawn through the age–abundance plane (in black), and the 1σ dispersion around this line is shown by the shaded gray regions around this central fit. Note that the top left-hand sub-panel indicates the selection around $[\text{Fe}/\text{H}] = 0 \pm 0.05$ dex and the remaining panels are the elements themselves and corresponding line of best fit given this selection. The typical error on the measurement is shown with the error bar at the bottom left of each sub-panel and the error in the y-axis direction is included in brackets in the right-hand corner. The slope of the best-fit straight line is shown, for reference, at the top of each sub-panel, to indicate the scale of the age–abundance correlation for each element (note the y-axis ranges are scaled to best show the data for each element). The bottom right-hand bracketed number is the total mean rms dispersion of the data around the best-fit line (σ_{TOTAL}).

Figure 4 shows the measured intrinsic dispersion for each element (in square scatter points), ordered by increasing value of intrinsic dispersion as measured for our sample. A dashed line is drawn through the mean intrinsic dispersion value at ≈ 0.02 dex. The x-axis lists all elements with respect to $[\text{X}/\text{Fe}]$, except for Fe which is $[\text{Fe}/\text{H}]$. The filled gray bars for each element show the total 1σ dispersion around the best-fit age–abundance polynomial. The color of the square markers indicates the overall gradient of the calculated intrinsic dispersion: blue represents a net positive age–abundance gradient, red represents a negative gradient, and black represents a gradient consistent with zero, within the errors. For comparison, the measured intrinsic dispersion values from a set of ≈ 100 solar-twin stars (around $[\text{Fe}/\text{H}] \approx 0$ dex) are shown as the circles (Bedell et al. 2018). The intrinsic dispersion values we measure from the red clump giants are remarkably similar to the main-sequence solar-twin sample of stars from Bedell et al. (2018) for many elements. Similarly to the solar-twin study, these stars are restricted in their spatial

extent to the *Kepler* field. Na and Rb are excluded from our analysis, as they are poorly measured (see Section 2). For Na we note the age–abundance trend is also counter to the expectation of other studies (e.g., Bedell et al. 2018). Therefore we do not wish to draw conclusions regarding this element.

Our abundance errors are important in the calculation of our intrinsic dispersion results, and we have taken care to ensure these are the most accurate estimates given our data. We find that the elements Co, Al, S, Si, V, Ca, Ni, Ti, Mg, and Cr all have lower than mean intrinsic dispersion values. This is not necessarily surprising for Si, Ti, Mg, and Ca: these are α -elements and we have conditioned on the α -element abundance of stars, selecting only those stars in the low- α sequence. Thus, selecting on a narrow range of these abundances renders them only fractionally additionally informative beyond being indicators of age. The elements Fe, K, Mn, O, N, C, and P all have intrinsic dispersion measures that are larger than the mean of the set of elements, up to the highest value measured of

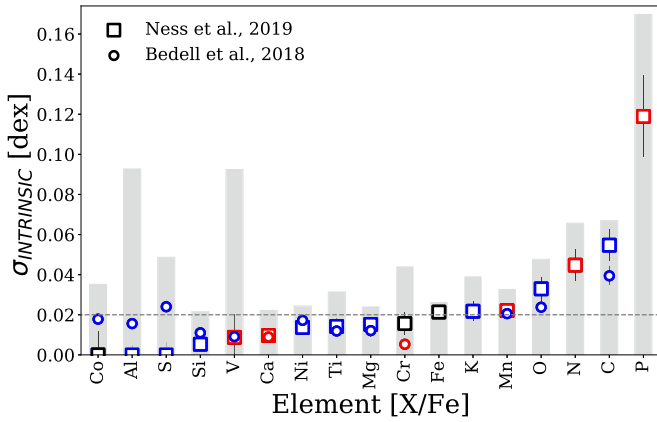


Figure 4. Intrinsic dispersion for stars at $[\text{Fe}/\text{H}] = 0 \pm 0.05$ dex, measured for each element in Figure 3 around the age–abundance relation fit by a second-order polynomial model to the data, as shown by the square markers. The elements are arranged in order of increasing intrinsic dispersion, $\sigma_{\text{INTRINSIC}}$. This is the dispersion around the age–abundance trend not explained by the measurement error, $\sigma_{\text{MEASUREMENT ERROR}}$. The gray filled bars show the total mean rms dispersion for each element around the best-fit line (σ_{TOTAL}). Our results for the red clump stars are very similar to the intrinsic dispersion measurements for the solar-twin study of Bedell et al. (2018), also shown for comparison by the round markers. The markers are colored according to the sign of the gradient calculated from a linear fit to the data; black is consistent with zero, red is negative, and blue is positive. The dashed line indicates the mean intrinsic dispersion measure for the elements, of ≈ 0.02 dex.

≈ 0.12 dex for P. The intrinsic dispersion results are, however, very sensitive to the accuracy of our uncertainty. The element that is measured most imprecisely, P, has the highest intrinsic dispersion. This may be a consequence of any inaccuracy in the uncertainty estimate, which will propagate with a higher amplitude in the intrinsic dispersion calculation, of the rms difference of the total (larger scatter) and the total measurement error. No comparative measurement is available for this element from the solar-twin study. We note, however, that P is produced in massive stars and chemical evolution models cannot reproduce the distribution of this element in the Milky Way (Cescutti et al. 2012). We have no reason to expect our measurement uncertainties are not accurate. We have taken into account the signal-to-noise dependency of the error and the parameter space of our red clump stars across T_{eff} , $\log g$ and $[\text{Fe}/\text{H}]$. Our findings of small intrinsic dispersion values for the elements around the age–abundance relations are consistent with what has been already found in the solar-twin study of Bedell et al. (2018), where the mean intrinsic dispersion across 66 element measurements is ≈ 0.02 dex (and ranges from 0–0.06 dex).

3.2. Constructing a Model of Ages from Abundances to Measure Intrinsic Dispersions Across (Age, $[\text{Fe}/\text{H}]$)

We now construct an abundance model to infer ages and overall metallicity, $[\text{Fe}/\text{H}]$, for the 15,000 APOGEE red clump stars in the low- α sequence, that span Galactic radii 4–16 kpc. We construct this model similarly to that of Ness et al. (2015), but instead of making a spectral model to generate stellar flux given abundances, we make a model to generate the abundance vector given age and $[\text{Fe}/\text{H}]$. We use the ≈ 600 low- α red clump stars with ages measured from the *Kepler* mission as our training set, for which we have APOKASC ages, with fractional age errors of $< 30\%$ and signal-to-noise ratio (S/N) > 150 , as well as high-fidelity $[\text{Fe}/\text{H}]$ measurements precise to

≈ 0.03 dex. From this training set, we make an abundance model, using the 18 abundance features of $[\text{C}, \text{N}, \text{O}, \text{Mg}, \text{Al}, \text{Si}, \text{S}, \text{K}, \text{Ca}, \text{Ti}, \text{V}, \text{Mn}, \text{Ni}, \text{P}, \text{Cr}, \text{Co}, \text{Rb}, [\text{C}/\text{N}]]$, with their associated uncertainties. For our model labels, we take both age and $[\text{Fe}/\text{H}]$. We exclude Na due its anomalous behavior, although note that including it does not change the results. We leave Rb in our set of abundances since, although it is poorly measured as indicated in Section 2, our inference is marginally improved (at the 5% level) with its inclusion.

Our model is then characterized by a coefficient vector θ_i for our i abundances, which allows us to predict each abundance X_{ni} for our n training objects, for a given label vector, ℓ_n , of age, τ , and overall metallicity, $[\text{Fe}/\text{H}]$. We write our abundance model as a linear function of a vector ℓ_n built from the labels:

$$X_{ni} = \theta_i^T \cdot \ell_n + \text{noise} \quad (1)$$

In this equation, θ_i is the set of spectral model coefficients at each $[\text{X}/\text{Fe}]$ for the i abundances. The noise is an rms combination of the associated uncertainty variance, σ_{ni}^2 of each of the abundance measurements (itself an rms sum of the cross-validation and formal uncertainty from the optimizer we use in python, *curve_fit*) and the intrinsic variance or scatter of the model of the fit at each abundance, s_i^2 . Our quadratic model fits a set of coefficients at each feature, in this case each abundance measurement, given our labels, in this case ages and $[\text{Fe}/\text{H}]$. The training stage optimizes not only for the coefficients but also a scatter term. This scatter term is equivalent to the (normally distributed) deviations from the approximate model for the true feature (abundance) variance around the best-fitting (second-order that we chose here) polynomial function to the data. This scatter is used by the model at the inference stage as an effective weighting of the features to infer the labels. This model assumes that the noise model is

$$\text{noise} = [s_i^2 + \sigma_{ni}^2] \xi_{ni} \quad (2)$$

where each ξ_{ni} is a Gaussian random number with zero mean and unit variance.

We use a polynomial model of our two labels of age and overall metallicity, τ and $[\text{Fe}/\text{H}]$, such that there are six coefficients at each abundance feature $\ell_n = [1, [\text{Fe}/\text{H}], \tau, ([\text{Fe}/\text{H}] \cdot \tau), ([\text{Fe}/\text{H}])^2, (\tau)^2]$, where the first element “1” permits a linear offset in the fitting. For the $[\text{Fe}/\text{H}]$ and age labels τ , these values are scaled to a mean of 0 and a variance of 1 in order to keep the model stable and pivoting around a reasonable point in label space.

We test our model using cross-validation, training on 95% of the data and testing on 5%, which we perform 20 times. The result of this inference for age is shown in the left-hand panel of Figure 5. The asteroseismic age is denoted as “True Age” on the x -axis and our inferred age from our model is on the y -axis. We infer age to a precision of about $\lesssim 1.6$ Gyr with almost no overall bias, although there is some apparent systematic effect at an asteroseismic (APOKASC) age of around 6 Gyr, for which a range of ages is inferred, up to < 12 Gyr. We infer the metallicity, $[\text{Fe}/\text{H}]$, for the stars to an rms precision of 0.07 dex with negligible bias. Our model is generative, which means we can generate the abundance vector given our ages and overall metallicity for each star, τ and $[\text{Fe}/\text{H}]$. The goodness of fit of our model compared to our data for the 15,000 stars in the low- α sequence for which we infer ages is shown in the right-hand panel of Figure 5. We quantify our goodness of fit with a χ^2

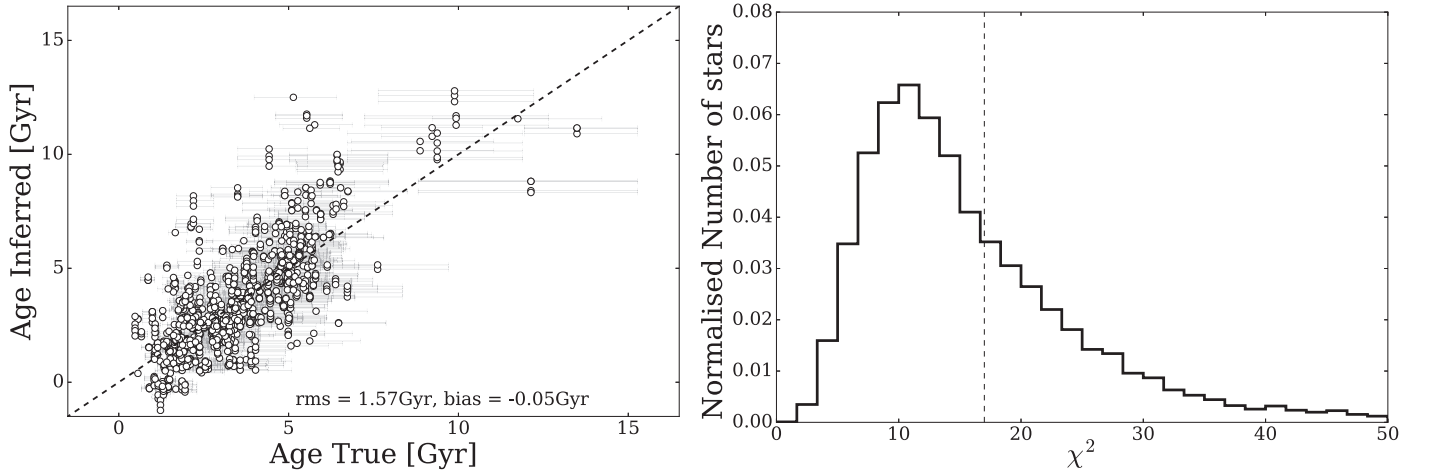


Figure 5. Left: the cross-validation results for our age inference using our quadratic model of the abundances given the ages and $[\text{Fe}/\text{H}]$, for our ≈ 600 stars in the training set (see Equation (1)). We remove 5% of stars to create 20 models each with a different missing 5% to generate this figure. Each of the 20 models is used to infer ages for the removed stars (shown). This training set of stars comprises stars with $S/N > 150$ and asteroseismic age errors $< 30\%$. The model returns ages precise to $\lesssim 1.6$ Gyr, and $[\text{Fe}/\text{H}]$ to 0.07 dex precision (not shown). Right: the χ^2 distribution of the model compared to the data for the $\approx 15,000$ red clump stars in the low- α sequence for which we derive ages and $[\text{Fe}/\text{H}]$ from our model. The dashed line indicates the roughly expected peak of the χ^2 distribution given the 18 abundance features that are used to infer the age and $[\text{Fe}/\text{H}]$ labels. Our distribution peaks below this expectation, which is likely a consequence of correlated features and errors, not accounted for in our simple χ^2 metric. Nevertheless this is demonstrative that our model generates a good fit to the data for the red clump stars in the disk, with only a small tail of stars with high χ^2 values where the model poorly fits the data.

metric where, for each star,

$$\chi^2 = \sum_{i=1}^{18} \frac{(X_{i,\text{model}} - X_{i,\text{data}})^2}{(s_i^2 + \sigma_i^2)} \quad (3)$$

for the $i = 1-18$ abundance features. The χ^2 value for the full set of low- α red clump stars with $-0.5 < [\text{Fe}/\text{H}] < 0.45$ dex peaks at $\chi^2 = 12\%$ and 90% of our stars have χ^2 values of < 35 , which indicates that our model is a good fit to the data for our sample, where we fit for 18 individual abundance features. Note the peak of the χ^2 is lower than the nominal expectation from this many label variables (18), and we expect this to be a consequence of the element correlations. We exclude stars with a χ^2 metric of > 35 for our subsequent analysis. These more poorly fit stars represent only a small ($< 8\%$) fraction of stars for which the inferred age and $[\text{Fe}/\text{H}]$ labels do not generate a good fit to the abundance data. This quality cut in χ^2 ensures we only include ages where the generated model to the abundance data is a fairly good fit, at worst. We examined the χ^2 fit of the model compared to the data in detail and, for a given age, this does not change significantly as a function of orbital property or Galactic location. This leads us to conclude that our local model of the age–abundance relation holds globally across the disk. If the low- and high- α sequences were combined for such analysis this would not be apparent, as the high- α stars show a different abundance–age relation to the low- α sequence of stars (see Bedell et al. 2018). The relative fraction of high- and low- α stars changes as a function of Galactic radii which, combined, would look like a changing abundance–age relation across the disk which our local model would not be able to fit.

Figure 6 shows the scatter, s , for each element from our quadratic model fit (see Equation (2)), in order of increasing scatter (shown as black squares). The scatter term is effectively a global measure (across all $[\text{Fe}/\text{H}]$ and ages) of the overall intrinsic dispersion of each element, measuring how much

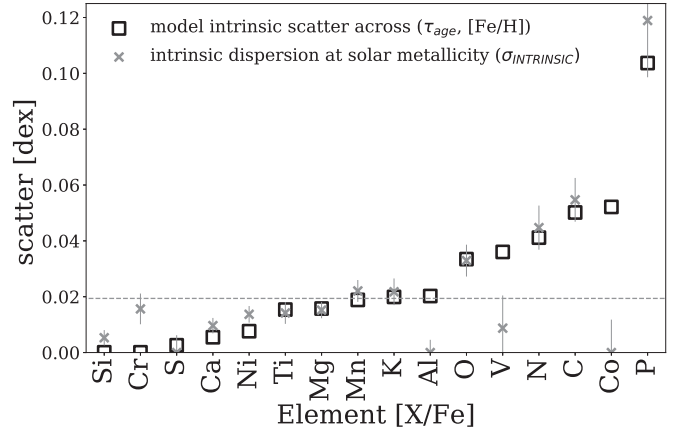


Figure 6. Model’s scatter terms for each element, which parameterize how well the second-order polynomial model fits the data for each element, given age and $[\text{Fe}/\text{H}]$ labels. This model is constructed using ≈ 600 stars. The elements are arranged in order of increasing scatter and the magnitude of the scatter is indicated by the square markers. The scatter term represents the intrinsic dispersion of each element not accounted for by the error on the abundance label, when fitting for $[\text{Fe}/\text{H}]$ and age simultaneously. The intrinsic dispersion of each abundance around the abundance–age relation at solar metallicity calculated from ≈ 70 stars is shown as gray crosses for comparison (from Figure 4).

additional variance around the model of abundances given ages and $[\text{Fe}/\text{H}]$ there is not accounted for by the (abundance) label errors. The intrinsic dispersion reported in Figure 4 by comparison is only for solar metallicity stars, $[\text{Fe}/\text{H}] = 0 \pm 0.05$ dex. Here, we are generating the abundances using both age and $[\text{Fe}/\text{H}]$ simultaneously. The scatter shown in Figure 6 is formally an upper limit of the intrinsic dispersion, given our second-order polynomial model, as we do not take into account the contribution of the age errors in our model. However, the contribution of the age uncertainties are small for the APOKASC sample. We ensure this by excluding stars with high age errors of $> 30\%$, and the average age error is 0.7 Gyr, which for the weak age–abundance relations does not

propagate into a significant contribution to the total measurement uncertainty. We confirmed in our analysis in Section 3.1 that the age uncertainties have <0.01 dex contribution to the total dispersion around the best-fitting second-order polynomial age–abundance fit to solar metallicity red clump stars. We compare the scatter from our model across age and $[\text{Fe}/\text{H}]$ to the intrinsic dispersion we find for the solar metallicity stars in Section 3.1 in Figure 6 (shown as gray crosses). In general these are identical or similar, with Co showing the largest discrepancy. Presumably, the higher scatter of the model compared to the intrinsic dispersion at solar $[\text{Fe}/\text{H}]$ for Co is due to outliers farthest away from the mean for this element, for the larger sample across the broad metallicity range. We found a single outlier in $[\text{Co}/\text{Fe}]$ (that had made our quality cuts) at solar $[\text{Fe}/\text{H}]$ alone significantly increased the measured intrinsic dispersion and so was subsequently removed. Such outliers could be representative of the true distribution of Co across the full metallicity range (which is broad and deviates from Gaussianity), or else indicate this element is not measured robustly. We do not attempt to remove any outliers in our modeling of our larger sample across the broad range in $[\text{Fe}/\text{H}]$, as the stars we analyze meet our quality cuts and we take into account the measurement errors (and these outliers do not have high errors indicative of highly uncertain measurements). Another potential source of the higher model scatter compared to the intrinsic dispersion at solar $[\text{Fe}/\text{H}]$ may be the temperature-dependent systematic that is reported for Co in APOGEE (Holtzman et al. 2018).

The goodness of fit of our generated model is not a function of the orbital or radial properties of our sample. That is, the model fits the stars near the Sun as well as the stars in the inner and outer disk. Therefore, we can also conclude that our model is a good approximation for the age–abundance relations across the disk. Correspondingly, the scatter measurement for each abundance not only represents the intrinsic dispersion at a given (age, $[\text{Fe}/\text{H}]$) but also across the radial extent of the disk from 4 to 16 kpc which these stars span. This is demonstrative that, moving beyond the solar neighborhood (i.e., the solar-twin study of Bedell et al. 2018 and the analysis in Section 3.1 and Figure 4), the chemical evolution of the low- α disk has been largely homogeneous and consistent with what we see in the solar neighborhood. This places strong constraints on mixing and gas accretion in the disk.

3.3. The Age–Abundance Trends for Stars across the Disk

We have generated a model to infer age and $[\text{Fe}/\text{H}]$ given abundances and subsequently determined ages for the red clump sample of stars that cover the parameter space of our training set. The high-fidelity set we can model well (with $\chi^2 < 35$) corresponds to about 15,000 stars which are shown colored by age in the right-hand panel of Figure 2. Our cross-validation test using the training set described in Section 3.2 indicates that these inferred ages are precise to $\lesssim 1.6$ Gyr (see Figure 5). From Figure 2 there is a clear trend (even for the low- α stars only) with $[\text{Mg}/\text{Fe}]$, with the abundance value increasing with age, which is indicative as to why α -element enhancement has been a long-used age proxy (e.g., Adibekyan et al. 2013; Haywood et al. 2013; Silva Aguirre et al. 2018) (although more typically to distinguish “young” α -rich and “old” α -poor stars). However, at any given $[\text{Mg}/\text{Fe}]$ value, there are a range of ages spanned, with age increasing as a function of $[\text{Fe}/\text{H}]$ at a given $[\text{Mg}/\text{Fe}]$. At any given $[\text{Fe}/\text{H}]$,

there is also a range of ages spanned, from the highest- to lowest-metallicity stars.

We show the age versus abundance results for metal-rich compared to more metal-poor stars in Figure 7. In this figure we show the best-fit age–abundance trends (again using a second-order polynomial model) for ≈ 1800 stars with $[\text{Fe}/\text{H}] = -0.35 \pm 0.05$ dex and ≈ 2300 stars with $[\text{Fe}/\text{H}] = 0 \pm 0.05$ dex. The second-order polynomial fit lines to the age–abundance trends for element for the two metallicity bins are shown in red for the metal-rich stars ($[\text{Fe}/\text{H}] = 0 \pm 0.05$) and in blue for the metal-poor stars ($[\text{Fe}/\text{H}] = -0.35 \pm 0.05$). The running means of the 1σ dispersion of the abundance measurements are shown in the thicker shaded regions either side of the best-fit lines, colored red and blue respectively. It is clear from this large set of stars (more so compared to the analysis in Section 3.1 and Figure 4) that, for many elements, the age–abundance trend is not described by a simple linear function. Most elements increase and flatten (e.g., C) or even turn over (e.g., Ni) at older ages. The age–abundance relationships also vary in amplitude and slope, as a function of $[\text{Fe}/\text{H}]$. This figure also reveals that there are small differences in the dispersion values of $[\text{X}/\text{Fe}]$ across age, as a function of $[\text{Fe}/\text{H}]$.

Taking $[\text{O}/\text{Fe}]$ as an example from Figure 7, this element shows metallicity-dependent age–abundance trends. The more metal-poor stars have a higher O enhancement and show a steeper rate of abundance decrease with age compared to the more metal-rich stars. By comparison the iron-peak elements V and Cr show basically no differences for the two metallicity bins. Of the α -elements, O, Mg, Si, and Ti show positive age–abundance relationships and Ca shows a negative age–abundance trend. Of all α -elements, Ti is the only element where the metal-poor stars are less Ti-enhanced than the metal-rich stars and show a flatter age–abundance slope. The element Si shows the most separation of the α -elements in the age–abundance relation between the two metallicity bins. This highlights that α -elements have individual age–abundance behaviors as a function of both $[\text{Fe}/\text{H}]$ and age, and themselves trace the detailed chemical evolution history of the low- α sequence of stars. Such comparative relationships are extraordinarily constraining for chemical evolution models and yield tables. Figure 8 also highlights that, for studies of the disk, if stars are not considered in bins of both age and $[\text{Fe}/\text{H}]$, then the net effect of the different trends with $[\text{Fe}/\text{H}]$ and age will be to dilute the underlying relationships that carry the information about the detailed star formation history, which we can see here is clearly richly described in the data. We reiterate this is for the low- α sequence of stars only; the full distribution of stars increases the complexity of these relationships and proceeding by considering them separately is well motivated given their likely entirely separate origins (e.g., Clarke et al. 2019; Gandhi & Ness 2019; Mackereth et al. 2019a).

4. Angular Momentum Gradients with Abundances in the Low- α Disk

We have investigated and shown the age–abundance trends for our set of red clump stars for which we have inferred ages precise to $\lesssim 1.6$ Gyr. We now wish to investigate the relationship between age, abundance, and stellar orbital properties.

There is a known age gradient in the disk, and in disks of spiral galaxies (e.g., Ness et al. 2016; Gadotti et al. 2019). Furthermore, there is a gradient in $[\text{Fe}/\text{H}]$ that exists at fixed

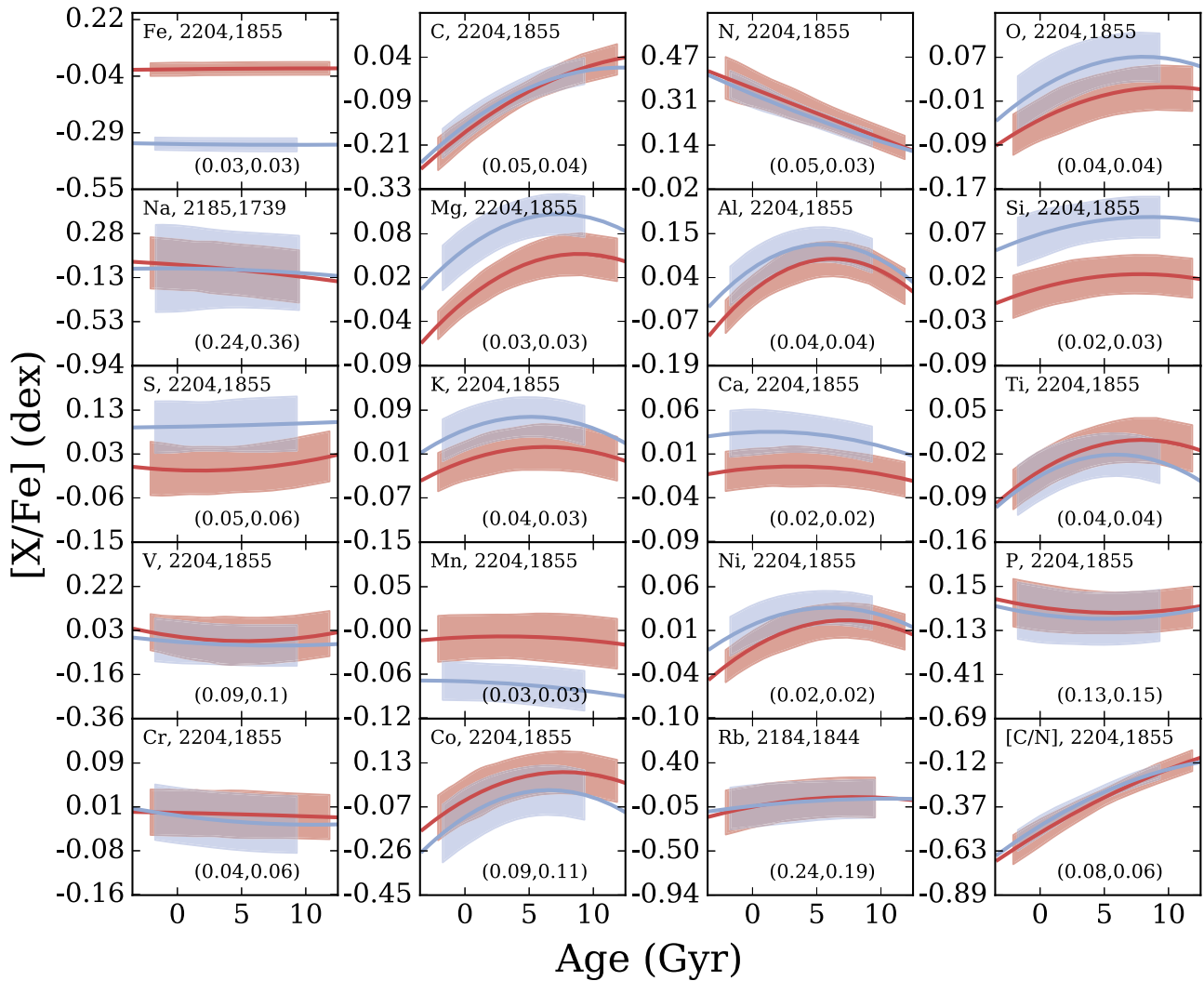


Figure 7. Best-fit second-order polynomial lines to the ≈ 1800 stars in our red clump sample with $[\text{Fe}/\text{H}] = -0.35 \pm 0.05$ dex (in blue) and the ≈ 2200 stars with $[\text{Fe}/\text{H}] = 0 \pm 0.05$ dex (in red), with the number of stars used to determine these lines indicated in each sub-panel. These stars shown within these narrow metallicity ranges are taken from the $\approx 15,000$ stars in the low- α sequence of the red clump stars with ages determined by modeling the relationship between the asteroseismic ages and abundances. The thicker regions around each line of best fit show the 1σ dispersion of the data. For most elements, the dispersion values for metal-rich and metal-poor stars are comparable. However, the mean value of any element at a given age is dependent on the overall $[\text{Fe}/\text{H}]$ and, for most elements, and the abundance-age slopes are typically steeper for the more metal-poor stars. Only stars with a $\chi^2 < 35$ are shown. Above this χ^2 threshold we deem the model to be a relatively poor fit to the data.

age (Ness et al. 2016). There is also an established correlation between stellar α -element abundances and age (e.g., Adibekyan et al. 2013; Haywood et al. 2013). In this work we show that most elements in fact correlate with age. From Figure 7, we expect that any gradients that we see in element abundance enhancements across the Galactic radius will in large part be due to the trends between abundances and age at a given $[\text{Fe}/\text{H}]$.

We now have access to not only the stellar radius of our stars, which is their present-day orbital position, but from *Gaia* proper motion measurements their angular momentum, J_ϕ , which enables a more accurate description of a star's average orbital radius in the disk. We use the angular momentum to describe this average stellar location, where small J_ϕ corresponds to the inner Galaxy, and large J_ϕ , to the outer Galaxy. Our red clump stars span a Galactic radius of about 4–16 kpc and, correspondingly, a J_ϕ range of about 500–3500 km s^{-1} kpc. Figure 8 shows the J_ϕ -age distribution of the low- α stars colored by $[\text{Fe}/\text{H}]$ at left, and $[\text{Mg}/\text{Fe}]$ at right. We show $[\text{Mg}/\text{Fe}]$ as this is representative of an

α -element, which is a well studied and mapped element across the disk, and has also been used as an age proxy. The dashed line in the left-hand panel shows the running mean of J_ϕ as a function of age, which shows the age gradient across the disk, with the mean stellar age decreasing with increasing distance from the Galactic center.⁷ The gray-shaded area around the line is the sampling error, representing the uncertainty of this mean J_ϕ measurement across age. The relationship between the abundance plane, J_ϕ , and age is clearly demonstrated in this figure. There is a strong trend between J_ϕ and $[\text{Fe}/\text{H}]$: this is the metallicity gradient in the disk with radius, which has been previously measured from Cepheid tracers (e.g., Luck & Lambert 2011) and known to be around $-0.06 \text{ dex kpc}^{-1}$, with the $[\text{Fe}/\text{H}]$ increasing from the outer to the inner disk (from larger to smaller J_ϕ). The total mean

⁷ This age gradient is fairly shallow, but note we are examining only the low- α stars. The age gradient is far steeper considering both low- and high- α stars simultaneously, as the high- α stars are on average older and concentrated in the inner Galaxy whereas the low- α stars are on average much younger and concentrated in the outer Galaxy.

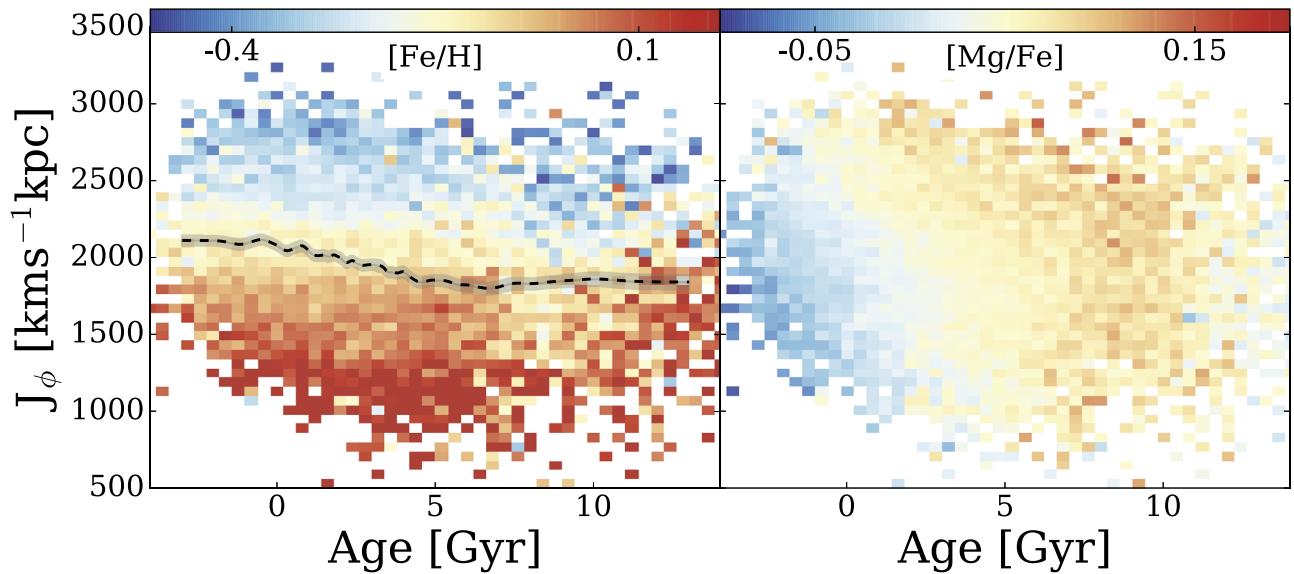


Figure 8. Summary of the abundance projections ($[\text{Fe}/\text{H}]$, $[\alpha/\text{Fe}]$) into the age- J_ϕ plane, for the 15,000 low- α red clump stars for which we have determined ages. This plane is shown colored (left) by $[\text{Fe}/\text{H}]$ and (right) by $[\text{Mg}/\text{Fe}]$, where the color bar is shown at the top of each sub-panel. The dashed line in the left-hand panel shows the running mean of J_ϕ across age, which reveals the gradient across age in J_ϕ : the mean age of the disk decreases with increasing Galactic radius. The left-hand panel shows that $[\text{Fe}/\text{H}]$ is strongly correlated with J_ϕ : the disk has a large overall metallicity gradient that decreases with increasing Galactic radius. At fixed J_ϕ there is also an $[\text{Fe}/\text{H}]$ gradient across age: younger stars are more metal rich than older stars at fixed J_ϕ . The right hand-panel shows that Mg enhancement is an excellent age proxy: at fixed J_ϕ there is a strong gradient with $[\text{Mg}/\text{Fe}]$ across age. At fixed age, there is also a weak trend in $[\text{Mg}/\text{Fe}]$: the $[\text{Mg}/\text{Fe}]$ abundance of stars increases with increasing radius.

change in $[\text{Fe}/\text{H}]$ across $J_\phi = 1000\text{--}3000 \text{ km s}^{-1} \text{ kpc}$ is ≈ 0.5 dex. This is consistent with the circular velocity of 220 km s^{-1} (e.g., McMillan 2017), whereby the J_ϕ gradient for a star on a circular orbit will be $\approx -0.6 \text{ dex kpc}^{-1} / (220 \text{ km s}^{-1}) = -0.27 \text{ dex} / (1000 \text{ km s}^{-1} \text{ kpc})$.

There is a very slight gradient with age in that, at fixed J_ϕ , higher-metallicity stars are younger than more metal-poor stars. For the abundance enhancement $[\text{Mg}/\text{Fe}]$, there is also a gradient seen between J_ϕ and $[\text{Mg}/\text{Fe}]$ as a function of age: at fixed J_ϕ the $[\text{Mg}/\text{Fe}]$ increases as stellar age increases. There is also a slight gradient seen across $[\text{Mg}/\text{Fe}]$ with J_ϕ : at fixed age, $[\text{Mg}/\text{Fe}]$ decreases with decreasing J_ϕ .

We next examine the full set of element abundance enhancement- J_ϕ relationships for the red clump stars across the disk. Figure 9 shows the ≈ 2000 solar metallicity, $[\text{Fe}/\text{H}] = 0 \pm 0.05$ dex stars from our sample. For these stars, we show the second-order polynomial line of best fit to the abundance- J_ϕ measurements (as the black dashed lines). We also show the best-fit second-order polynomial line (as the filled red lines in Figure 9) to the difference between the age-abundance trend (shown in Figure 7) and the value of each star’s abundance (which we call the age-abundance residuals), and its corresponding J_ϕ measurement, for each element. The gray shaded and red hatched regions around these lines are the 1σ dispersion of the measurements (for the abundance measurements and the residuals away from the age-abundance fit, respectively). The elements we exclude as being unreliable are removed from this figure, which is otherwise in the same order as the previous Figure 7. The magnitude of the abundance change over the range of $J_\phi = 1600\text{--}2200 \text{ km s}^{-1} \text{ kpc}$ is shown in the top of each sub-panel for the abundance data and the residual data, respectively. We expect and see that elements most correlated with age show the largest gradients in their abundance- J_ϕ trends (i.e., C and N). We see that, for elements with small intrinsic dispersion values in particular, removing the age-abundance trend, as shown by

plotting the best-fit line to the residuals, results in a flat abundance residual- J_ϕ trend compared to the abundance- J_ϕ trend. Note that the element Co shows a residual trend not accounted for by the age-abundance relation. This may be a consequence of the known systematic measurement errors of this element with T_{eff} reported by APOGEE (Holtzman et al. 2018).

5. The Orbital Separation of Stars Using Abundances at Fixed $[\text{Fe}/\text{H}]$

We have seen that the abundance architecture of stars is linked to their age and J_ϕ . We now seek to examine if the detailed abundances of stars indicate their orbital properties. We proceed by examining the mean orbital actions of stars that are grouped by their abundance similarity, at a fixed $[\text{Fe}/\text{H}]$. We take mono- $[\text{Fe}/\text{H}]$ bins of our sample of disk stars. We take these bins across the metallicity range of $-0.4 \leq [\text{Fe}/\text{H}] \leq 0.25$ dex (where there are a sufficient number of stars for this test). The width of each of the $[\text{Fe}/\text{H}]$ bins is approximately the error in $[\text{Fe}/\text{H}]$: $\sigma_{[\text{Fe}/\text{H}]} = 0.035 + 0.05 \times |[\text{Fe}/\text{H}]|$. The mean radii for these metallicity selections span from $R_{\text{GAL}} = 8.6$ to 10.3 kpc (with the mean radii increasing from the more metal-rich to more metal-poor stars). Here, we also condition on (i) vertical action $J_z < 10 \text{ km s}^{-1} \text{ kpc}$,⁸ thus excluding stars that make excursions away from the plane of the disk, and (ii) age, examining both only the youngest stars ($< 1 \text{ Gyr}$) and then comparing this to intermediate-age stars ($> 5 \text{ Gyr}$).

For the stars assigned within each metallicity bin, we use the k-means clustering algorithm and the $[\text{X}/\text{Fe}]$ information, similarity to Hogg et al. (2016), but here to separate the stars into *two* groups at each fixed $[\text{Fe}/\text{H}]$, which we denote as $N = 1$ and $N = 2$. Each grouping is independent, across $[\text{Fe}/\text{H}]$ and across each action at fixed $[\text{Fe}/\text{H}]$. We allow each

⁸ This corresponds to a maximum vertical excursion $z_{\text{max}} \sim 0.7 \text{ kpc}$ at $R_{\text{GAL}} = 8 \text{ kpc}$ and $z_{\text{max}} \sim 0.78 \text{ kpc}$ at $R_{\text{GAL}} = 10 \text{ kpc}$.

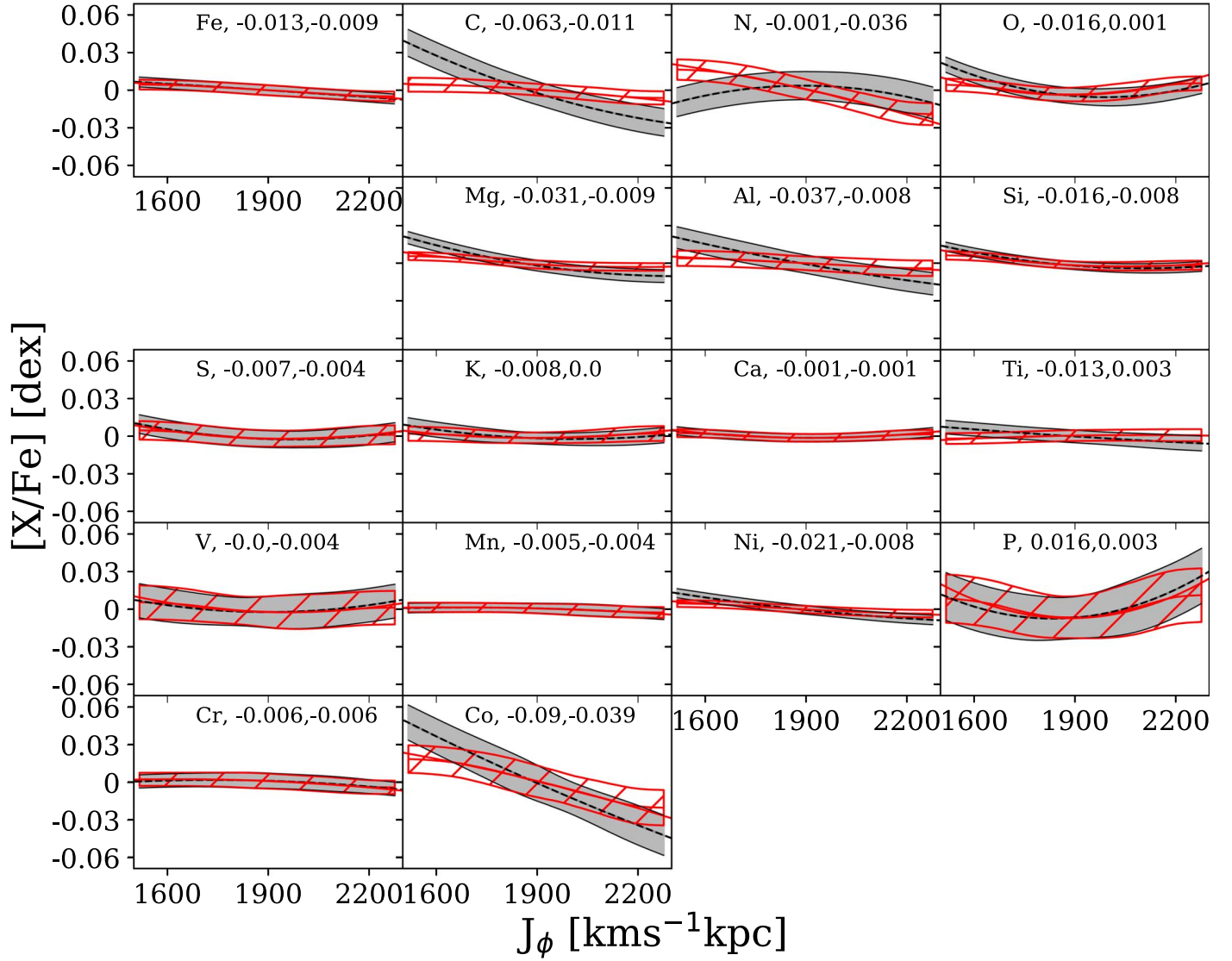


Figure 9. Approximately 2000 red clump APOGEE stars across the disk, with $[\text{Fe}/\text{H}] = 0 \pm 0.05$ dex. The black dashed line shows the best-fitting second-order polynomial to the abundance- J_ϕ measurements for each element. The gray shaded region shows the 1σ dispersion of the data around this fit. The red line shows the best-fitting second-order polynomial to the difference between the age-abundance trend (shown in Figure 7) and the value of each star’s abundance (which we call the age-abundance residuals), and its corresponding J_ϕ measurement, for each element. The red hatched region shows the 1σ dispersion of the data around this fit. The red hatched trends therefore represent the relationship between each element and J_ϕ , with the age-abundance relation removed. For almost all elements, the removal of the age-abundance trends (almost entirely) flattens the correlation of the element with J_ϕ , as shown in the best-fit to the age-abundance residual values in red. This demonstrates that the gradients seen between abundances and orbital radius in the disk are largely accounted for by the fact that elements indicate age. The trends shown look very similar for the lower metallicity selection of stars ($[\text{Fe}/\text{H}] = -0.35 \pm 0.05$ dex) shown in Figure 7.

element to have a weighting coefficient and optimize the coefficients for each action at each $[\text{Fe}/\text{H}]$, to maximize the separation between $N = 1$ and $N = 2$ for each orbital action (note that such a coefficient is not strictly necessary, but the otherwise equal weighting of the scaled abundances is arbitrary). These two groups, $N = 1$ and $N = 2$, represent populations of nearest-neighbor abundance similarity and are determined from the $[\text{X}/\text{Fe}]$ of the stars (not including, but rather at fixed $[\text{Fe}/\text{H}]$), for the low- α sequence (given our additional age and vertical action cut described). The $N = 1$ and $N = 2$ populations therefore represent groups of stars most chemically similar to one another: stars within cluster $N = 1$ and within cluster $N = 2$ are more similar to each other than stars between the $N = 1$ and $N = 2$ clusters. We assert two clusters as we are restricted by our number of stars, but this

exercise could be generalized to examining the orbital separation of an arbitrary number of groups.

The mean actions J_ϕ , J_R , and J_z for each of the two clustered assignments of stars into the $N = 1$ and $N = 2$ groups at each fixed $[\text{Fe}/\text{H}]$ are shown in Figures 10 and 11, for each mono- $[\text{Fe}/\text{H}]$ bin, for the stars with ages < 1 Gyr and > 5 Gyr, respectively. The x-axis for these figures shows the mean $[\text{Fe}/\text{H}]$ of the groups, where most groups have near-identical mean $[\text{Fe}/\text{H}]$ values, typically to the second decimal place, and the y-axis shows the mean action, respectively, for each action. The mean action values for each pair of stars within groups $N = 1$ and $N = 2$ are colored by the age difference of the pair, normalized by the standard deviation of the age range of the stars (we do the normalization so as to use a consistent scale for our < 1 Gyr selection and > 5 Gyr selection of stars).

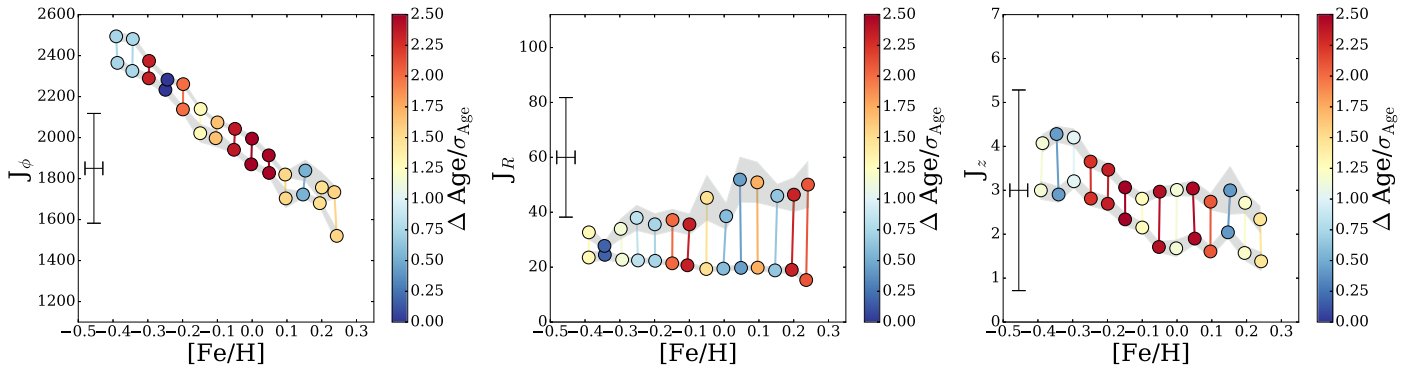


Figure 10. Mean actions (J_ϕ , J_R , J_z), for low- α disk stars divided into two groups based on their element abundance similarity, $[X/Fe]$, at fixed $[Fe/H]$. The division of the stars at each $[Fe/H]$ into two groups is done using the k-means algorithm where each grouping is independent across $[Fe/H]$ and each action. Each pair at fixed $[Fe/H]$ is joined with a line. These groups are comprised of ≈ 6500 stars in total, with ages < 1 Gyr and span an $[Fe/H]$ range $-0.5 \text{ dex} < [Fe/H] < 0.45 \text{ dex}$. The stars have also been selected to be confined to the disk, with $J_z < 10 \text{ km s}^{-1} \text{ kpc}$. The circles, which show the mean action values for the two groups at each $[Fe/H]$, and the line that joins them, are colored by the age difference between the pair, normalized by the 1σ standard deviation of the age range of the stars in this selection. The errors on each mean action measurement are shown as the thick gray line joining the open circles. This figure shows that populations of stars with different $[X/Fe]$, at a fixed $[Fe/H]$, separate out into groups of different mean ages with different mean actions. The error bar at the left shows the average 1σ standard deviation of the individual action measurements of the stars that comprise the groups, around their mean.

Note that the members in each pair $N = 1$ and $N = 2$ are different for each action plot (with corresponding different mean ages and age differences). This is because for each mono- $[Fe/H]$ bin for each action, we independently optimize the weighting of each element to maximize the action separation. The purpose of this is to test if abundances can indicate age and orbital properties for each component of the orbit (as opposed to testing if groups separated in, say, J_ϕ are discrete in their other orbital actions, which under radial migration we expect to not be the case). Note that we are also not trying to examine groups of stars of a common cluster birth origin, to test how their orbits may behave. We are not in the regime where we can reconstruct common birth sites of stars. We instead can examine if there is structure in the abundance–orbital plane, and we see that this structure links to stellar age. Without using the age measurements for the stars themselves, we can see that the vector of abundances reveals the structure in the dynamical configuration of the disk.

We find that those abundances that are the most important to maximally separate each action (at each $[Fe/H]$) change for each action. This must link to the initial radial, vertical, and azimuthal properties of the gas disk from which the stars were formed, and the timescales of its enrichment and assembly. Over time, this link is confounded by how the stars have subsequently moved.

Each pair of groups of stars, $N = 1$ and $N = 2$, are joined with a solid line, which is colored by the normalized age difference of the pair. The 1σ standard error on the mean action measurement of the groups is shown as the gray line joining the markers across all of $N = 1$ and across all of $N = 2$. The typical 1σ standard deviations around the mean action measurements are indicated at the far left of each sub-panel.

In both Figures 10 and 11, the two groups $N = 1$ and $N = 2$, based on abundance similarity, have different mean ages for each pair (with a couple of exceptions). This is not surprising since abundances indicate age, as seen in Section 3. Groups based on similarity of abundances are effectively groups in age. The groups $N = 1$ and $N = 2$ in Figures 10 and 11 also show significant differences in their mean actions. Notably, the separation in J_ϕ between $N = 1$ and $N = 2$ is larger for the stars > 5 Gyr in Figure 11, compared to the stars < 1 Gyr in

Figure 10. This is indicative of stars moving from their birth locations over time. The stars with ages < 1 Gyr have simply had less time to migrate across the disk. Presumably this separation of the groups $N = 1$ and $N = 2$ is a qualitative measure of the migration of stars over time.

The J_R result for young stars with ages < 1 Gyr in Figure 10 shows an increasing action separation between pairs of $N = 1$ and $N = 2$ groups for metal-rich compared to metal-poor stars. This trend is not seen for stars with ages > 5 Gyr in Figure 11. Unlike the J_ϕ result, the separation between the groups in J_R is similar for stars with ages < 1 Gyr and > 5 Gyr. Overall, the mean J_R amplitude increases with stellar age. The J_z result for young stars with ages < 1 Gyr in Figure 10 shows that J_z decreases for these stars with increasing $[Fe/H]$. This is not seen for stars > 5 Gyr in Figure 11. Similarly to J_R , while the mean J_z increases for both $N = 1$ and $N = 2$ groups for ages < 1 Gyr to > 5 Gyr, the mean action separation between the two groups does not significantly change (although the presumably random variation in action difference across $[Fe/H]$ appears higher for older stars). We note that the magnitude of the difference in age does not, in general, correspond to the action separation. This is with the exception of the two pairs of stars in groups $N = 1$ and $N = 2$ with around zero age difference at ages < 1 Gyr, for J_ϕ and J_R . These pairs of stars show no measurable action difference in J_ϕ and J_R . The intrinsic dispersion around the mean actions is high (on order of the mean action values themselves for J_R and J_z), and significantly larger for older stars. This indicates that, just because the stars separate in mean orbital actions at a given metallicity, individual orbital actions themselves will not necessarily be able to be well recovered from the abundance vector for a star (see also Beane et al. 2018).

6. Discussion

There has been an immense increase in the size and dimensionality of chemical and dynamical data of stars. In this work, we seek to interpret these data by understanding the relationship between individual abundance enhancements of stars and age. This is motivated by our understanding of the formation of galaxies, which gives us a picture of stars in our

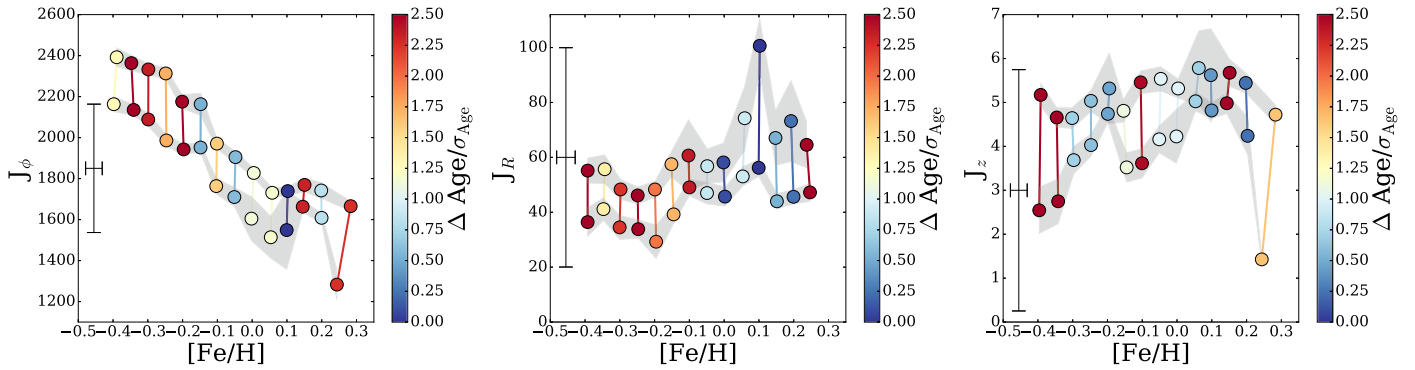


Figure 11. Same as Figure 10, only taking the ≈ 3500 intermediate-to-old stars with ages > 5 Gyr, across the $[\text{Fe}/\text{H}]$ range $-0.5 \text{ dex} < [\text{Fe}/\text{H}] < 0.45 \text{ dex}$. Again, the stars have also been selected to be confined to the disk, with $J_z < 10 \text{ km s}^{-1} \text{ kpc}$. As per Figure 10, this shows that populations of stars defined using their $[\text{X}/\text{Fe}]$ abundances at a fixed $[\text{Fe}/\text{H}]$ separate out into groups of different mean ages with different mean actions. Compared to Figure 10, older stars show greater separation in J_ϕ compared to younger stars, presumably as a consequence of having more time to migrate from their birth origin (Frankel et al. 2018). On average, comparing to Figure 10, both young and old stars shown similar average separation in J_R and J_z . However, typically there is larger seemingly random variation in the action difference between the pairs for the older stars shown here compared to stars < 1 Gyr old.

own Milky Way forming in a gas disk. Our understanding of galactic evolution suggests that this gas disk, and hence the radial extent of stellar birth sites, will grow with time. Our understanding of nucleosynthesis within, and chemical enrichment from, those stars implies that each generation of stars will contain signatures of prior generations (Rix & Bovy 2013; Bland-Hawthorn & Gerhard 2016). These insights combine into a picture where stars of a given age form from gas which is chemically segregated into mono-abundance rings, whose spatial extent gradually increases over time. Hence, the chemical abundance distributions of stars across the Galaxy are expected to tell us the story of how our Galaxy formed, while their locations or orbits tell us something about where.

This expectation has been exploited with forward models of chemical enrichment. These models attempt to capture the physics of galaxy formation and chemical enrichment, given the observed abundance distributions (e.g., Nidever et al. 2014; Loebman et al. 2016; Sanderson et al. 2018; Clarke et al. 2019; Mollá et al. 2019; Valentini et al. 2019). Two key uncertainties limit this approach: the extent to which the stars will lose their memory of the location of their birth sites as they migrate around the Galaxy through secular (e.g., resonances with bars and spiral arms) and non-secular (e.g., interactions with Galactic satellites or gas clouds) dynamical processes, and the sources, timescales, and yields of enrichment. A second approach to interpretation is to chemically tag stars to take advantage of the expectation that stars formed in the same gas cloud will have identical chemical signatures (Freeman & Bland-Hawthorn 2002; Krumholz et al. 2019). The main uncertainty here is to what extent star clusters will be chemically distinct within the measurement errors (Ting et al. 2015; Ness et al. 2018; Bovy et al. 2019; Liu et al. 2019).

This paper takes a step back from directly reconstructing Galactic history to instead explore where the information is in the actual data. We see from the data that individual element abundances of stars show rich and varied correlations with age, spatial location, and orbital actions. We have examined the relationships between age and abundance for the low- α stars of the disk, using 18 elements measured from the APOGEE red clump spectra. Our typical measurement precisions are < 0.03 dex and we do not have to consider diffusion and other astrophysical abundance offsets as a function of stellar

parameter space (e.g., Liu et al. 2019; Souto et al. 2019). We find that most abundances indicate age, albeit some only weakly. The elements C and N have the largest age gradients, presumably as a consequence of mass-dependent dredge-up on the giant branch that renders them age indicators. The elements Al, O and Mg are the next most correlated elements with age (although many of these element age-abundance trends are non-linear and flatten, and even appear to turn over at large ages). Presumably the age trends seen for Al, Mg, and O are a consequence of Galactic chemical evolution rather than surface abundance changes (as for C and N).

We measure the intrinsic dispersion around the age-abundance trends for each abundance, $[\text{X}/\text{Fe}]$, that is, the dispersion around the age-abundance relation that is not explained by the measurement errors. We measure this intrinsic dispersion in two ways. First, we consider only the ≈ 70 red clump stars at solar metallicity with asteroseismic ages. For these stars, we calculate the abundance dispersion around a second-order polynomial line of best fit to each age-abundance trend and determine what proportion of this dispersion is not explained by the measurement errors. We find that, for the low- α stars in the disk, most abundances have small intrinsic dispersion values at a given age at solar $[\text{Fe}/\text{H}]$, in line with what is found from high-fidelity boutique analyses of solar-twin stars (Bedell et al. 2018). The solar-twin analyses examined about 100 low- α stars in the solar neighborhood, making 66 abundance measurements from $R = 115,000$ data with $\text{S/N} > 100$, finding a mean intrinsic dispersion of ≈ 0.02 dex and a range of 0–0.06 dex. We find quite consistent results, with a mean intrinsic dispersion of ≈ 0.02 dex for our 17 elements, with a range of ≈ 0 –0.08 dex (with the exception of P which has an intrinsic dispersion of 0.12 dex, but which is not measured in the solar-twin sample).

For our second approach, we use the 600 APOGEE stars with asteroseismic ages $0.3 < \tau_{\text{age}} < 11.7$ Gyr, and metallicities $-0.5 < [\text{Fe}/\text{H}] < 0.45 \text{ dex}$, to model the relationship between the abundance vector and the labels of $(\tau_{\text{age}}, [\text{Fe}/\text{H}])$, across the parameter space of these 600 stars. We do this using a second-order polynomial model and calculate the intrinsic scatter around this model for each abundance feature, across age and $[\text{Fe}/\text{H}]$, given the errors on each of the abundances. This is an effective measure of the mean intrinsic dispersion

around this model, for each element, across metallicities from $-0.5 < [\text{Fe}/\text{H}] < 0.45$ dex and across ages $0.3 < \tau_{\text{age}} < 11.7$ Gyr (although distributed around a mean age of 3.3 Gyr). The intrinsic scatter from this model should represent an upper bound on the intrinsic dispersion at solar $[\text{Fe}/\text{H}]$ and our two measures of the intrinsic dispersion are consistent within the errors, given the intrinsic scatter is an upper limit, for all except one element, Cr, and only by a marginal amount. We report a mean intrinsic scatter across $(\tau_{\text{age}}, [\text{Fe}/\text{H}])$ of ≈ 0.02 dex for 16 elements, with a range of ≈ 0 –0.10 dex. The elements Si, Cr, S, Ca, Ni, Ti, and Mg all have scatters < 0.02 dex (i.e., below the mean of our sample). That the α -elements have very small scatter values is not surprising, given that we are conditioning on low α -enhancement. The element O is anomalous among the α -elements with relatively high scatter of 0.04 dex, although we note it is often reported as behaving dissimilar to the other α -elements (e.g., Buder et al. 2018). The elements Mn, K, and Al all have scatters of ≈ 0.02 dex, which is the mean scatter of the measurements. The elements P, Co, C, N, V, and O have the largest scatters of > 0.02 dex (i.e., above the mean scatter of our sample), and range from ≈ 0.04 to 0.10 dex. Some of these elements may also have minor systematic abundance temperature offsets that are limited across the parameter range here of the red clump (Holtzman et al. 2018). P has an anomalously high scatter of 0.10 dex and an intrinsic dispersion at solar $[\text{Fe}/\text{H}]$ of 0.12 dex. However, this element is also poorly measured, so any inaccuracy in the measurement uncertainty will more dramatically propagate into this determination. We have no reason, however, to suspect our error estimate on P is incorrect, or an anomalous underestimation, and this element is not available for comparison from the solar-twin study. Furthermore, there are very few studies for this element and the available literature indicates that massive stars produce three times more P than predicted, as no yields can even nearly produce the solar abundance of this element (Cirasuolo et al. 2014).

We find that our model of stellar ages and $[\text{Fe}/\text{H}]$ given abundances, and its corresponding scatter vector, can well generate the abundance vector for our large set of $\approx 15,000$ low- α red clump stars that are distributed across the disk. We find no preferentially poorer fit of the model as a function of Galactic location. Therefore, the model’s scatter determination is a measure of the intrinsic dispersion of the abundances around the age–abundance relation for stars distributed across a large Galactic radius, from 4 to 16 kpc, well beyond the solar neighborhood. For most elements, the abundances $[\text{X}/\text{Fe}]$ of Galactic disk stars can be predicted well (on average ≈ 0.02 dex), across a wide range of Galactocentric radii, given age and $[\text{Fe}/\text{H}]$.

The intrinsic dispersion around the age–abundance relation across age and $[\text{Fe}/\text{H}]$ also measures the prospect of distinguishing different sites of star formation at a given time, which may have different abundance signatures. This also therefore represents the precision with which each element must be measured to pursue chemical tagging of individual disk stars (Bland-Hawthorn et al. 2010), that is, to assign stars to a unique birth site from their vector of element abundances. However, there is also empirical element variation within star formation sites (open clusters), which is on the order of the intrinsic dispersion itself, for many elements (Bovy 2016; Ness et al. 2018; Liu et al. 2019). Therefore, the intrinsic dispersion does not necessarily represent the full variance *between* different individual star formation sites. It is possible to

recover simulated clusters from their abundances alone, which presumably links to their different ages, as traced by abundance differences (Price-Jones & Bovy 2019). A further test will be to see if simulated individual clusters formed at the same time can be recovered, with abundances drawn from their age relations as shown, for example, in Figures 3 and 8.

Our results, whereby most elements have small intrinsic dispersion measurements, do not diminish the value or importance of stellar abundances to reconstruct the formation and evolution history of the Milky Way’s disk. On the contrary, all of the trends we see reveal the richness of the abundances to indicate age and orbits, projected across a large spatial extent. Although we need extremely high precision to reconstruct clusters, we already have enough precision to quite clearly tackle Galactic evolution using the distribution of means (abundance, age, radius, orbit). The ground-breaking *Gaia* mission has shown that the Galactic disk displays numerous signatures of dynamical structure on local and global scales (e.g., Antoja et al. 2018; Trick et al. 2019; Bland-Hawthorn et al. 2019; Laporte et al. 2019; Mackereth et al. 2019b). The extent of the chemo-dynamical architecture across the disk is presumably linked to numbers of initial disk-forming clusters, cluster masses, and migratory events of stars across the disk over time. We see detailed chemo-dynamical structure in our data that presumably links to this. We find that a simple clustering of vertically confined stars of the disk using their abundances (at fixed $[\text{Fe}/\text{H}]$) separates groups of stars into populations of different mean ages and mean orbital actions (see also Beane et al. 2018). We also see the effect of radial migration, projected into the increasing J_ϕ separation of pairs of stars grouped by their abundance similarity at fixed $[\text{Fe}/\text{H}]$ over time.

The significant opportunity and challenge in tracing the formation of the disk, given these data, will be how to best combine stars in order to measure underlying chemo-dynamical structure to constrain its formation and evolution. The ≈ 0.02 dex (on average) precision on individual abundance measurements, that we find is needed to get at any inter-birth cluster element variation using individual stars, is an unattainable aim with current technology for most large surveys. This is particularly so over any span in parameter space in an evolutionary state (note here we use only the red clump to obtain our high-precision abundance measurements shown in Figure 1). However, by combining stars, a precision of $\ll 0.02$ dex on a population basis (i.e., the mean abundance value for many stars) will be straightforward to achieve.

Perhaps most importantly, the abundance measurements and abundance–age correlations we report here are strongly constraining with respect to stellar nucleosynthetic yields and chemical evolution models of the low- α disk across its radial extent. Examining the ratio of elements within nucleosynthetic families (and also between them), for groups of stars across the disk for single-age populations, is one powerful approach to isolate particular chemical enrichment events and access the corresponding signatures of the star formation history over time (e.g., Blancato et al. 2019; Weinberg et al. 2019). Correspondingly, there is a significant opportunity to use the measurements we have in hand to develop data-driven approaches to set nucleosynthetic yield tables and to build chemical evolution models from the actual data.

7. Conclusion

We have reported the rich and varied relationships between 18 chemical abundances of stars across the low- α disk in APOGEE and their ages as a function of $[\text{Fe}/\text{H}]$, for stars spanning $-0.5 < [\text{Fe}/\text{H}] < 0.45$ dex. The abundances, including those within single nucleosynthetic families (e.g., α -elements, light elements, iron-peak elements), have unique trends with age. We find that many individual abundances for stars contain only marginal additional information beyond being indicators of age, at fixed $[\text{Fe}/\text{H}]$. Therefore, for a good fraction of the low- α disk, age and $[\text{Fe}/\text{H}]$ alone can predict the other abundances measured by APOGEE to high precision. Furthermore, this suggests that a star's $[\text{Fe}/\text{H}]$, age, and membership in the high- or low- α sequence are all (or nearly all) that are needed to determine a star's birth radius (Frankel et al. 2018). Subsequently, for the low- α sequence, stars with the same $[\text{Fe}/\text{H}]$ and age most likely were born at the same radial annulus.

Element abundance precisions of <0.1 dex are a realistic and achievable goal for multi-million star surveys. In the case of APOGEE quality spectra, this precision is achieved fairly trivially, at $S/N > 30$ for most elements (e.g., Ness 2018; Ting et al. 2018). A precision of <0.02 dex, which is the mean intrinsic dispersion we measure around the age–abundance trends, for our set of elements from APOGEE spectra is, however, not an achievable goal for large surveys, given current stellar models and infrastructure. In the regime of large stellar samples, the lesser and more achievable precision of <0.1 dex, particularly across a wider parameter space than that considered here with our red clump sample, still affords tremendous opportunity to reconstruct the Milky Way's formation. From the many millions of densely sampled stars from future surveys, such as Sloan V's Milky Way Mapper (Kollmeier et al. 2017), WEAVE (Bonifacio et al. 2016), 4MOST (de Jong et al. 2019), GALAH (De Silva et al. 2015), PFS (Tamura et al. 2018), LAMOST (Newberg et al. 2012), and MOONS (Cirasuolo et al. 2014), we have tremendous opportunity to work out how to optimally combine stars (e.g., spatially, dynamically, temporally). In doing so, we can expect to reveal the empirical characteristics of the chemo-dynamical structure in the Milky Way across a vast range of scales (e.g., Xiang et al. 2018; Kamdar et al. 2019).

We thank the anonymous referee for their close reading of the manuscript and constructive feedback.

We thank the Flatiron astronomical data group for helpful discussions. (<https://astrodata.nyc>).

M.N. is in part supported by a Sloan Foundation Fellowship. This work was performed in part during the *Gaia*19 workshop at the Kavli Institute for Theoretical Physics at the University of California, Santa Barbara.

K.V.J.'s contributions were supported in part by the National Science Foundation under grants NSF PHY-1748958 and NSF AST-1715582. Her work was performed in part during the *Gaia*19 workshop and the 2019 Santa Barbara *Gaia* Sprint (also supported by the Heising-Simons Foundation), both hosted by the Kavli Institute for Theoretical Physics at the University of California, Santa Barbara.

K.B. is supported by the NSF Graduate Research Fellowship under grant number DGE 16-44869.

H.W.R. gratefully acknowledges insightful conversations with Yuan-Sen Ting and Neige Frankel on the relations between birth-orbits, ages, and abundances of stars in the Galactic disk.

K.H. is partially support by a TDA/Scialog Grant from the Research Cooperation “A *Gaia*-Enabled View of Chemical Homogeneity.” K.H. is also partially support by the Center for Astrophysical Plasma Properties (CAPP) sponsored by the Department of Energy.

A.B. was supported in part by the Roy & Diana Vagelos Program in the Molecular Life Sciences and the Roy & Diana Vagelos Challenge Award.

ORCID iDs

M. K. Ness  <https://orcid.org/0000-0001-5082-6693>

H-W. Rix  <https://orcid.org/0000-0003-4996-9069>

A. Beane  <https://orcid.org/0000-0002-8658-1453>

References

- Adibekyan, V. Z., Figueira, P., Santos, N. C., et al. 2013, *A&A*, **554**, A44
- Antoja, T., Helmi, A., Romero-Gómez, M., et al. 2018, *Natur*, **561**, 360
- Armillotta, L., Krumholz, M. R., & Fujimoto, Y. 2018a, *MNRAS*, **481**, 5000
- Armillotta, L., Krumholz, M. R., & Fujimoto, Y. 2018b, *MNRAS*, **481**, 5000
- Beane, A., Ness, M. K., & Bedell, M. 2018, *ApJ*, **867**, 31
- Bedell, M., Bean, J. L., Melendez, J., et al. 2018, *ApJ*, **865**, 68
- Bensby, T., Alves-Brito, A., Oey, M. S., Yong, D., & Meléndez, J. 2012, in ASP Conf. Ser. 458, *Galactic Archaeology: Near-Field Cosmology and the Formation of the Milky Way*, ed. W. Aoki et al. (San Francisco, CA: ASP), 171
- Bensby, T., Feltzing, S., Gould, A., et al. 2017, *A&A*, **605**, A89
- Binney, J., & Tremaine, S. 2008, *Galactic Dynamics* (2nd ed.; Princeton, NJ: Princeton Univ. Press)
- Blancato, K., Ness, M., Johnston, K. V., Rybizki, J., & Bedell, M. 2019, *ApJ*, **883**, 34
- Bland-Hawthorn, J., & Gerhard, O. 2016, *ARA&A*, **54**, 529
- Bland-Hawthorn, J., Krumholz, M. R., & Freeman, K. 2010, *ApJ*, **713**, 166
- Bland-Hawthorn, J., Sharma, S., Tepper-García, T., et al. 2019, *MNRAS*, **486**, 1167
- Bonifacio, P., Dalton, G., Trager, S., et al. 2016, in SF2A-2016: Proc. Annual Meeting of the French Society of Astronomy and Astrophysics, ed. C. Reylé et al. (Paris: SAF), 267
- Bovy, J. 2016, *ApJ*, **817**, 49
- Bovy, J., Leung, H. W., Hunt, J. A. S., et al. 2019, arXiv:1905.11404
- Bovy, J., Rix, H.-W., Liu, C., et al. 2012, *ApJ*, **753**, 148
- Buder, S., Asplund, M., Duong, L., et al. 2018, *MNRAS*, **478**, 4513
- Cescutti, G., Matteucci, F., Caffau, E., & François, P. 2012, *A&A*, **540**, A33
- Cirasuolo, M., Afonso, J., Carollo, M., et al. 2014, *Proc. SPIE*, **9147**, 91470N
- Clarke, A. J., Debattista, V. P., Nidever, D. L., et al. 2019, *MNRAS*, **484**, 3476
- de Jong, R. S., Agertz, O., Agudo Berbel, A., et al. 2019, *Msngr*, **175**, 3
- De Silva, G. M., Freeman, K. C., Bland-Hawthorn, J., et al. 2015, *MNRAS*, **449**, 2604
- Duong, L., Freeman, K. C., Asplund, M., et al. 2018, *MNRAS*, **476**, 5216
- 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
- Gadotti, D. A., Sánchez-Blázquez, P., Falcón-Barroso, J., et al. 2019, *MNRAS*, **482**, 506
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *yCat*, **616**, A1
- Gandhi, S. S., & Ness, M. K. 2019, *ApJ*, **880**, 134
- García Pérez, A. E., Allende Prieto, C., Holtzman, J. A., et al. 2015, arXiv:1510.07635
- Genovali, K., Lemasle, B., Bono, G., et al. 2014, *A&A*, **566**, A37
- Gilmore, G., Randich, S., Asplund, M., et al. 2012, *Msngr*, **147**, 25
- Hawkins, K., Leistedt, B., Bovy, J., & Hogg, D. W. 2017, *MNRAS*, **471**, 722
- Hawkins, K., Ting, Y.-S., & Walter-Rix, H. 2018, *ApJ*, **853**, 20
- Hayden, M. R., Bovy, J., Holtzman, J. A., et al. 2015, *ApJ*, **808**, 132
- Haywood, M., Di Matteo, P., Lehnert, M. D., Katz, D., & Gómez, A. 2013, *A&A*, **560**, A109
- Ho, A. Y. Q., Rix, H.-W., Ness, M. K., et al. 2017, *ApJ*, **841**, 40
- Hogg, D. W., Casey, A. R., Ness, M., et al. 2016, *ApJ*, **833**, 262
- Holtzman, J. A., Hasselquist, S., Shetrone, M., et al. 2018, *AJ*, **156**, 125
- Holtzman, J. A., Shetrone, M., Johnson, J. A., et al. 2015, *AJ*, **150**, 148

- Kamdar, H., Conroy, C., Ting, Y.-S., et al. 2019, arXiv:1902.10719
- Kollmeier, J. A., Zasowski, G., Rix, H.-W., et al. 2017, arXiv:1711.03234
- Krumholz, M. R., McKee, C. F., & Bland-Hawthorn, J. 2019, *ARA&A*, **57**, 227
- Laporte, C. F. P., Minchev, I., Johnston, K. V., & Gómez, F. A. 2019, *MNRAS*, **485**, 3134
- Leung, H. W., & Bovy, J. 2019, *MNRAS*, in press
- Liu, F., Asplund, M., Yong, D., et al. 2019, *A&A*, **627**, A117
- Loebman, S. R., Debattista, V. P., Nidever, D. L., et al. 2016, *ApJL*, **818**, L6
- Luck, R. E., & Lambert, D. L. 2011, *AJ*, **142**, 136
- Mackereth, J. T., Bovy, J., Leung, H. W., et al. 2019a, *MNRAS*, **489**, 176
- Mackereth, J. T., Schiavon, R. P., Pfeffer, J., et al. 2019b, *MNRAS*, **482**, 3426
- Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, *AJ*, **154**, 94
- Martig, M., Fouesneau, M., Rix, H.-W., et al. 2016, *MNRAS*, **456**, 3655
- Masseron, T., & Gilmore, G. 2015, *MNRAS*, **453**, 1855
- McMillan, P. J. 2017, *MNRAS*, **465**, 76
- Minchev, I., Anders, F., Recio-Blanco, A., et al. 2018, *MNRAS*, **481**, 1645
- Minchev, I., Famaey, B., Combes, F., et al. 2011, *A&A*, **527**, A147
- Mollá, M., Díaz, Á. I., Cavichia, O., et al. 2019, *MNRAS*, **482**, 3071
- Ness, M. 2018, *PASA*, **35**, e003
- Ness, M., Freeman, K., Athanassoula, E., et al. 2013, *MNRAS*, **430**, 836
- Ness, M., Hogg, D. W., Rix, H.-W., et al. 2016, *ApJ*, **823**, 114
- Ness, M., Hogg, D. W., Rix, H.-W., Ho, A. Y. Q., & Zasowski, G. 2015, *ApJ*, **808**, 16
- Ness, M., Rix, H.-W., Hogg, D. W., et al. 2018, *ApJ*, **853**, 198
- Newberg, H. J., Carlin, J. L., Chen, L., et al. 2012, in ASP Conf. Ser. 458, Galactic Archaeology: Near-Field Cosmology and the Formation of the Milky Way, ed. W. Aoki et al. (San Francisco, CA: ASP), 405
- Nidever, D. L., Bovy, J., Bird, J. C., et al. 2014, *ApJ*, **796**, 38
- Pinsonneault, M. H., Elsworth, Y. P., Tayar, J., et al. 2018, *ApJS*, **239**, 32
- Price-Jones, N., & Bovy, J. 2019, *MNRAS*, **487**, 871
- Rix, H.-W., & Bovy, J. 2013, *A&ARv*, **21**, 61
- Roškar, R., Debattista, V. P., Quinn, T. R., Stinson, G. S., & Wadsley, J. 2008, *ApJL*, **684**, L79
- Sanderson, R. E., Wetzel, A., Loebman, S., et al. 2018, arXiv:1806.10564
- Schlesinger, K. J., Johnson, J. A., Rockosi, C. M., et al. 2014, *ApJ*, **791**, 112
- Sellwood, J. A., & Binney, J. J. 2002, *MNRAS*, **336**, 785
- Silva Aguirre, V., Bojsen-Hansen, M., Slumstrup, D., et al. 2018, *MNRAS*, **475**, 5487
- Souto, D., Allende Prieto, C., Cunha, K., et al. 2019, *ApJ*, **874**, 97
- Steinmetz, M., Zwitter, T., Siebert, A., et al. 2006, *AJ*, **132**, 1645
- Tamura, N., Takato, N., Shimono, A., et al. 2018, *Proc. SPIE*, **10702**, 107021C
- Ting, Y.-S., Conroy, C., & Goodman, A. 2015, *ApJ*, **807**, 104
- Ting, Y.-S., Freeman, K. C., Kobayashi, C., De Silva, G. M., & Bland-Hawthorn, J. 2012, *MNRAS*, **421**, 1231
- Ting, Y.-S., Hawkins, K., & Rix, H.-W. 2018, *ApJL*, **858**, L7
- Trick, W. H., Coronado, J., & Rix, H.-W. 2019, *MNRAS*, **484**, 3291
- Valentini, M., Borgani, S., Bressan, A., et al. 2019, *MNRAS*, **485**, 1384
- Weinberg, D. H., Holtzman, J. A., Hasselquist, S., et al. 2019, *ApJ*, **874**, 102
- Xiang, M., Shi, J., Liu, X., et al. 2018, *ApJS*, **237**, 33
- Yanny, B., Rockosi, C., Newberg, H. J., et al. 2009, *AJ*, **137**, 4377
- Zasowski, G., Johnson, J. A., Frinchaboy, P. M., et al. 2013, *AJ*, **146**, 81